

# Natural Source Protein Purification: A Primer

*Purifying proteins from natural sources, such as urine, saliva, blood and other tissues, is more complex than recombinant protein purification. Certain information is needed at the outset and having a working understanding of the basics of the process will allow potential research partners to appreciate why the information is needed and how it will be used by the Guild BioSciences (GBS) staff. The following document presents a number of different aspects of natural source protein purification, focusing on the characteristics that influence the success or failure of the purification process. Ideally, the understanding acquired from this short primer will help potential partners gain a realistic expectation of the cost, duration, and final product of a given project. The final section serves as a reference, describing the major pieces of information that would be of use and may be requested by the GBS staff in the course of services.*

*While proteomics generally utilizes fairly standardized methods to analyze a set of proteins, purifying proteins from natural sources requires a more individualized and rigorous approach, as procedures must be uniquely developed for every protein. Therefore, the GBS staff scientists need to collect as much information as possible about the target protein in order to have the best chance of isolating and purifying the molecule in a timely and efficient manner. Because this type of procedure is more complex than analysis services such as 2DE proteome analysis or mass spectrometry services, it requires a considerable amount of correspondence with GBS staff during both quote preparation and the work period. During correspondence in the quoting process, in addition to asking about the protein of interest, the GBS staff scientists will inquire as to any health and safety issues that are expected to be associated with handling any of the material related to the protein's purification. Health and safety issue information, although not directly related to the purification process, is necessary for shipping regulations and bio-safety laboratory operations.*

## The Purification Process

### Detecting the target

In order to isolate and purify a protein, there must be some means of identifying which fractions of the separation contain the target protein and which do not. Detection of the protein of interest is usually accomplished via some known specific biochemical characteristic, such as molecular weight, iso-electric point, a specific metal or detectable cofactor (e.g. heme) that is bound, an uncommon quantity of a specific amino acid (e.g. high cysteine), an enzymatic assay specific to the protein, an affinity assay that can identify the protein (e.g. western blot, ELISA), or a combination of these. A preferable detection method would be one able to detect the protein in the starting material (if a solution), so that recovery can be determined during purification. The sensitivity, specificity,

reliability, and robustness of the detection method chosen will strongly influence the effectiveness of the purification efforts.

### Basic purification procedure

With the exception of affinity-based methods, purification of a protein to homogeneity by chromatographic methods is a multi-step process that typically requires 3-5 chromatography separations, each requiring a distinct type of chromatographic method. Each separation has 2-4 process steps:

*Step 1: Confirming detection of the target in unfractionated material.* The detection method suggested by a research partner is tested in-house to ensure acceptable performance when conducted by GBS staff. Verification is typically only required for the first separation step, but can vary considerably in the length of time required, depending on the sensitivity

and robustness of the detection method and the amount of target protein in the starting material.

*Step 2: Initial purification run.* When all available information about the target protein is collected and reviewed, the GBS staff will determine an appropriate starting procedure. Typically, the method chosen will be rather low resolution (i.e. a steep eluting buffer gradient) to reduce the number of fractions that must be analyzed. The purpose is to identify the point in the separation process at which the target protein elutes (washes out) from the chromatography column.

*Step 3: Separation optimization:* After the initial run, the conditions of the separation are varied in order to establish how to achieve the best separation and recovery possible. Manipulations include altering elution gradient slopes, creating complex gradients, altering starting conditions, and changing eluting components. Optimization is usually the most time-consuming step of the separation, depending upon the characteristics of the specific protein.

*Step 4: Collecting sufficient target for the next separation step.* After the separation has been optimized, sufficient quantities of partially purified protein must be produced to allow progression to subsequent separations using different types of chromatography. Appropriate amounts of partially purified protein can be generated either by (1) loading a larger amount of starting material or (2) if the capacity of the column would be surpassed, conducting multiple chromatographic runs and pooling the target-rich fractions. Concentration of the target-rich fractions and/or changes to the constitution of the dilution buffer may also be needed to

facilitate the next separation step. Such handling can be difficult when target protein concentrations are low.

## Chromatography 101

This section is not meant to be a detailed discussion on chromatography (of which numerous complete books have been published), but rather a superficial introduction to the subject. With the exception of size exclusion chromatography (also called gel filtration), nearly all forms of protein chromatography are based upon the same basic principles.

A chromatography column contains a solid resin having specific chemical characteristics. When a mixture of proteins in a specific buffer cocktail (the loading buffer) is applied to this column, some of the proteins will bind to the column, while others will wash through without binding (called the void peak). The strength with which proteins will bind depends on how the unique surface chemistry of each protein interacts with the column chemistry. Once the proteins have bound, a second (eluting) buffer is introduced to disrupt the attachment of the proteins to the resin. The eluting buffer is not introduced as a single addition at a specific concentration, but rather in the form of a gradient, whereby the percentage of the eluting buffer relative to the percentage of the loading buffer incrementally increases at a designated rate. This manner of application results in bound proteins experiencing a steadily increasing concentration of the eluting buffer over time. Proteins that bind weakly to the column separate at low concentrations of the eluting buffer, while progressively stronger binding proteins require increasingly higher concentrations of the eluting buffer to

detach. At the end of the flow path, different fractions containing different concentrations of the eluting buffer and the respective detached proteins are collected. Thus, the protein mixture is separated (or fractionated) into discrete units containing individual proteins or groups of proteins that dissociated from the column under specific eluting buffer concentrations. The above presentation, although considerably oversimplified, explains the fundamental concepts for most forms of chromatography.

