Technological University Dublin
ARROW@TU Dublin

Books/Book Chapters/ Proceedings

School of Food Science and Environmental Health

2017

Differential Precipitation and Solubilisation of Proteins

Barry J. Ryan

Gemma K. Kinsella

Follow this and additional works at: https://arrow.tudublin.ie/schfsehbk

 Part of the Bioinformatics Commons, Food Science Commons, and the Molecular Biology Commons

This Book Chapter is brought to you for free and open access by the School of Food Science and Environmental Health at ARROW@TU Dublin. It has been accepted for inclusion in Books/Book Chapters/ Proceedings by an authorized administrator of ARROW@TU Dublin. For more information, please contact arrow.admin@tudublin.ie, aisling.coyne@tudublin.ie, gerard.connolly@tudublin.ie.

This work is licensed under a Creative Commons Attribution-Noncommercial-Share Alike 4.0 License

Differential Precipitation and Solubilisation of Proteins.

Barry J. Ryan and Gemma K. Kinsella

School of Food Science and Environmental Health, Dublin Institute of Technology, Cathal Brugha Street, Dublin 1, Republic of Ireland.

Email: barry.ryan@dit.ie

Summary/Abstract

Differential protein precipitation is a rapid and economical step in protein purification and is based on exploiting the inherent physicochemical properties of the polypeptide. Precipitation of recombinant proteins, lysed from the host cell, is commonly used to concentrate the protein of choice before further polishing steps with more selective purification columns (e.g. His-Tag, Size Exclusion etc.). Recombinant proteins can also precipitate naturally as inclusion bodies due to various influences during over-expression in the host cell. Although this phenomenon permits easier initial separation from native proteins, these inclusion bodies must carefully be differentially solubilised so as to reform functional, correctly folded proteins. Here, appropriate bioinformatics tools to aid in understanding a protein's propensity to aggregate and solubilise are explored as a backdrop for a typical protein extraction, precipitation and selective resolubilisation procedure, based on a recombinantly expressed protein.

Key Words

Ammonium Sulphate precipitation, Bioinformatics, Inclusion body solubilisation, Protein refolding, Trichloroacetic Acid precipitation, Three Phase precipitation.

1. Introduction

Protein precipitation can be caused by the differential solubility between a protein-rich soluble phase and a solid chemical precipitant. Soluble proteins can be insolubilised by interaction with a suitable precipitant that decreases the protein's attraction to the solvent and increases the protein's attraction to other protein molecules, resulting in protein accumulation and eventually precipitation. The addition of low molecular weight substances, such as glycerol, polyethylene glycol and sucrose, and high molecular weight substances such as serum albumin, can have significant effects on protein structure and stability. Preferential hydration of a protein molecule caused by the presence of these additives can increase the protein's stability. Certain salts can also exert a stabilising effect by 'salting out' hydrophobic residues of a protein, causing the molecule to adapt a more compact, stable structure **(*1*)** frequently resulting in precipitation. The use of such protein precipitating molecules is an empirical process, the effects of any given substance on a protein must be determined experimentally. The use of additives can not only be used as a simple approach to increase the stability of a given protein, but also to actively effect protein precipitation. Protein precipitation can be used as a crude protein clean-up method from cell lysates, readily employed after bacterial over-expression of recombinant proteins.

Differential solubilisation of proteins is often employed for proteomic analyses **(*2, 3*)**, but it too can offer an alternative purification technique for non-soluble recombinant proteins expressed in heterologous hosts. Recombinant proteins expressed as inclusion bodies can be readily separated from the host cell protein matrix, however careful solubilisation and refolding are critical for obtaining suitable recombinant proteins for further downstream processes.

Protein modelling and *in silico* analysis can assist in the experimental design process, for example, in reducing the empirical experimentation required. Bioinformatic analyses can be undertaken to gain an understanding of the physicochemical properties of the target protein (e.g. amino acid composition, secondary structure prediction) and its propensity to aggregate at the outset of a protein expression project. The solubility of a protein upon expression is heavily dependent on its primary amino acid sequence and initial solubility prediction methods were based on the content of charged and turn-forming residues. In recent years, more advanced prediction software has been developed utilising training sets of soluble and insoluble proteins. These tools are used to predict solubility before performing wet lab experiments thus saving effort, time and cost [recently reviewed elsewhere; ***(4)***].

By way of a worked example, a typical recombinant protein precipitation and resolubilisation procedure, preceded by judicious bioinformatic analyses, is outlined.

2. Materials

Notes:

1. All consumables may be sourced from Sigma-Aldrich unless otherwise stated.
2. A computer and Internet access is all that is required for bioinformatic analyses.

2.1 Recombinant / Native Protein Extraction.

- Centrifuge (Bucket Type, e.g. *J2-21*, Beckman and microfuge, e.g. *5415D* Eppendorf).
- pH meter (e.g. *M-240*, Corning).
- Sonicator (e.g. *Vibra Cell*, Sonics Scientific).
- Water bath (temperature controlled).
- Vortex.
- Resuspension Buffer One: 50 m*M* Na_2HPO_4-NaH_2PO_4, pH 8.0, 0.3 *M* NaCl.
- -80 ºC freezer

2.2 Protein Precipitation using Ammonium Sulphate.

- Resuspension Buffer One: 50 m*M* Na_2HPO_4-NaH_2PO_4, pH 8.0, 0.3 *M* NaCl.
- Saturated Ammonium Sulphate: Add 750 g of $(NH_4)_2SO_4$ to 1 L of double distilled water in a beaker or flask. Stir the solution at room temperature with a magnetic stirrer for 15 min or until saturation. Gently decant the clear supernatant solution after the undissolved solids settle on the bottom of the flask.
- Graduated Pipette (10 mL).

