ALTA BIOSCIENCE LTD

PROTEIN SEQUENCING

Technical Document

AltaBioscience offers N-terminal protein sequencing to ISO 17025:2017 standard as defined by the United Kingdom Accreditation Service (UKAS).

The structure and function of a protein is determined by the composition and arrangement of its amino acids. Protein sequencing is the process of identifying the order of these amino acid arrangements and has become an important tool for protein and peptide identification for de novo sequencing of novel proteins, to probe design for molecular cloning and analysis of samples from GMP processes.

Why use AltaBioscience?

AltaBioscience is accredited to ISO 17025:2017 for N-Terminal Protein Sequencing. N-Terminal Sequencing, otherwise known as Edman degradation, is recognised as the definitive sequencing technique and is accepted as being the most reliable, robust and accurate method, providing an unambiguous sequence read.

In addition to our quality guarantee, you will receive the following:

- Easy to understand reports
- Fast turnaround or results with optional express service
- Sensitivity at picomole levels
- Expertise and advice from our highly trained staff

Applications

- Determination of the N-terminal amino acids of a protein or peptide
- Confirmation of the correct translation of a recombinant protein
- Confirmation of the sequence of a recombinant protein
- Quality control check for synthetic peptides
- Determination of complete protein sequence in combination with additional methods
- Drug discovery for de novo sequencing of new novel proteins
- Identification and characterisation of proteins and peptides
- Analysis of samples from GMP processes
- Probe design for molecular cloning

Advancing science through analytical chemistry and synthesis services in the life and food science sectors worldwide

Methodology

Automated N-terminal sequence analysis involves a series of chemical reactions that derivatise and remove one amino acid at a time from the N-terminal of purified peptides or intact proteins.

Sequences of 40 amino acids or more are possible with this technique.

The first stage in the process involves immobilisation of the protein, either by blotting onto a PVDF membrane or by adsorption onto a biobrene treated glass fibre filter. An unmodified α-amino group is required at the N-terminal end of the molecule to be able to carry out this cyclic process. The free N-terminal amino acid is reacted with the Edman reagent, phenylisothiocyanate (PITC), to form a PTC derivative. The amino acid derivative is removed by acid cleavage as its phenylthiohydantoin (PTH) derivative and a new α-amino group on the next amino acid is subsequently available to react with PITC and continue the next cycle. The coupling and cleavage process is repeated for as many times as is required.

Each cleaved amino acid PTH derivative is separated using HPLC and detected using a UV detector. The amino acids are identified by comparison to a standard mixture.

Features

- Sequencing as low as 5-10 picomole of sample
- Identification of a purified protein/peptide
- Semi-quantitative analysis
- Isobaric (Ile/Leu) and near isobaric (Gln/Lys) residues are reliably determined
- Turnaround of analysis and report within 5-7 working days,
- Express service for urgent results available

- Sequence reads of up to 40 amino acids in one operation (sample dependent)
- Compatible with SDS-PAGE following blotting to PVDF membrane
- GMP process compatible
- Analysis does not require a sequence database

Length of read

It is possible to obtain reads of over 40 amino acids, sequence and sample dependent

- 5 - 6 amino acids are sufficient to identify a known protein
- 8 - 10 amino acids are required to search a database for a unique match
- 12 - 15 amino acids are needed to design an oligo primer for DNA sequencing

Sample Types

Proteins on stained PVDF blots

Membranes with the bands of interest clearly marked or alternatively the bands of interest excised and sent in a micro-centrifuge tube. Photographs or scans of the gel, or whole blot, can be sent with the sample to help determine how much signal is expected.

Proteins in PAGE gel slices

The gel slices of interest cut out and sent in a microcentrifuge tube. Several bands of the same protein can be sent to enhance concentration.