Unfortunately, it is rare for a single chromatography separation to fully separate a target protein from a complex protein mixture (such as a cell lysate). Therefore, multiple chromatography methods, each relying on a different set of physical and/or chemical properties, are usually necessary to achieve complete purification of a protein. The most common forms of protein chromatography are:

- 1) ion (either anion or cation) exchange chromatography separates proteins by surface charge characteristics

- 2) size exclusion chromatography separates proteins by size (or more accurately by Stokes radii)
- 3) hydrophobic interaction chromatography (HIC) separates proteins by their surface hydrophobicity, usually using a high-salt loading buffer and a low-salt eluting buffer (compare to RPC below). Thus, conditions are generally mild, allowing separation of proteins with native structure intact.
- 4) reversed phase chromatography (RPC) separates proteins by their surface hydrophobicity as well, but utilizes an acidic loading buffer (i.e. trifluoroacetic acid) and an organic modifier (acetonitrile, methanol, etc) for elution. Although RPC is a bit harsher than HIC, it typically results in higher resolution.

GBS is capable of conducting these mainstream forms of chromatography as well as more specialized chromatographic forms, including:

- 1) group affinity chromatography techniques (hydroxyapatite chromatography, dye-ligand chromatography, immobilized metal-

### **A Note on Difficult Proteins**

Certain protein characteristics may render purification more challenging. Difficult proteins are most commonly proteins that are easily lost during handling (also see the section on Starting Quantities). Typically, losses occur when 'sticky' proteins adhere to the materials used during purification and subsequent handling, including the tubing of the HPLC system, the chromatography column, tubes used for collecting fractions, and membranes used to concentrate or change buffers (i.e. dialysis tubing, ultrafiltration membranes, etc). Strongly hydrophobic (e.g. integral membrane proteins, very large proteins) or highly charged (especially positively charged) proteins commonly have this character. A second protein characteristic that results in protein loss during purification is a tendency toward denaturation, and subsequent aggregation and precipitation. Proteins that denature easily often reposition hydrophobic regions, which are normally internalized, at the protein surface, transforming them into sticky proteins. In addition to adhering to equipment, these proteins may bond to each other or other proteins, forming aggregates and even precipitates that cannot be recovered once produced. The problem often becomes more prevalent as proteins are increasingly purified and concentrated, but sometimes can be avoided or minimized by the proper use of special additives. Examples of proteins likely to have this tendency would be proteins that exist in large complexes under physiological conditions. Once separated from the complex, they bind nonspecifically with other proteins and to each other, forming aggregates.

ion chromatography, and lectin affinity chromatography)

- 2) specific affinity chromatography techniques (immunoaffinity chromatography, and other ligand-specific affinity chromatography)

Utilizing the information available about the target protein's biochemistry, the chemical conditions at receipt, and staff expertise on the strengths and weaknesses of the various forms of chromatography, GBS personnel will devise a sequence of purification separations that will efficiently isolate and purify any target protein.

### Information that May Be Requested from GBS

Now that you have read through the basic processes involved in purification of proteins from natural sources, we hope you have a better understanding of the kinds of information that our staff may need. To guide you, the following paragraphs describe information that may be necessary or helpful when considering or processing your purification project. While all the suggested information may not be required for or available to you, the more information that you can provide, the more efficiently and cost-effectively we may derive a quote and complete the protein purification procedure.

#### Literature protocols

If the identity of the protein is known, it is advantageous for you to engage in some preliminary investigation of the literature to determine if others have already published purification methods for your target protein. Supplying this information to our staff at GBS may lessen the time we need to investigate and therefore, reduce your costs.

#### Known Biochemistry

Any biochemical information on your target protein that you can make available to the GBS staff, whether from the literature or your own research, will also reduce the amount of preliminary work, and hence costs to you. To facilitate timely processing, purification and analysis work that you have performed on the protein should be disclosed once a contract or non-disclosure agreement has been signed by GBS. We understand that you may be reluctant to share information about the protein target before the contract is signed, but please recognize that the accuracy of a given quote is limited by the amount of information supplied for its production. Thus, limited information can often result in a quote that is significantly over- or underestimated. Useful information that you might include would be:

- 1) molecular weight
- 2) iso-electric point
- 3) relative hydrophobicity
- 4) requirement of cofactors (metals, organic ligands, co-enzymes, etc)
- 5) type of protein (e.g. integral membrane, membrane-associated, or secreted)
- 6) structural biochemistry, such as whether it forms dimers or oligomers, or is normally part of a complex
- 7) biological function (e.g. enzymatic)
- 8) problematic characteristics, such as whether it is known to be sticky, or to easily denature or precipitate, and under what conditions. If the nature of the stickiness (hydrophobic or highly charged) is known or suspected, please inform the GBS staff as this can influence our selection of products to use in the purification process. Be assured that

we protect your samples by using a Waters 626 HPLC system, optimized for protein purification. With a flow path constructed of bio-friendly materials, this specialized equipment reduces any protein losses from your target sample. However, even with this system, difficult proteins will have lower recoveries and, consequently, require more optimization to minimize losses or more chromatography runs to compensate for the losses, thus increasing expected costs.

