

Applications of N-terminal Edman Sequencing for Proteins and Peptides

By Ejvind Mortz, Thanh Ha Nguyen, and Thomas Kofoed. Alphalyse A/S, Unsbjergvej 4, Odense, Denmark.

Email correspondence: info@alphalyse.com

Analytical characterization of purified proteins, in particular pharmaceutical proteins, requires detailed analysis of the protein terminals and confirmation of the expected sequence start and end. N-terminal protein sequencing by Edman degradation has been the method of choice for many years, but has gradually been replaced by mass spectrometric peptide mapping and MS/MS sequencing. Alphalyse provides N-terminal sequence analysis by both Edman and Mass spectrometry of therapeutic proteins, monoclonal antibodies and protein vaccines. In our view, Edman sequencing and mass spectrometric analysis provide complementary information. In this application note, we highlight the benefits and discuss applications of N-terminal Edman sequencing for protein characterization.

INTRODUCTION

The sequence of amino acids in a protein or peptide can be analyzed from the N-terminal by Edman sequencing. The Edman degradation chemistry was developed more than 60 years ago by Pehr Edman (1, 2). Since then faster and less expensive sequencing methods have been developed, including top-down mass spectrometric sequencing (3) and DNA sequencing methods. Still N-terminal Edman sequencing has advantages for protein analysis that cannot readily be obtained by other analysis methods (4).

Advantages of Edman sequencing:

The Edman chemistry cleaves off the N-terminal amino acid of the protein. The first Edman sequencing cycle thus identifies the exact N-terminal amino acid.

The released amino acids are identified and quantified by chromatography. Thereby amino acids with identical molecular weights can be identified, i.e. I and L both have a mass of 113 Da, and Q/K both have a mass of 128 Da, but different HPLC retention time.

Edman sequencing can be performed on PVDF blots from 1D and 2D gels. This enables N-terminal sequencing of proteins in mixtures.



APPLICATIONS

Edman sequencing is used in the following applications:

Verification of correct protein expression - confirmation of the correct insertion and N-terminal expression in recombinant proteins. Quality control of pharmaceutical proteins according to the ICH Q6B guideline for analysis of the protein terminals

Cell line development and fermentation - compare different cell lines for expression and processing of the leading methionine residue and signal sequence.

Protein degradation analysis – identify the new N-terminal and proteolytic cleavage site in the protein fragments.

De-novo sequencing – for analysis of novel proteins and peptides where sequence databases are not available for MS/MS database searching. And for de-novo sequencing of peptides containing I/L and Q/K.

ANALYSIS PRINCIPLES

Proteins purified by liquid chromatography are analyzed by applying the protein in-solution onto a TFA filter and then loaded onto the Edman sequencing instrument.

Proteins in mixtures are first separated by 1D SDS PAGE or 2D PAGE and then blotted onto a PVDF membrane. The proteins are detected by Coomassie blue, Amido black or Poncau S staining and the proteins of interest cut out and the PVDF membrane piece loaded onto the Edman sequencer (Fig 1).

The Edman sequencing reaction is a cyclic procedure where the N-terminal amino acid residue is cleaved off and identified by chromatography (Fig 2). Then the next amino acid is cleaved off and identified and so forth. Thereby, the amino acid sequence is read from the N-terminus. The coupling and cleavage reaction efficiency is not 100% and in practice limits the number of residues that can be sequenced to 5-50 amino acids.