2.3 Protein Precipitation using Trichloroacetic acid.

- 2% Deoxycholate (DOC): Add 2 g of DOC to 100 mL of double distilled H_2O, mix well.

- 100% Trichloroacetic acid (TCA): Add 1 g of TCA to 454 µL double distilled H_2O and mix carefully. Store in a light proof bottle at 4ºC until required for use. TCA is a harmful skin and eye irritant. Always use correct personal protective equipment when handling it.
- Acetone (ice cold). Store acetone at -20ºC. Use directly from -20ºC.
- Vacuum Dryer (e.g. *DNA 110 Speed Vac*, Savant)
- Resuspension Buffer: 50 m*M* Na_2HPO_4-NaH_2PO_4, pH 8.0, 0.3 *M* NaCl.
- Centrifuge (Bucket Type, e.g. *J2-21*, Beckman and microfuge, e.g. *5415D* Eppendorf).

2.4 Protein Precipitation using Three Phase Partitioning.

- Saturated ammonium sulphate (see Section 2.2).
- *t*-Butanol (HPLC grade).
- Aspirator (e.g. *FB70155*, Fisherbrand) or siphon tube.
- Centrifuges (Bucket Type, e.g. *J2-21*, Beckman; microfuge, e.g. *5415D* Eppendorf).

2.5 Protein Solubilisation.

- DN*ase* I (100 U/mL).

- Resuspension Buffer Two: 50 m*M* Na_2HPO_4-NaH_2PO_4, pH 8.0, 0.3 *M* NaCl, 5m*M* DTT (*see* **Note 1**), 0.35 mg/mL Lysozyme, Proteinase Inhibitor Cocktail (*see also* **Chapter 4**). Make up as fresh prior to use.
- Triton X-100.
- Solubilisation Buffer: 50 m*M* Na_2HPO_4-NaH_2PO_4, pH 8.0, 0.3 *M* NaCl, 25 m*M* DTT, 6 *M* Guanidine HCl.
- PBS-T: Phosphate buffered saline (PBS, 1x) containing 1% (v/v) Triton X-100.
- Centrifuges (Bucket Type, e.g. *J2-21*, Beckman; microfuge, e.g. *5415D* Eppendorf).
- Vacuum Concentrator (*Speed Vac*, Savant)

2.6 Protein Refolding.

- Refolding Buffer: 50 m*M* Na_2HPO_4-NaH_2PO_4, pH 8.0, 0.3 *M* NaCl, 2.5 m*M* reduced Glutathione, 0.25 m*M* Oxidized Glutathione, 0.2 *M* Arginine.
- Dialysis Buffer: 50 m*M* Na_2HPO_4-NaH_2PO_4, pH 8.0, 0.3 *M* NaCl.
- Dialysis apparatus (dialysis tubing and clips, magnetic mixer, large clean container).
- Guanidine Hydrochloride: 6 *M* stock, made in double distilled H_2O.
- Amicon protein concentration device (e.g. *Ultra-15 Centrifugal Filter Units*, Amicon).
- Gradient maker apparatus (*see* **Fig. 1**).

3. Methods

3.1 Computational prediction of protein solubility

1. Obtain the protein of interest's Accession number (a unique identifier) and its amino acid sequence (in FASTA format) by searching the Uniprot database (www.uniprot.org).
2. Use TMHMM 2.0 (http://exon.niaid.nih.gov/cas/manual/TMHMM.html) or another membrane protein webserver to predict the protein's secondary structure and whether it is expected to be membrane bound.
3. If the protein of interest is not a membrane protein, utilise PROSO II (PROtein SOlubility: http://mips.helmholtz-muenchen.de/prosoII; ***5***) or ESPRESSO (EStimation of PRotein ExpreSsion and Solubility: http://mbs.cbrc.jp/ESPRESSO/Submission.php; ***6***) to classify the protein as soluble or insoluble (*see* **Notes 2 and 3**).

3.2 Optimisation of protein solubility

After bioinformatic analysis, if the protein is predicted to be insoluble this could result in unwanted precipitation. An approach towards preventing this is through random or rational mutagenesis of specific codons to enhance solubility (*see* **Note 4**).

1. Search the Protein Databank (PDB, http://www.rcsb.org/pdb/home/home.do) for a structure (crystal/nuclear magnetic resonance) of the protein of interest and download the relevant PDB file. If a template structure is unavailable, a comparative model may be available (e.g. ModBase, http://modbase.compbio.ucsf.edu/modbase-cgi/index.cgi).

2. Visualise the protein structure using suitable software, for example, Pymol software (www.pymol.org) or Swissprot Viewer (http://spdbv.vital-it.ch/).
3. Analyse solvent accessible residues and determine possible hydrophobic surface amino acids for mutation (*see* **Note 5**).
4. Design a mutated gene for improved protein solubility (*see* **Note 6**). Translation software such as GeneDesigner (https://www.dna20.com/resources/genedesigner) is available to aid in this process **(*7*)**.

3.3 Recombinant / Native Protein Extraction.

The source of the protein will determine the optimal technique to release the protein from the tissue or cells in which it is contained. The typical freeze-thaw cell lysis procedure (below) is generally sufficient to lyse most bacterial cell types although other options are available (*see* **Notes 7, 8, 9**).

1. Collect the bacterial cells by transferring the bacterial culture to a pre-chilled sterile centrifuge tube and centrifuge at low speed (5 min, 800 x *g*) in a previously cooled centrifuge (4°C).

