EUROSEQUENCE bv

The Art of Protein Sequence Analysis

PRODUCT INFORMATION N-TERMINAL PROTEIN SEQUENCE ANALYSIS

Method and General Remarks

The sequence analysis is performed by Edman degradation [1,2] with an automated sequenator (Model 494 Procise Applied Biosystems)[3] using protocols, reagents, chemicals and materials from Applied Biosystems (Warrington, U.K. and Foster City, CA, U.S.A.). Step-wise released phenylthiohydantoin amino acids are identified with an on-line RP-HPLC on the basis of their elution times ('Chromatogram Reports'). For this reason, one or more calibrations are performed prior to each sequence analysis. The chromatograms have been checked by us, and, where necessary, corrections, additions and/or remarks have been made; symbols written on the chromatograms only serve as a memory aid for us during our interpretation of the results and are of no further significance.

Some parameters to judge the sequence analysis process are the 'Initial Yield' and 'Repetitive Yield'. The **'Initial Yield'** is the quantity of sequenceable protein material; it is **usually about 50%**, but may vary between **20%-80%** of the quantity of protein that is used for analysis, but lower for proteins blotted onto PVDF **(5%-50%)**. Because of the spread in values for the Initial Yield, sequence analysis data cannot be used as an accurate determination of the protein quantity. However, within the same analysis, relative quantities can be reliably be determined, based on the ratio of the Initial Yield values. The 'Repetitive Yield' is the average percentage of recovery of PTH-amino acid per Edman degradation cycle. For proteins this value is usually about 94%; for short peptides, about 86%.

Edman degradation of N-alkylated proteins may result in preview [4]. Since certain procedures for reduction and alkylation of cysteine/cystine may lead to partial N-alkylation, about 10% n-1 mers may be found in samples, which have been subjected to Edman degradation after gas-phase pyridylethylation.

Explanation of the interpretation report, and raw data delivered

The scientific interpretation of the results is summarized in a separate report signed by the responsible scientist/specialist. On this report you will find our product code and order number, which are also printed at the bottom of the corresponding chromatograms (possible text after the product code and order number are for internal use). The page numbering of the raw data (chromatograms) is such that both the first page and the last page are not delivered (first is only a blank cycle; last page is an internally used page containing a computer prediction of the likely sequence).

Amino acids in parentheses are so-called tentative assignments; in case of more than one tentative assignment per position, the amino acids are separated by a slash [/]. At positions where no amino acid is identified (indicated by 'Xaa' or '???'), or where questionable, tentative assignments are made, the following amino acids are also possible: - Cysteine/Cystine (if not chemically modified to a stable derivative); - Tryptophan (which may have been destroyed by oxidation, especially after electroblotting onto PVDF); - amino acids with carbohydrate attachments (Serine, Threonine or Asparagine if consensus sequence Asn-Xaa-Thr/Ser and eukaryotic expression system present) and; - other, less frequent, post-translationally or chemically modified amino acids.

We would like to emphasize that the amino acid sequence of your sample is based on the best possible interpretation of the chromatograms obtained. However, contamination of a sample especially with protein material, originating either from other protein material or from (extensive) internal chain cleavage, as well as chemical modifications and/or cleavages of the protein material prior to sequence analysis, may lead to misinterpretations. Therefore, information about the sample history is also essential for obtaining the best possible interpretation.

Furthermore, in case the 'Initial Yield' (see above) is much lower then expected, the possibility exists that the protein in your sample is N-terminally blocked and that possible interpreted signals may originate from a contaminant in your sample or from an internal chain cleavage product.

References

[1] Edman, P. (1956) Acta Chem. Scand. **10**, 761-768.

[2] Ilse, D. & Edman, P. (1963) Aust. J. Chem. **16**, 411-416.

[3] Hewick, R.M., Hunkapiller, M.W., Hood, L.E. & Dreyer, W.J. (1981) J. Biol. Chem. **256**, 7990-7997.

[4] Hempell, J., Nilsson, K., Larsson, K. and Jörnvall, H. (1986) *FEBS Lett.* **194**, 333-337.