Therefore, it is important that this type of information is disclosed early, so that cost prediction may be applied accurately.

Remember, the more information provided, the more accurate the quote will be, and the quicker the purification will result in a deliverable.

### Starting material

Potential starting materials include intact tissue, homogenate, soluble extract, partially purified material, etc. If a tissue is your starting material, the specific type, source, and any notable health issues should be disclosed as well as (ideally) how you would like it disrupted. If a homogenate, protein extract, or partially purified preparation (i.e. ammonium sulfate precipitation purified) is your starting material, the specific reagents and their approximate concentrations in the extract solution should be disclosed, if possible, as this can significantly affect the type of chromatography GBS personnel will choose for the first step of the purification. Also, information regarding prior treatments, such as a DNase treatment or another processing technique that could alter the original lysate, is helpful. If your starting

material is a form of protein extract, providing us with the total protein concentration ensures appropriate loading of the chromatography column, accurate determination of target recovery, and target purity.

### Protein Quantities

Key considerations in any GBS purification plan are the quantities of the target protein present in the starting material and required in the final purified product. The size of these two values determines much of the cost of purification.

**STARTING QUANTITY:** Although arguably the most important quantity that must be determined, a quantification method need not be extremely precise, and a rough estimate is generally sufficient. Sometimes this quantity cannot be estimated beforehand, introducing a significant amount of uncertainty to the process, and likely resulting in your quote being considerably higher due to the extensive planning required by our staff. The additional time and effort needed is related to the difficulty of effectively purifying low abundance proteins. Low abundance proteins are challenging because they (1) require more time to initially detect – process Step 1 – depending on the detection method, (2) require much more starting material to concentrate sufficient target protein for subsequent purification steps, and (3) are subjected to higher proportional losses during purification procedures (losses to the column, tubing, etc).

**FINAL QUANTITY DESIRED:** Also an important piece of the process, the larger the quantity of purified material desired, the more starting material will be required. Please carefully consider

whether you can provide an adequate amount of starting material to achieve the final amount of purification product you desire. Bear in mind that it may require multiple chromatographic runs at each separation step to produce enough target-rich material for subsequent steps, which leads to higher costs. Also, in some cases, highly concentrated pure proteins can be more prone to aggregation and precipitation, which can lower yields, necessitating more runs to achieve the desired final quantity, and more optimizing to improve recovery – again, elevating costs.

### **Purity**

When discussing your purification needs with a member of the GBS staff, please indicate the level of purity needed in the final product. Different post-purification applications demand different degrees of protein purity. For example, if you plan to use your purified protein as an antigen for monoclonal antibody development, it will need to be essentially homogenous (as pure as possible), while a purified enzyme for an enzymatic assay might only need to be free of contaminants that would interfere with its enzymatic action. Typically, the greater the desired purity, the more purification steps will be necessary, the more time will be needed to reach a deliverable, and the higher the cost. Therefore, in the interest of minimizing

expenditures, consider the final use of your protein and only request a level of purity required for that function. If you prefer a specific method for assessing purity, please let us know. You may choose from SDS PAGE, SDS PAGE with western blotting, two-dimensional electrophoresis, or mass spectrometry (outsourced). The standard technique used at our facility is SDS PAGE.

### **Native structure/biological activity**

When reviewing your proposed project, consider whether the final product needs to have its native structure maintained and its biological activity intact. For example, if your target protein is an enzyme that will be used in an enzymatic assay, both its native structure and biological activity will need to be preserved. If your protein is an enzyme you will use for the production of an antibody, its native structure must be preserved, but not necessarily its biological activity. If your protein is to be digested so that you may determine its glycosylation state by mass spectrometry, then neither the native state nor biological activity are essential. The choice of chromatographic methods depends in part on the final use of the product. Some types of chromatography are gentler than others, so the desired condition of the purified product, or its intended use, can influence the series of steps we choose for purification of your particular protein.

### **Final thought**

As mentioned previously, we understand that you may only be able to supply some of the information encouraged. Success may be possible with only a small portion of the suggested information. However, the amount of time and cost may increase accordingly. The information above will hopefully prove useful to you, so that you may inquire about our services with a basic understanding of the types of information we may request and an idea of the costs that factor into the quote that you receive. We appreciate your consideration of Guild BioScience's Protein Services and wish to work with you in the future.