2. Carefully remove the culture media from the bacterial cell pellet, ensuring the pellet is not disturbed.

3. Resuspend the cell pellet in Resuspension Buffer One, in 10% of the original culture volume (*see* **Note 10**).

4. Freeze the resuspended cells to -80°C by placing the resuspension solution (still in the plastic centrifuge tube) into a pre-equilibrated -80°C freezer (or liquid Nitrogen), then warm the cells to 37°C (using a pre-equilibrated water bath) for 10 min. Repeat this freeze thaw process three times (*see* **Note 11**).

Sonication can also be used if the protein is not released during the freeze thaw steps. It is crucial to maintain the cell suspension on ice during the sonication process (*see* **Notes 12, 13, 14**).

5. Sonicate at 10 amplitude microns for 10-20 s.

6. Allow the cell suspension to stand on ice for 30 s.

7. Repeat steps 5 and 6 three more times.

8. Check the recombinant protein induction/expression by loading and analysing a representative sample (typically 50 µg protein) onto a sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) gel (*see* **Note 15**).

3.4 Protein Precipitation using Ammonium Sulphate.

A common and inexpensive first step to isolate and concentrate proteins during protein purification is precipitation with an external additive. This additive alters the physicochemical properties of the protein causing it to fall out of solution. Selecting a

precipitating agent is a matter of identifying the one that provides the desired protein in the most optimal final state. Ammonium sulphate (*see* **Note 16**) is commonly used for large-scale precipitations.

1. Gently stir the protein mixture with the aid of a magnetic stirring bar at 4 ^{o}C. Add, using a graduated pipette, the saturated ammonium sulphate solution drop-wise to the protein solution until precipitates start to form (*see* **Notes 17, 18 and 19**).

2. Once sufficient saturated salt solution has been added to cause precipitation of the protein of interest (indicated by collection of precipitate at the bottom of the container), centrifuge the mixture at 10,000 *g* for 15 min in a pre-cooled rotor. Collect the precipitate by carefully discarding as much supernatant as possible (*see* **Note 18**).

3. Resuspend the protein pellet at 4^{o}C in Resuspension Buffer One for further downstream processes (*see* **Notes 20 and 21**).

3.5 Protein Precipitation using Trichloroacetic acid.

Trichloroacetic acid (TCA) is routinely employed for small-scale operations or precipitations of protein preparations that are at low concentration; however, it should be noted that this procedure is protein denaturing, and **caution** must be exercised when working with TCA (*see* **Notes 22, 23, 24**).

1. To one volume of protein solution, add 1/100 volume of 2% DOC (sodium deoxycholate).

2. Vortex and incubate for 30 min at 4°C.

3. Add 1/10 volume of 100% Trichloroacetic acid (TCA). Vortex the solution and incubate statically overnight at 4°C.

4. Centrifuge the sample for 15 min at 4°C (10,400 x *g*). Gently remove the supernatant and retain the pellet. Carefully dry the tube by inversion on tissue paper (note: the pellet may be difficult to see).

5. Optional: Wash the pellet twice with one volume of ice-cold acetone. Vortex and re-pellet the samples by centrifugation 10,400 x *g* for 5 min at 4°C between washes (*see* **Note 24**).

6. Dry the samples under vacuum (e.g. *Speed Vac*, Savant) or allow to air dry.

7. Resuspend the protein pellet in a buffer of choice for further downstream processes.

3.6 Protein Precipitation using Three Phase Partitioning (TPP)

TPP is a three-stage batch method, which is a hybrid of salting out and alcohol precipitations. Its key advantages are that a lower concentration of ammonium sulphate is required and the technique can be performed at room temperature (*see* **Note 25**).

1. Whole cells are treated with aqueous 1.6 *M* ammonium sulphate solution buffered at pH 8.0 with 50 m*M* Tris-HCl, pH 8.0, with shaking.
2. Add 1 volume of tertiary butanol (*t*-butanol). Shake vigorously for 1 min at room temperature (RT).
3. Centrifuge (10,000 *x g*) for 15 min at room temperature. Three phases separate during centrifugation.
4. Remove the upper (alcohol) layer by aspiration (*see* **Note 26**). Remove the lower precipitate layer, leaving the interfacial aqueous layer.
5. Fresh *t*-butanol is added to the remaining aqueous layer (with vigorous shaking once again)
6. Precipitation again and collect layers as above.
7. Add ammonium sulphate to the precipitate layer (see Section 3.4).
8. Resuspend the protein pellet at 4 ^{o}C in Resuspension Buffer One for further downstream processes.

3. 7 Protein Solubilisation.

Recombinant protein expression in a heterologous host frequently can result in insoluble and inactive proteins (*see* **Note 27**). Regularly, protein overexpression results in the production of inclusion bodies, which are insoluble aggregates of misfolded protein. Although these inclusion bodies can be purified, further characterisation of this protein mass is often impossible without solubilisation of the protein of interest and refolding into an active form (*see* **Note 28**). A typical inclusion body solubilisation, using Triton

X100 as a detergent, and refolding protocol is outlined below (*see also* **Note 29**). An additional detergent screening methodology, to identify superior alternatives to Triton X100, is also detailed.

1. Carry out steps 1-3 as outlined in Section 3.3, except resuspend the cell pellet in 10% of the original culture volume of Resuspension Buffer Two.