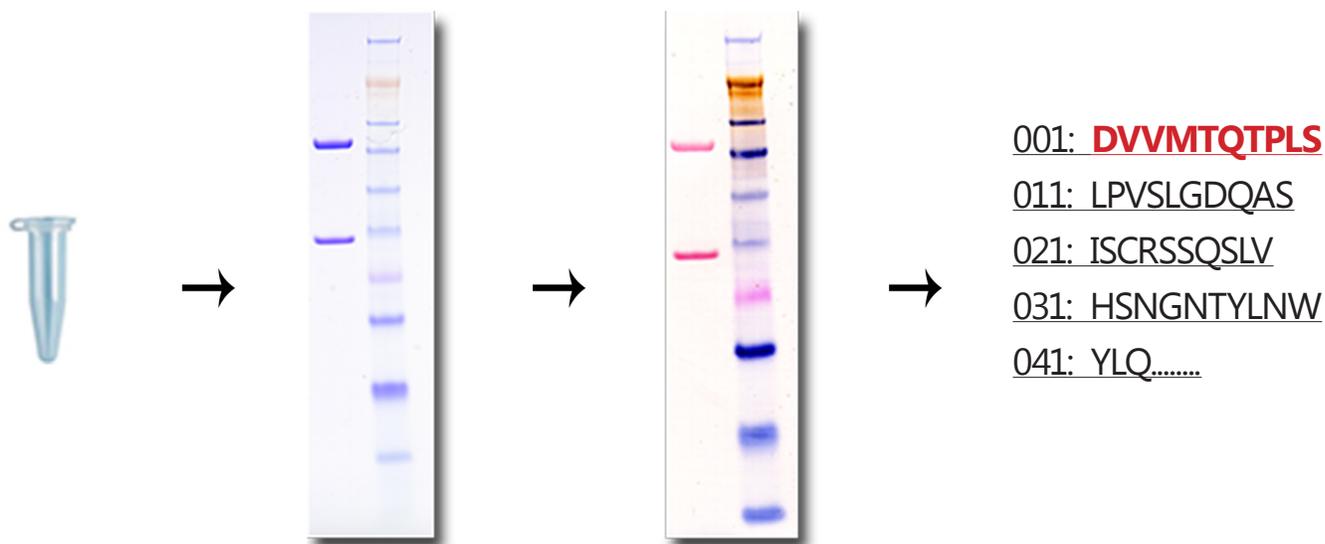


FIGURE 1. FLOW-CHART FOR N-TERMINAL EDMAN SEQUENCING OF ANTIBODY BANDS FROM 1D SDS PAGE GEL

A) The heavy and light chains of an antibody are separated by 1D SDS PAGE according to protein molecular weights. **B)** The HC and LC are transferred to a PVDF membrane by electroblotting, and the proteins detected by Coomassie, Amido black or Ponceau S staining. For Edman sequencing, the protein of interest should be clearly stained and contain 10-100 picomoles. The protein should be well-separated from other protein bands to get clear sequence read. **C)** The PVDF bands are cut out with a scalpel and loaded onto the ABI 494 Precise Protein Sequencing System for automated N-terminal Edman sequencing providing 5-20 residues for both the HC and the LC (sequence for LC shown in figure).

MATERIALS AND METHODS

Gel electrophoresis

Protein samples for 1D-SDS PAGE were dissolved in LDS sample buffer with DTT and heated at 70°C for 10 min. The samples were loaded onto a 12% Bis-Tris gel and separated using a MOPS running buffer at 200V for 1 hr. (NuPAGE gel system, Invitrogen).

Electroblotting

Proteins were electroblotted for 1 hr onto a PVDF membrane

(Hybond-P, 0.45 μm, GE Healthcare) using the Xcell Surelock Mini-Cell system (Invitrogen), and the NuPAGE transfer buffer with 10% methanol.

Ponceau S Staining

The proteins were detected by staining the PVDF membrane with a freshly made solution of 1% Ponceau S, 1% acetic acid for 2 min. The PVDF membrane was destained in distilled water until the protein bands were visible.

Coomassie Blue and Amido Black Staining

Alternatively, PVDF membranes for Edman sequencing can be stained by Coomassie or Amido Black.

For Coomassie stain, use freshly prepared 0.1% Coomassie Blue R-250 in 40% methanol/1% acetic acid for 1 min. Destain in 50% methanol until bands are visible and background clear.

For Amido Black stain, use 0.1% Amido Black dissolved

in 30% methanol, 10% acetic acid. Destain by several washes in distilled water (5).

Edman sequencing

Edman sequencing was performed on a Procise Protein Sequencing System (Applied Biosystems), equipped with a 494 Procise sequencer, the ABI model 140C HPLC pump and a variable wavelength UV detector. Sequence data are analyzed using the Sequence Pro software.