Proteins in solution

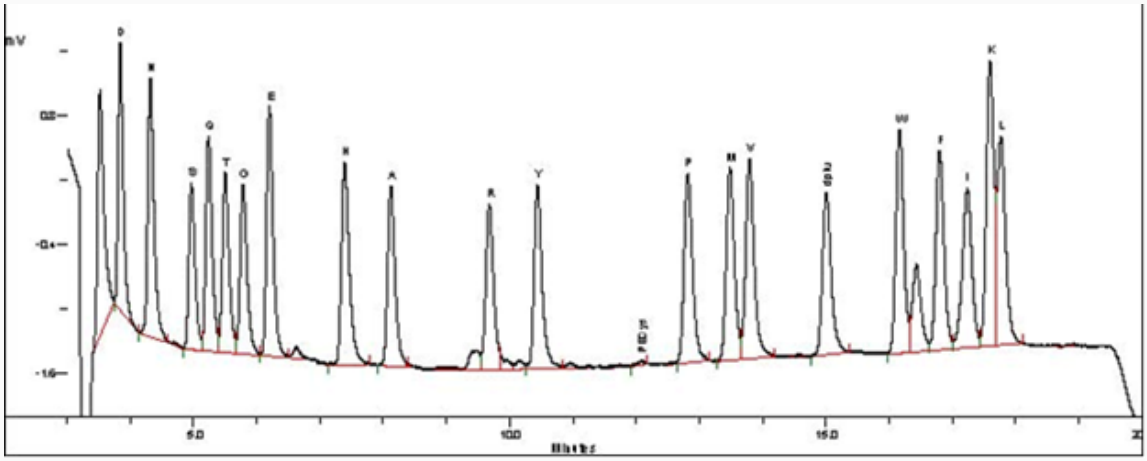
For samples in solution, it would be helpful to provide details of the approximate concentration. It is not advisable to dry samples of protein, as once dry they can be very difficult to re-dissolve.

Peptides in solution or dried

Peptides from HPLC

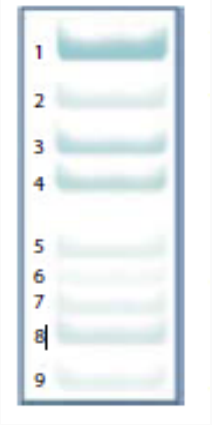
Peptide libraries on beads

Sequencing Standard Profile



Sample Amounts

The illustration below shows a typical single lane of a gel with different amounts in each band.



A typical stained blot:

- **Band 1** would give a very strong signal with an easy to read sequence.
- **Band 2** may need 2 lanes
- **Bands 3 and 4** would each give a clear read.
- **Bands 5, 6, 7** and **9** would be at the limit of detection and may need multiple lanes for a satisfactory sequence read.
- **Band 8** would give a workable signal for about 10 amino acids

Sample Requirements

We can sequence samples from various media types:

- PVDF membrane (blots)
- Gel slices
- In solution
- As solids

Samples provided on membranes need to be on PVDF. Nitrocellulose blots must not be used as they degrade within the instrument resulting in chargeable repairs. We advise the use of pre-cast gels or alternatively allow the gel to polymerise overnight before running your sample. This will help eliminate any N-terminal blockage due to unpolymerised acrylamide in the gel.

t

Sample Preparation

We advise that where a sample contains more than one soluble protein for sequencing, it should be more than 80% pure.

Reagents

All reagents used to resolve and solubilise proteins should be of the highest quality. Reagents that contain primary amines (glycine and ethanolamine) should be avoided. Dialysis membranes are often a source of interfering contaminants. If dialysis is necessary in the purification process, only high quality, thoroughly cleaned membranes should be used. The sample should be in a volatile solvent or buffer. Salts and detergents should be avoided where possible.

Staining Samples

Any of the common stains such as Ponceau, Amido Black, Coomassie blue and Page blue will be suitable for sequencing.

Ponceau and Amido black, in our experience, provide excellent results.