2. Slowly add Triton X-100 (to a final concentration of 1% v/V), and mix gently (*see* **Note 30** and Section 3.7.1).

3. Carry out the sonication procedure as outlined in Section 3.3, steps 5 -7.

4. Incubate the cell debris with DN*ase* I (100 U/mL) for 1 h at 37 ºC.

5. Collect the inclusion bodies by centrifugation at 30,000 x *g* for 30 min at 4 ºC.

6. Wash the inclusion body pellet twice with PBS-T, followed by centrifugation at 30,000 x *g* for 30 min at 4 ºC.

7. Solubilise the pelleted inclusion bodies in the solubilisation buffer and allow total solubilisation to occur at 4 ºC for 1 h, with occasional gentle mixing.

8. After 1 h solubilisation, remove all remaining insoluble material by centrifugation 30,000 x *g* for 10 min at 4 ºC (*see* **Note 31**).

9. Determine the protein concentration and adjust to 1 mg/mL by dilution in solubilisation buffer and proceed directly to re-folding at 4 ºC (Section 3.8).

3.7.1 Protein Solubilisation Detergent Screen

Selecting the appropriate detergent can be crucial for effective purification, particularly of membrane proteins. An additional detergent screening methodology is provided if required.

1. Lyse cells as previously described. A membrane preparation may be performed if required (*see* **Note 32**).
2. Small aliquots of lysed cells/cellular membranes are mixed with an equal volume of Resuspension Buffer Two containing different detergents (e.g. DDM, LDAO, DM, FC-12, C12E9; see reference ***8***). To avoid the production of foam, do not vortex the solubilisation mixtures but mix by gently pipetting up and down.
3. Gently mix using an end-over-end rotator at 4 °C for 12–14 h (*see* **Note 33**).
4. Take a 10 µL sample for analysis by SDS-PAGE and store at 4 °C.
5. To pellet the unsolubilized material, centrifuge the remainder of the sample for 30 min at 100,000 x *g* and 4 °C.
6. Remove the supernatant from each tube taking care not to disturb the pellet and transfer it to another clean, chilled 1.5 mL tube.
7. Mix the contents of each tube by gentle pipetting up and down. Take a 10 µL sample for analysis by SDS-PAGE.
8. Examine samples by SDS-PAGE, and Western blot if appropriate, to identify the detergent and conditions that resulted in the best solubilisation (*see* **Note 34**)

9. Incorporate these detergent and incubation parameters for the purification of the target protein.

3.8 Protein Refolding.

In recent years there have been many novel approaches developed to induce protein refolding. These include high hydrostatic pressure **(*9*)**, dialysis against PE-PEG **(*10*)**, solid-phase refolding in cation-exchange resin with decreasing gradient of urea **(*11*)** and the use of micro-fluidic chips **(*12*)**. However, one of the most cost effective and easiest to execute is simple low concentration, gradient dialysis, as detailed below.

1. Dilute the solubilised proteins as quickly as possible (to yield a final protein concentration of 0.1 mg/mL, *see* **Note 35**) into pre-chilled Refolding Buffer (*see* **Note 36**).

2. Dialyze the diluted solubilised protein overnight against a 200-fold volume of dialysis buffer with slowly decreasing concentrations of guanidine hydrochloride (GuHCl) (typically decrease GuHCl concentration as follows: 6 *M*, 4 *M*, 2 *M*, 1 *M*, 0.5 *M* and then 0 *M* in a continual dialysis approach; *see* **Fig. 1** and **Notes 37, 38**).

3. Centrifuge the dialysate at 4 °C for 30 min at 30,000 x *g*.

4. Carefully remove the liquid protein rich layer, concentrate (e.g. *Amicon* filtration) and store at an appropriate temperature (*see* **Note 39**).

5. To ensure correct protein folding perform an appropriate bioassay and run a sample on SDS-PAGE.

[Insert Figure 1 here]

Notes

1. If the protein contains cysteine residues, oxidation could lead to protein aggregation. A reducing agent, such as DTT, is added to the buffer to prevent this.

2. PROSO II is a machine-learning model trained on an experimentally available dataset of proteins. This is a public webserver to classify proteins into “soluble” and “insoluble” classes. ESPRESSO implements both a sequence/predicted structural property-based method and a sequence pattern based method to predict protein solubility.

3. If the protein is predicted as being soluble, proceed with recombinant expression and purification; if predicted as being insoluble, mutagenesis (Section 3.2) is an option or the protein can be expressed insolubly and resolubilised (Section 3.3 onwards).

4. In general, the amino acids on the protein’s surface (those not affecting function) are primary mutagenesis targets, with the mutagenic approach taken to minimise the weak interactions that result in inclusion body formation. This is typically achieved by replacing the hydrophobic

surface amino acid residues with charged, polar hydrophilic or less hydrophobic residues **(*13*)**. The success of rational molecular engineering depends on the availability of a suitable resolution 3D structure of the candidate protein for selection of the surface residues to mutate.

5. Computational prediction of a proteins aggregation propensity is based on its amino acid sequence. Available analysis tools include AGGRESCAN (http://bioinf.uab.es/aggrescan/) or AMYLPRED2 (http://biophysics.biol.uoa.gr/AMYLPRED2). Molecular engineering, such as the addition of a translation fusion Green Fluorescent Protein (GFP) tag for example, can facilitate easier downstream analysis through florescence detection **(*14*).**

6. Alternatively, an approach is to determine folding free-energy estimations of proposed mutated structures and focus on those site mutations likely to enhance solubility e.g. PoPMuSiC software (***15***; http://dezyme.com/).