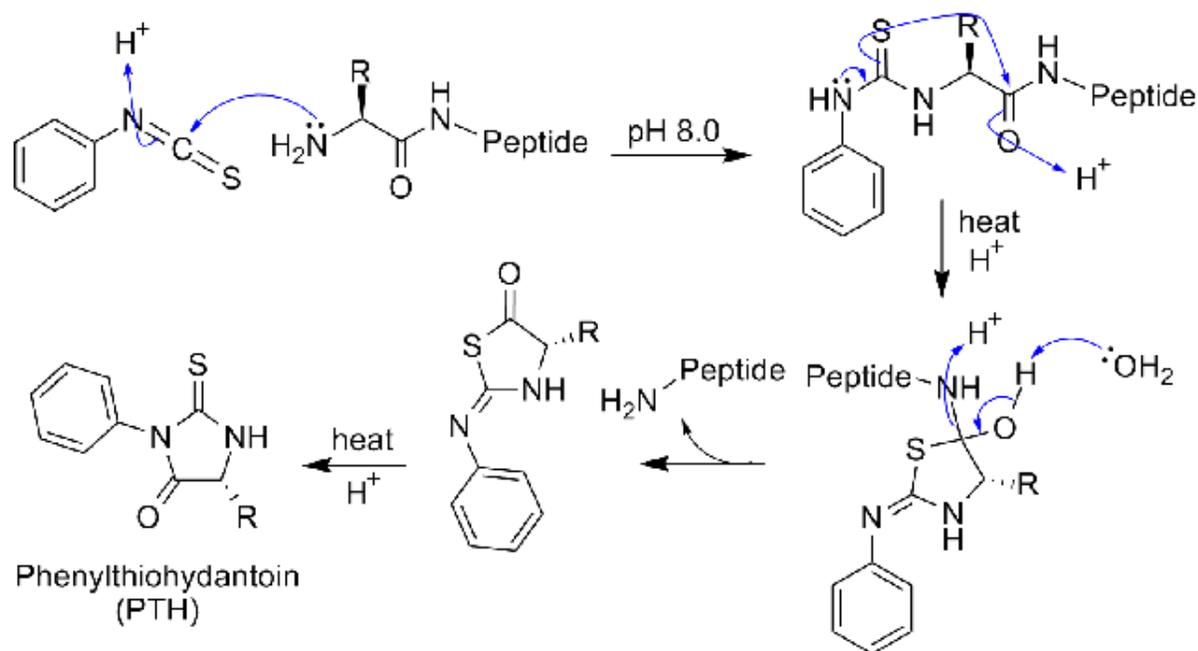


FIGURE 2. N-TERMINAL EDMAN SEQUENCING CHEMISTRY

There are 3 steps in the cyclic Edman degradation procedure. In step 1, the PITC reagent (phenylisothiocyanate) is coupled to the N-terminal amino group under alkaline conditions. In step 2, the N-terminal residue is cleaved in acidic media. In step 3, the PITC coupled residue is transferred to a flask, converted to a more stable PTH-residue and transferred to an on-line HPLC and identified by chromatography (6).

RESULTS

Verification of Protein Expression

In the expression of recombinant proteins, in the development of high yield cell lines, and in the optimization of fermentation conditions, it is frequently observed that the N-terminal is wrong. The correct and complete removal of signal sequences and excision of the N-terminal translation initiator Methionine is important for the function and stability of recombinant proteins (7). According to the N-end rule, the protein half-life is determined and altered by the N-terminal amino acid residue (8). Therefore it is important to analyze and verify the expected N-terminal of the recombinant protein. For pharmaceutical proteins it is a requirement in the

ICH Q6B Guideline that the protein terminals should be analyzed.

As example, in a cell line development project, individual cell clones were analyzed by running the whole cell lysates on separate lanes on 1D SDS PAGE (Figure 3). The proteins were electroblotted onto PVDF and the drug substance bands analyzed by Edman sequencing. One clone showed exclusively the correct N-terminal amino acid Glycine (Fig 3A), whereas one clone showed a mixture of the correct Glycine and the incorrect N-terminal Methionine (Fig 3B).