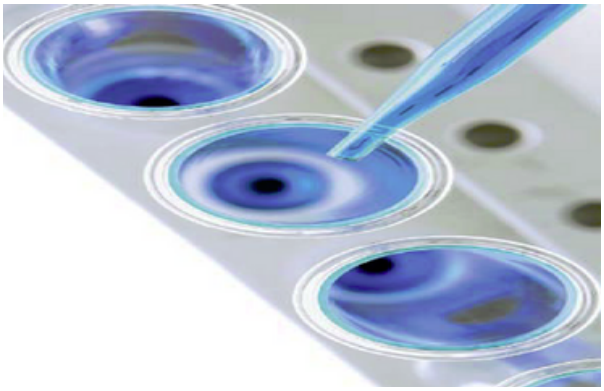
We would advise against using gold or silver stains, as these can block the N-terminus, preventing sequencing. In addition, these high sensitivity stains are not suitable for our sequencing technique.

Method Considerations

The first amino acid is often hard to identify because of the higher background often present during the first cycle, this is particularly true for Glycine.

Some sequences such as the Pro-X bond are abnormally stable resulting in partial resistance to the cleavage reaction and creating sequence 'carry overs'. AltaBioscience has adopted special measures to overcome this problem.

Proteins with a His tag at the N-terminus give poor sequences with a high degree of carryover.



N-terminal Blockage

Proteins and peptides which are N-terminally blocked do not have a free N-terminal amino group, therefore cannot be sequenced using this method. Some proteins naturally have chemically modified N-terminii. More than 50% of all eukaryote proteins are blocked, the majority of these groups being permanently attached with no effective means of removal. No sequence data can be obtained from these samples without prior fragmentation into smaller components.

Many proteins may be blocked at the N-terminus either by natural processes or by modification during manipulation and purification. Modification to the N-terminus may result from the following:

- Handling steps at elevated pH (>9.0)
- Inferior grade reagents and water
- Exposure to elevated temperatures if glutamine is the N-terminal residue
- Protease inhibitors which may react with amino groups
- Use of formic acid may formylate the N-terminus
- Non deionizing reagents such as urea

Depending upon the nature of the modification, some N-terminal blocking groups, if known, may be removed with mild acid treatment. Failing this, the solution is to carry out chemical or enzymatic cleavages followed by HPLC purification. Deblocking procedures can be performed on request but require significantly more protein. However please note that deblocking is not always successful even when the type of blockage is known.

Prokaryotic proteins tend not to be blocked.

Sequencing Considerations

Positive identification of cysteine requires special modification prior to N-terminal sequencing. Please enquire about this as an additional service. Glycosylated amino acids and phosphorylated amino acids may result in blank cycles, reduced peaks or altered retention times.

Fragmentation of Proteins

As described above, a large proportion of eukaryotic proteins cannot be sequenced, therefore an alternative is to digest the intact protein into fragments, separate, and sequence one of them. The easiest and much preferred method is digestion by cyanogen bromide (CNBr). CNBr cleaves cleanly and specifically after methionine residues, usually resulting in large fragments which are easily resolved on a PAGE gel. AltaBioscience offer a simple procedure for running this technique.

Tryptic digests should be avoided as they give large numbers of difficult to resolve fragments. Enzymes such as Lys-C and Asp-N are more specific and give manageable numbers of fragments. Unlike the CNBr method, there is a risk of background contamination from the enzyme reaction or the risk/possibility of sequencing the enzyme itself.

Modified Amino Acids

AltaBioscience can perform sequence analysis of post translationally modified amino acids. Post translational modifications such as hydroxyproline are easily identifiable.

If the presence of a phosphoserine is anticipated, here at AltaBioscience we use an efficient method to positively identify it.

You will receive

Results are delivered in a detailed and easy to understand signed pdf report Chromatograms can be provided on request.

The raw and processed data are stored electronically prior to archiving. Any archived data will be kept for a period of seven years.

About AltaBioscience

AltaBioscience, was founded in 1973, originally within The University of Birmingham. The ISO 9001 certified company provides services for academic research and for the pharmaceutical, academic, healthcare and biotechnology sectors worldwide

CONTACT US

For more details on sending samples including our sample submission form, visit **www.altabioscience.com**

To discuss any aspect of this service please contact us by phone on **+44 (0)1527 584495**
or email us at **info@altabioscience.com**

Advancing science through analytical chemistry and synthesis services in the life and food science sectors worldwide