7. There are several methods to achieve this, including repeated freezing and thawing, sonication, homogenization at high pressures, enzymatic lysis or permeabilization by organic solvents. The method of choice depends on how fragile the protein is and how robust the host cell is **(*16*)**. Commonly cited examples of cells with modified robustness include the use of easily lysed *E. coli* expression cell lines (e.g. pLYSs mutants or alternative mutants; ***17***) or wall-less strains, L-form strains, lpp deletion strains and the use of co-expression of lysis-promoting proteins **(*18*)**. The use of molecular engineering to introduce a signal peptide (e.g. pelB from *Pectobacterium carotovorum*, formally known as *Erwinia carotovora,* or SP from Bacillis; ***19***) may also offer a method to direct protein expression to an easily extracted portion of the fermentation. Methods of extraction from the various cellular, and extra cellular, compartments are detailed in **Notes 8 and 9**. Molecular methods can be further enhanced by simple inclusion of

specific additives and compounds to increase targeted protein accumulation **(*20*)** and downstream solubility **(*21*)**.

8. Proteins can also be selectively released from the various compartments of a bacterial host, for example proteins expressed in the periplasmic envelope can be selectively lysed by a method similar to that described by French and co-workers **(*22*)**. Pellet the bacterial cells to be disrupted by centrifugation at 800 x *g* for 3 min. Resuspend the pellet in Fractionation Buffer (F1) buffer in 20% of the original culture volume. The F1 comprises (final concentrations): 500 µg/mL lysozyme, 20% w/v sucrose, 1 m*M* EDTA, pH 8.0, 200 m*M* guanidine hydrochloride and 200 m*M* Tris-HCl, pH 8.0, at room temperature. Statically incubate the resuspended cells at room temperature for 15 min, after which add an equal volume of ice-cold water. Stand the mixture at room temperature for 15 min. Remove the cell debris by centrifugation at 10,400 x *g* for 10 min. Transfer the supernatant (containing the periplasmic fraction) to a clean container for further purification.

9. Proteins expressed and transported to the culture supernatant can also be conveniently concentrated by a method outlined by Caldwell and Lattemann **(*23*)**. In brief this method involves adding an equal volume of a PRMM solution (0.05 m*M* pyrogallol red, 0.16 m*M* sodium molybdate, 1.0 m*M* sodium oxalate, 50.0 m*M* succinic acid, 20% v/v methanol in H_2O, adjusted to pH 2.0 with HCl) to cleared (0.22 µm filtered) culture supernatant. Adjust the pH of the solution to 2.8 (±0.1), and allow the proteins to precipitate for 1 to 2 h at room temperature, followed by an overnight incubation at 4°C. Sediment the precipitate by centrifugation at 10,000 x *g* for 1 h, and carefully remove the supernatant. Repeatedly rinse the precipitate with 1 mL of

acetone. Remove all traces of acetone by evaporation at room temperature. Solubilise the precipitate by adding 100 µL of 2x SDS-PAGE sample buffer (25% glycerol, 8% SDS, 4% β-mercaptoethanol, 0.02% bromophenol blue, 100 m*M* Tris-HCl, pH 6.8). These samples can be applied directly to an SDS-PAGE gel, if this type of analysis is required.

10. A protein is least soluble when at its isoelectric point (p*I*). Hence, selection of an appropriate pH buffer range is important e.g. Hepes (pH 6.8-8.2); Tris-HCl (pH 7.5-9); phosphate (pH 5.8-8). To determine a given protein's p*I*, its accession number (or amino acid sequence in FASTA format) should first be retrieved from the Uniprot database (www.uniprot.org). Subsequently a webserver such as the bioinformatics resource portal can be utilised to determine a theoretical p*I* (http://web.expasy.org/compute_pi/) and aid in appropriate buffer selection (e.g. choose a pH buffer one unit above or below the p*I* of the protein of interest).

11. If the expression host of choice is a yeast, one of the most widely used methods for the disruption is agitation with glass beads. Several cycles of agitation are interspersed with cooling cycles to avoid overheating.

12. The sonication process can generate large amounts of heat, which is why pulses are limited to ~20 s. In between pulses, cool the tube in ice or ice-water slurry for 30 s. If a large volume is required to be sonicated, split the cell suspension into two tubes, and alternate the sonication and cooling steps. Sonicate on ice where possible in order to maintain protein activity **(*24*, *25*)** during sonication. See reference **(*26*)** for an in-depth review of sonication optimisation during recombinant protein lysis.

13. The extraction process also releases proteases, which will digest all proteins in the solution. If the protein is sensitive to proteolysis, it is desirable to employ a protease inhibitor (*see* **Chapter 4**), to proceed quickly, to keep the extract cooled to minimise proteolysis and to select an easily lysed cell line that is ideally protease-deficient (*see* **Note 7**).

14. Lysozyme (500 μg/mL, to assist cell wall degradation) and DN*ase* I (100 U/mL, to degrade genomic DNA) can be added to the lysis buffer.

15. An alternative approach to quantify protein induction and over expression is to use a Nanodrop spectrophotometer (Thermo Scientific). Similar to a SDS-PAGE gel analysis it is best to compare an expressed sample lysate against a replicate non-induced sample as a control. If the protein of interest is catalytically active a crude lysate test of functional activity, or an appropriate bioassay may also give an indication of successful over expression and cell lysis.