The relative ratio of the 2 N-terminal forms can be determined from the relative peak areas of the chromatographic peaks for Glycine and Methionine.

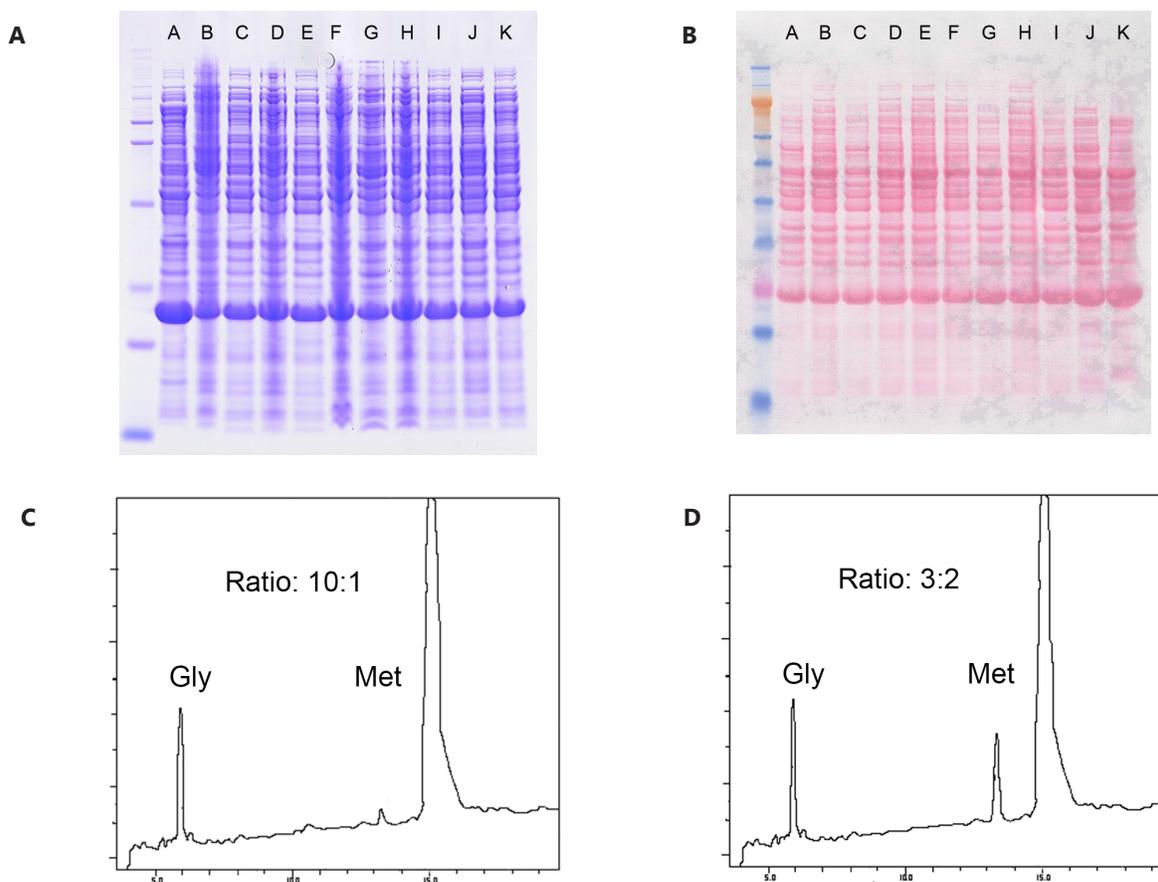


FIGURE 3. MIXED CORRECT AND INCORRECT PROCESSING OF N-TERMINAL METHIONINE

Whole cell lysates of individual cell lines (A to K) analyzed by 1D SDS PAGE (Fig A). Drug substance bands from the PVDF blot (Fig B) were analyzed by N-terminal Edman sequencing. A correct N-terminal Glycine residue is observed for cell line A (Ratio 10:1) (Fig C). Cell line B shows a mixture of N-terminal Methionine and Glycine (Ratio 3:2) (Fig D).

