PROTEIN CONCENTRATION BY AMINO ACID ANALYSIS (AAA)

Introduction

To determine the exact concentration of a pure protein sample, amino acid analysis (AAA) is considered by many to be the "Gold Standard" measurement. It can also be used to give a reasonable concentration of a heterogenous sample and is often used in the food industry for this purpose. (The technique can be used on proteins that are not in solution to give data on the ratio of amino acids the sample contains as a form of identity confirmation, but this is not the focus of this document).

While UV-Vis absorption at 280 nm is the most common way of measuring protein concentration (see our protocol on Determining Protein Concentration by UV), the method can have certain drawbacks, especially for proteins with low aromatic residue content or that have contaminants or buffer components that absorb at 280nm and therefore give a false reading. AAA does not suffer from these particular drawbacks but does have other considerations that are discussed below.

Other options to determine concentration that use a UV or fluorescence spectrometer are available. See our protocol on Protein Concentration Assays for the various pros and cons of these methods.

AAA technique overview

Protein samples from a diverse range of sources can be subjected to AAA. For example, it is used to quantify protein and amino acids in serum, fermentation culture media, food, drink, and animal feed. We at Peak Proteins use it for the accurate measurement of concentration of our purified protein samples.

The process to determine an accurate concentration of a protein sample is as follows:

1. Sample submission (usually in duplicate)
2. Sample preparation and acid hydrolysis
3. Recovery of amino acids
4. Calculation of concentration based upon amount recovered

Considerations:

- Need to know the exact amino acid composition of your sample to calculate the concentration (i.e., for this particular application it is not useful for unknown samples)
- Needs specialist equipment (but there are companies that provide AAA as a service. We use AltaBioscience in the UK).

Pros:

- It is the most accurate method of determining concentration of a pure protein sample

Cons:

- Need sample to be pure (generally >95%)
- Need specialist equipment
- Sample is destroyed during analysis
- Time consuming

AAA Procedure

Protein Sample preparation considerations

Liquid samples: For maximum accuracy the following points should be noted

- When using AAA to determine protein concentration, the protein sample should be as pure as possible, >95% at least. Otherwise contaminating proteins will affect the ratio of amino acids recovered after hydrolysis.
- Glycine buffers should be avoided as it obviously confounds the measurement of the glycine residue content of a sample. Glycine is also difficult to remove from the equipment post analysis.
- Phosphate salts or divalent metals should be avoided as they can give low and variable yields of most amino acids during hydrolysis.
- Glycerol, Acrylamide and Mannitol should not be included due to the risk of explosion during hydrolysis.
- Ammonium salts should not be used in the latter stages of any purification protocol as ammonia can cause the analysis reagent to precipitate out in the reaction coil.
- Polyethylene glycol (PEG) should be avoided as it can react with certain amino acids and adversely affect their recovery.
- When Ion exchange is used for the recovery method, high salt buffers should be avoided as they can carry through the hydrolysis and affect the recovery. Also, UV absorbing compounds e.g., mercaptans and detergents can give peaks that interfere with quantification.
- For dilute samples pre-concentration without loss is a problem. The best option is to use either TCA/Deoxycholate or acid-acetone precipitation.

Notes for solid samples:

- Hygroscopic proteins are difficult to weigh and can therefore impact the accuracy of any subsequent measurement.
- For adsorbing samples on to a membrane use PVDF membranes or filters. Samples on nitrocellulose can't be analysed.

Acid hydrolysis

6M HCl / 0.1% phenol for 20-24 hr at 110°C under vacuum are the most common conditions used for hydrolysis of protein samples.

Longer times can be used for large or difficult to hydrolyse proteins, but that impacts the recovery of certain amino acids even more (see below)

Separation of amino acids

There are 2 methods commonly used for separating the AAs post hydrolysis, ion exchange (IEX) and reverse phase high pressure liquid chromatography (rp-hplc). For the IEX, post column derivatisation with ninhydrin, o-phthaldehyde or fluorescamine is often used to allow detection of the AAs. For the rp-hplc pre-column derivatisation with phenyl isothiocyanate giving phenylthiocarbamyl (PTC) amino acids is used.

Recovery of amino acids

When analysing the data obtained from the AAA, the stability and recovery of the various amino acids through the AAA process needs to be considered when back calculating the concentration of the original sample.