16. The addition of high concentrations of salt to a protein solution causes precipitation by removing water from hydrophobic patches on the protein's surface, resulting in these patches aggregating together causing the protein to come out of solution. A number of salts can be used for this process; NaCl, Na_2SO_4, KCl, $CaCl_2$ and $MgSO_4$, however, $(NH_4)_2SO_4$ is by far the most commonly used additive. This is due to several advantageous characteristics of the salt, including the fact that it has a high solubility in water (4 *M* saturation) and it has a low density at saturation allowing precipitated proteins to be collected by centrifugation. Hence the first proteins to be purified during ammonium sulphate precipitation are water-soluble proteins **(*27*)**. Alternatively, an empirically derived dual additive approach can also enhance selective protein precipitation based on the synergistic effect of electrostatic interactions in combination with traditional salting-out effects **(*28*)**.

17. The concentration of ammonium sulphate needed for precipitation varies from protein to protein and is dependent on salt/protein concentration, pH, temperature and time. It should be determined empirically **(*29*)**. In brief, this entails placing a volume of cell extract, e.g. 10 mL, into each of 5 test tubes. Adding, with gentle mixing, amounts of solid ammonium sulphate to yield 20 %, 30 %, 40 %, 50 % and 60 % w/v saturation and allowing to sit for 30 min to promote precipitation. The insoluble material is collected by centrifugation and the resulting pellets dissolved in an appropriate buffer, dialysed if required, and assayed for enzyme activity or analysed by SDS-PAGE.

Final concentrations of ammonium sulphate must be calculated using standard nomograms or with online tools ("Ammonium Sulfate Calculator", available at www.encorbio.com/protocols/AM-SO4.htm). Adding increasing amounts of ammonium sulphate causes the different fractions of a protein mixture to precipitate at different rates. One advantage of this method is that it can be performed inexpensively with very large volumes. Additionally, the high salt content of the precipitated protein permits its direct addition onto a hydrophobic interaction chromatography (HIC) purification column, thus speeding up the overall purification process.

18. Ammonium sulphate salt can be added either in saturated solution or directly as salt crystals. It may be advantageous to add pre-ground ammonium sulphate directly into the protein mixture as powdered solids during large-scale purification processes so that the effect of dilution by the salt solution is minimized. If a saturated salt solution is employed the amount of ammonium sulphate solution added must be recorded accurately, often this is achieved by dispensing from a graduated pipette. It is critical to avoid the spatial non-uniformity in the salt concentration during

the addition of the salt solution. Localized concentration "*hot-spots*" will prematurely initiate the precipitation of other proteins and inadvertently affect the precipitation process. Record the volume of the saturated ammonium sulphate solution required to precipitate the protein of interest. Also note that protein precipitation is not instantaneous; it may require more than 30 min to equilibrate.

19. Purification of integral membrane proteins requires the addition of a detergent such as sodium dodecyl sulphate (1% w/v SDS) to dissolve cell membranes and keep membrane proteins in solution during purification. It should be noted that SDS causes protein denaturation, hence milder detergents such as 1% Triton X-100 or 1% CHAPS can be used to retain the protein's native conformation during cell membrane dissolution.

20. Avoid frothing (as much as possible) when mixing, as air can promote oxidation of proteins and also cause protein denaturation at the air-water interface.

21. For many downstream processes it is important to de-salt and to remove/inhibit proteases (*see* **Chapter 4**) to avoid protein degradation. Alternatively, this salt rich preparation can be applied directly onto a hydrophobic interaction chromatography column, followed by a size exclusion column to effect purification and simultaneously remove precipitating salt (*see* **Note 17**).

22. TCA is a harmful skin and eye irritant. Always use correct personal protective equipment when handling it.

23. There are numerous options to effect other types of protein precipitation including (but not limited to) acetone precipitation (useful to simultaneously eliminate acetone soluble components

and increase protein concentration), ethanol precipitation (useful to simultaneously concentrate proteins and remove traces of GuHCl prior to SDS-PAGE analysis), acidified acetone/methanol (50/50 v/v; useful to simultaneously remove acetone and methanol soluble interferences such as SDS prior to IEF analysis) and chloroform/methanol (50/50 v/v; useful to simultaneously remove salt and detergents).

24. The presence of trace amounts of TCA, carried through from the precipitation, can acidify the resuspension sample buffer. If further downstream processes are pH sensitive the sample buffer should be titrated with 1 *M* NaOH or 1 *M* Tris-HCl, pH 8.5, to obtain the desired pH for the required process. Acidified SDS-PAGE sample buffer, for example, can give a yellow colour. Correct titration will result in reversion to the typical blue sample buffer colour. Hint: A simple method to overcome this is to resuspend the samples in a slightly basic SDS-PAGE loading buffer (e.g. pH 9.0). Hence, any residual TCA left it will be neutralized by the basic buffer allowing direct addition onto the SDS-PAGE gel. Additionally, excess TCA traces will cause the Coomassie dye to precipitate during SDS-PAGE. If this is a problem repeat the optional wash steps outlined.

25. Recently Three Phase Fractionation, and derivatives such as Ionic liquid-based three phase partitioning, have been cited as effective alternatives both singularly **(*30*)** or as part of a combination approach such as preparative crystallisation **(*31*)**. Three Phase Partitioning utilises saturated, buffered ammonium sulphate in combination with an equal volume of water miscible aliphatic alcohol, commonly *t*-butanol. The two solutions do not mix and a heterogeneous, two-phase liquid is produced that can be used to partition proteins from whole cells without the need

to pre-lyse the cells. This process can be readily scaled-up for larger scale protein production and precipitation.

26. Although *t*-butanol is miscible with water, it is insoluble in aqueous solutions having high concentrations of salt, especially ammonium sulphate solutions. Three phases separate during centrifugation. The upper phase contains t-butanol, while the precipitate settles below the organic layer.