Protein degradation analysis

In the studies of protein degradation it is important to identify the generated peptide fragments and the N-terminal cleavage sites. This is for example used in stability assays of pharmaceutical proteins, and in the determination of proteolytic cleavage specificity.

Again, the stability samples from different time points can be run into individual lanes in a 1D SDS gel. The fragments are blotted onto PVDF, stained for visual comparison, and the individual cleavage fragments identified by Edman sequencing (data not shown).

Beta-Lactoglobulin quality control sample

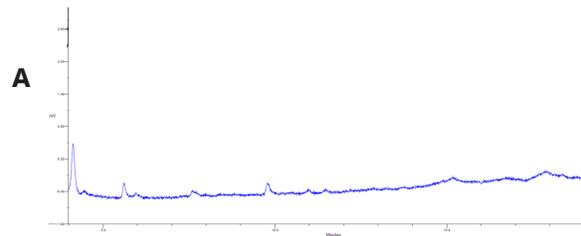
To illustrate important parameters of Edman sequencing, a 20 pmol beta-lactoglobulin sample (BLG) was analyzed. The BLG sample is used as a quality control standard to confirm the instrument performance and Edman chemistry.

In the first cycle, a blank HPLC run is performed without any sequencing reaction (Fig 4A).

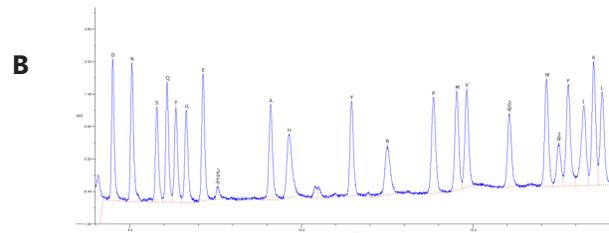
In the second cycle, a standard mixture with 10 pmol of each of the 19 PTH derivatized amino acids is injected on the HPLC to check aa retention time and signal response (Fig 4B).

In the third HPLC cycle, the N-terminal amino acid(s) from the first sequencing reaction is injected. From the subsequent 19 cycles, a sequence of 20 amino acids is read and checked with the known BLG sequence. The sequence length, signal strength and data quality is checked.