Amino Acid	Stability and recovery
Ala Leu Asp Glu Arg Phe Lys	Are stable and have typically show good recovery.
Pro Gly	Are reasonably stable and usually have good recovery.
Asn and Gln	Are hydrolysed quantitatively to Asp and Glu respectively, to release ammonia. This should therefore be accounted for in the predicted amount of Asp and Glu that is expected from the target sequence.
Thr	Typically shows losses of 5%
Ser	Typically shows losses of 10%
Trp	Largely destroyed during hydrolysis.
Tyr	Can be halogenated, inclusion of 0.1% w/v phenol prevents this.
Met	Oxidised to methionine sulphoxide. Addition of reductant (e.g. 2-mercaptoethanol prevents this).
Cys and Cys-Cys	Largely destroyed during hydrolysis, so can't be determined directly (but see below).
His Val Ile	Have poor recovery, so should not be used in back calculating the concentration of the original sample.

Further considerations

Ala-Ala, Ile-Ile, Val-Val, Val-Ile, Ile-Val, Ala-Val. Bonds are highly resistant to hydrolysis and may well give a low recovery for those AAs. This should be considered when calculating the predicted amounts of these AAs released from the original sample. One option is to perform a second hydrolysis for 92-120 hrs, but this destroys other AAs especially Ser and Thr.

Additional methods that can be used for unstable amino acids

Performic acid oxidation

Converts cysteine and cystine to cysteic acid and Met to methionine sulphoxide, derivatives that are stable during acid hydrolysis. Typical yields from this process are about 95%, so it is best calculated by reference to molar quantities of stable AAs. Note this method destroys Tyr and Trp and can affect the recovery of other amino acids, so needs to be run as a separate sample.

Reduction and Alkylation of cysteine

This is another method that can be used to determine the cysteine content of protein. Commonly alkylating reagents are Iodoacetic acid (IAA) and 4-vinylpyridine (4-VP). IAA produces S-carboxymethylcysteine which is then stable to acid hydrolysis (note it may cyclise though and give lower recovery yields). 4-VP produces S-P-(4-pyridylethyl)cysteine. This elutes after the ammonia peak on IEX analysis which is a variable baseline segment. Better data is obtained using the rp-hplc analysis method.

In general half of sample is taken through the reaction to modify and determine free thiols. The other half is reduced under denaturing conditions and alkylated to work out cysteines.

Tryptophan

Trp is difficult to analyse, but there are methods involving alkaline hydrolysis with barium hydroxide and there are other protocols that involve hydrolysis with mercaptoethanesulphonic or methanesulphonic acids.

Asparagine and Glutamine

If specific levels of Asn and Gln are required, Bis(1,1-trifluoroacetoxy)iodobenzene treatment before acid hydrolysis converts Asn and Gln to their diaminopropionic and -butyric residues. These elute near or with the Lys peak and so needs to be determined by the difference from previous analysis. Note, Thr, Met, Cys, Tyr and His content are also altered by this method.

Back calculation of sample concentration

The data from the AAA is usually reported as the recovered nmole/mL of each individual amino acid, based on integration of the recovered peak area from the IEX or rp-hplc trace. If an internal reference standard, such as norleucine has been included, this data can be corrected for the recovery of that standard.

There are different ways to convert this data back to the concentration of the original protein sample. One method is given below.

1. From the amino acid sequence of the target protein (including any tag sequences, deletions etc) calculate the total number of each amino acid that are contained in the sequence. The Expasy ProtParam tool is useful for this.
2. Also calculate the total average mass in daltons for the target protein sequence.
3. Adjust, these numbers to the totals expected,
    a. by subtracting any pairs or triplets of Ala-Ala, Ile-Ile, Val-Val, Val-Ile, Ile-Val, Ala-Val that are present in the sequence as these are resistant to standard hydrolysis conditions.
    b. Adjust No of Ser residues to be 90% of total in sequence to reflect 10% loss in AAA
    c. Adjust No of Thr residues to be 95% of total in sequence to reflect 5% loss in AAA
4. If needed correct the data for the % recovery of any internal reference standard (e.g norleucine) that was obtained.
5. Using the average mass of each amino acid, convert the data to μg/mL recovered. Note, The Cys peak is observed as cystine (dimer), the weight reported will be the mass of cystine.
6. For each amino acid then back calculate the total concentration of the original protein sample using the formula below

Total conc (back calculated) = AA mg/mL detected x ((mw protein / (predict no aa x mw aa))

7. Pick the amino acids that have good recoveries; typically, Ala, Leu, Asp, Glu, Arg, Phe and Lys. Also Ser and Thr (corrected for 10% and 5% losses respectively), and possibly Pro and Gly depending on the data. The average the total concentrations calculated for these amino acids then gives you the final original concentration of your sample.