27. Expression of insoluble recombinant proteins has been noted as greater than 30% of all recombinantly expressed proteins **(*32*)**.

28. The interactions between solvents and proteins, and also proteins and proteins, determine the solubility of any given protein. Interactions can be classified as either attractive or repulsive. A protein will be soluble in a particular solvent if the net free energy of the proteins interactions is adequately negative (i.e. attractive). Additionally, protein solubility is improved if protein-protein interactions have sufficiently positive net free energy (i.e. repulsive), although it should be noted that protein-protein interaction is modulated by the chemical nature of the solvent of choice. Conversely, insolubility typically results from net attractive forces between proteins and net repulsive forces between the solvent of choice and the protein of interest. Furthermore, a soluble protein can be insolubilised by a change in its free energy state in relation to the proteins, or the solvent, it interacts with and, hence, additions/subtractions to a protein solution should be carefully assessed on a small scale **(*33*)**.

29. The separation of one protein, or family of proteins, from other proteins by means of differential solubility with chemical reagents is based on the differential solubility between a

liquid phase and a solid phase. The optimisation of this procedure is empirical but, Lindwall and colleagues **(*34*)** outline an optimisation procedure based on a sparse matrix approach; solubilisation buffers are composed based on "*solubility space*" which is related to accepted protein solubilisation theories. This method assists in identifying suitable solubilisation conditions for most over-expressed proteins (*see* **Note 30**).

30. Non-ionic detergents such as Triton X-100 are less effective in solubilizing hydrophobic proteins **(*35*).** An initial detergent screen may be required and can be performed to determine the optimum conditions for both solubilisation efficiency and maintenance of protein function. For example, during solubilisation a membrane protein is transferred from its natural environment to a buffered detergent solution. Membrane proteins require detergents for solubilisation during isolation as well as to maintain solubility.

31. It is important to remove existing aggregates that can act as nuclei to trigger aggregation during folding.

32. In brief, a typical membrane preparation protocol is as follows. Wash cells of interest three times with ice cold PBS, and then pellet at 500 x *g* for 10 min at 4°C. Discard the supernatant and resuspend the cells in 2 mL hypotonic buffer (10 m*M* HEPES [pH 7.9], 1.5 m*M* $MgCl_2$, 10 m*M* KCl, 0.5 m*M* DTT, 1X protease inhibitor cocktail containing 50 m*M* NaF, 2 m*M* Na_3VO_4, 1 m*M* $Na_4P_2O_7$). Sonciate cells on ice for 1 min. Spin the lysate at 3,000 x *g* for 15 min at 4 °C, to remove cell debris. Transfer the supernatant to an Optiseal Polyallomer tube (Beckman 362185) and fill the tube with the lysate rich hypotonic buffer (fill volume 5 mL) and centrifuge at 100,000 x *g* for 1 h at 4 °C (in a NVT90 rotor using the Optima L-100 XP Ultracentrifuge, Beckman Coulter). After centrifugation, remove the supernatant and resuspend the membrane

pellets in 50 µL MES buffer (25 m*M* MES, 150 m*M* M NaCl, 1 % Triton X-100, pH 6.5). Finally, to aid complete solubilisation, sonicate the sample, on ice, for 10 sec.

33. The amount of time necessary for solubilisation may be less than this and can be optimized to 1 h once an appropriate detergent has been selected.

34. Detergent screening can be further facilitated if the recombinant protein of interest is tagged with GFP.

35. The final protein concentration should not exceed 0.05 to 0.1 mg/mL as dilute protein mixtures refold optimally at this concentration. A rapid and efficient mix is essential at this step. Constant literature reviews of the fundamental refolding processes are available; see **(*36*)** and **(*37*)** for comprehensive and accessible reviews of refolding solubilised proteins. Cited protocols range from empirical to rationally designed **(*38*)**. Alternatively, an open source collection of protocols, *REFOLD*, is available at http://refold.med.monash.edu.au/. These are practitioner derived standard protocols supplemented with hands-on practical considerations **(*39*)**.

36. The addition of a mild solubilising agent [e.g. 1 *M* 3(1-pyridinio)-1-propane sulfonate] during the refolding steps limits re-aggregation of re-folding proteins. For further information on this, Tsumoto and co-workers **(*40*)** compressively detail the effect of small molecule additives on refolding and aggregation of the proteins along with other considerations for refolding recombinant proteins.

37. Continual dialysis can be set up by using a gradient maker. In its simplest form, this consists of two containers of the same shape connected by a siphon. One container contains the low

concentration buffer, and the other contains high concentration buffer. The buffer is withdrawn from the low concentration container to the high concentration container. This will produce a linear gradient from high to low buffer concentrations over the total volume of the gradient. Once the "*low concentration buffer*" supply has been depleted, the dialysis tubing is removed from the larger vessel and placed in a similar, clean vessel containing fresh buffer at the same concentration as the original "*low concentration buffer*". The "*low concentration buffer"* vessel is replaced with a vessel containing buffer at the next lower concentration level and the process is allowed to continue until the "*low concentration buffer*" supply is depleted again. This process is repeated until the vessel containing the dialysis tubing has reached the desired final concentration, typically 0 *M* GuHCl.

38. See Rudolph and Lilie **(*41*)** for a comprehensive overview of protein refolding and **(*42*)** for an industrial viewpoint on that topic.