Blank run



PTH-aa standard



Residue 3

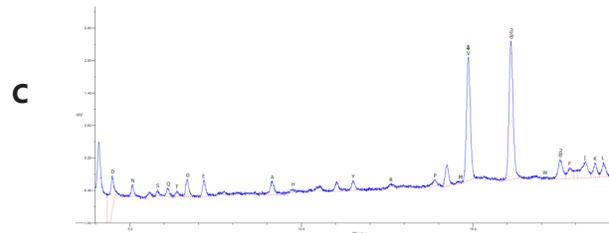


FIGURE 4: SELECTED CYCLES FROM BETA-LACTOGLOBULIN ANALYSIS

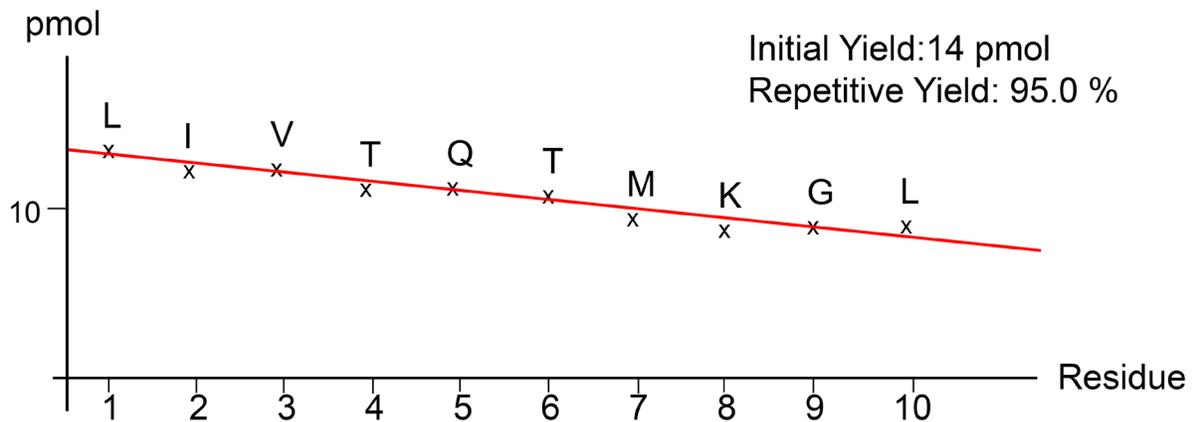


FIGURE 5: BETA-LACTOGLOBULIN RUN REPETITIVE YIELD AND INITIAL YIELD GRAPH

TECHNIQUE	CYCLIC EDMAN DEGRADATION
N-terminal sequence	Yes
C-terminal sequence	No
Sequence lengths	5-50 residues
Sample requirements	<ul style="list-style-type: none"> Purified protein SDS PAGE purified protein on PVDF membrane
Protein amounts	<ul style="list-style-type: none"> More than 10 picomoles, or 2-5 microgram protein On a PVDF blot, the protein band should be clearly stained with Ponceau S, Coomassie blue or Amido black.
Special features	<ul style="list-style-type: none"> Requires free amino group on the N-terminal amino acid. N-terminally blocked proteins cannot be sequenced, eg. N-terminals with acetylation or pyroglutamic acid The 19 known amino acids are identified by HPLC retention time. Modified residues are not identified, eg. glycosylated or phosphorylated residues, cys-cys residues, etc. Glycine is often observed in the first cycle due to the presence in many blotting buffers
Disadvantage	<ul style="list-style-type: none"> Low throughput analysis with cycle time 40 mins per residue Expensive chemicals results in high cost per residue Does not work for N-terminally blocked proteins
Advantage	<ul style="list-style-type: none"> <u>Can be combined with 1D and 2D electrophoresis for analysis of proteins in mixtures.</u> <u>True sequencing of individual aa makes <i>de-novo</i> sequencing possible.</u>

TABLE 1: FEATURES OF EDMAN SEQUENCING

CONCLUSION

In this application note we have demonstrated several unique features of N-terminal Edman sequencing that makes it very useful for specific applications.

The applications include:

- De-novo sequencing of novel amino acid sequences, and peptides containing I/L or Q/K residues with identical Mw's.

- Identification of protein degradation fragments and proteolytic cleavage sites
- Quality Control and analysis of protein N-terminals according to ICH Q6B Guideline
- Verification of correct excision of N-terminal Methionine and leading signal sequences in choice of cell line and fermentation conditions
- Analysis of proteins from protein mixtures separated by 1D SDS or 2D PAGE.

REFERENCES

- Edman P., Högfeldt E., Sillén LG., Kinell P. "Method for determination of the amino acid sequence in peptides". Acta Chem. Scand. 4: 283–293, 1950.
- Niall HD. "Automated Edman degradation: the protein sequenator". Methods in Enzymology 27: 942–1010, 1973.
- Mortz E, Nguyen TH., Crawford J., "N- and C-terminal sequencing by MALDI ISD". Alphalyse Application note #200902.
- Speicher KD, Gorman N., and Speicher DW. "N-Terminal Sequence Analysis of Proteins and Peptides" Curr. Protoc. Protein Sci., PMC August 5 2010.
- Yonan CR., Duong PT., Chang FN. "High-efficiency staining of proteins on different blot membranes". Analytical Biochemistry 338, 159–161, 2005.
- http://en.wikipedia.org/wiki/Edman_degradation
- Gigliione C, Boularot A, Meinnel T. "Protein N-terminal methionine excision." Cell Mol Life Sci. Jun; 61(12):1455-74, 2004.
- Varshavsky A.. "The N-end rule pathway of protein degradation". Genes to Cells 2 (1): 13–28, 1997.