39. Most proteins can be stored at 4°C, without significant denaturation, for up to 24 h. For intermediate storage times (24 h to one week) the protein should be filter sterilised (through a 0.22 μm filter) and stored at 4°C. Additional supplements, such as a bacteriostatic agent (e.g. 0.1% sodium azide) can be included to avoid bacterial growth. For storage times greater than one week (up to several months) it is advisable to freeze the protein preparation. Rapid freezing helps reduce protein denaturation. It is useful to freeze the solution in small aliquots to avoid repeated freeze/thaw cycles which may reduce the biological activity of the protein. Additional stabilizing agents can also be added prior to freezing, such as glycerol (5-50% w/v), serum albumin (10 mg/mL), reducing agents (such as 1 m*M* DTT), and ligands/co-factors (depending on the nature of the target protein). Extended protein storage (several months to years) should be carried out at

-80°C or in liquid nitrogen. The addition of 50% (w/v) glycerol is recommended for storage at this temperature. Alternative strategies include storing the protein as an ammonium sulphate precipitate at 4°C, or at lower temperatures in a lyophilized form (*see also* **Chapter 9** for protocols and discussion regarding the storage and lyophilisation of proteins).

References

1. O'Fágáin, C. (1997) Protein stability and its measurement, in *Stabilising protein function* (O'Fágáin, C., ed.), Springer Press, Berlin, pp 69-75.

2. Ramos, Y., García, Y., Llopiz, A., and Castellanos-Serra, L. (2008) Selectivity of bacterial proteome fractionation based on differential solubility: A mass spectrometry evaluation *Anal. Biochem*. **377**, 134–140.

3. Leimgruber, R. M. (2005) Extraction and solubilisation of proteins for proteomic studies, in *The Proteomics Protocols Handbook* (Walker J. M., ed.), Humana, Totowa, NJ, pp 1-18.

4. Habibi, N., Hashim, S. Z. M., Norouzi, A., and Samian, M. R. (2014). A review of machine learning methods to predict the solubility of overexpressed recombinant proteins in Escherichia coli. *BMC Bioinformatics*, **15**, 134.

5. Smialowski, P., Doose, G., Torkler, P., Kaufmann, S., and Frishman, D. (2012). PROSO II–a new method for protein solubility prediction. *FEBS Journal*, ***279***, 2192-2200.

6. Hirose, S., and Noguchi, T. (2013). ESPRESSO: a system for estimating protein expression and solubility in protein expression systems. *Proteomics*, **13**, 1444-1456.

7. Villalobos, A., Ness, J. E., Gustafsson, C., Minshull, J., and Govindarajan, S. (2006). Gene Designer: a synthetic biology tool for constructing artificial DNA segments. *BMC Bioinformatics*, **7**, 285.

8. Hays, F. A., Roe-Zurz, Z., and Stroud, R. M. (2010). Overexpression and purification of integral membrane proteins in yeast. *Methods in Enzymology*, **470**, 695-707.

9. Zheng, H., Miyakawa, T., Sawano, Y., Yamagoe, S., and Tanokura, M. (2013). Expression, high-pressure refolding and purification of human leukocyte cell-derived chemotaxin 2 (LECT2). *Protein Expression and Purification*, **88**, 221-229.

10. Dworeck, T., Petri, A. K., Muhammad, N., Fioroni, M., and Schwaneberg, U. (2011). FhuA deletion variant Δ1-159 overexpression in inclusion bodies and refolding with Polyethylene-Poly (ethylene glycol) diblock copolymer. *Protein Expression and Purification*, **77**, 75-79.

11. Rane, A. M., Jonnalagadda, S., and Li, Z. (2013). On-column refolding of bone morphogenetic protein-2 using cation exchange resin. *Protein Expression and Purification*, **90**, 135-140.

12. Yamaguchi, H., and Miyazaki, M. (2014). Refolding techniques for recovering biologically active recombinant proteins from inclusion bodies. *Biomolecules*, **4**, 235-251.

13. Pace, C. N., Grimsley, G. R., and Scholtz, J. M. (2009). Protein ionizable groups: pK values and their contribution to protein stability and solubility. *Journal of Biological Chemistry*, **284**, 13285-13289.

14. Trevino, S. R., Scholtz, J. M., and Pace, C. N. (2008). Measuring and increasing protein solubility. *Journal of Pharmaceutical Sciences*, **97**, 4155-4166.

15. Dehouck Y, Kwasigroch JM, Gilis D, Rooman M. (2011). PoPMuSiC 2.1: a web server for the estimation of protein stability changes upon mutation and sequence optimality. *BMC Bioinformatics*, **12**, 151.

16. Jonasson, P., Liljeqvist, S., Nygren, P. A., and Ståhl, S. (2002). Genetic design for facilitated production and recovery of recombinant proteins in Escherichia coli. *Biotechnology and Applied Biochemistry*, **35**, 91-105.

17. Chen, Z. Y., Cao, J., Xie, L., Li, X. F., Yu, Z. H., and Tong, W. Y. (2014). Construction of leaky strains and extracellular production of exogenous proteins in recombinant *Escherichia coli*. *Microbial Biotechnology*, **7**, 360-370.

18. Yoon S.H., Kim S.K., and Kim J.F. (2010). Secretory production of recombinant proteins in *Escherichia coli*. *Biotechnology*, **4**, 23–29.

19. Jonet, M. A., Mahadi, N. M., Murad, A. M. A., Rabu, A., Bakar, F. D. A., Rahim, R. A., and Illias, R. M. (2012). Optimization of a heterologous signal peptide by site-directed mutagenesis for improved secretion of recombinant proteins in *Escherichia coli*. *Journal of Molecular Microbiology and Biotechnology*, **22**, 48-58.

20. Fu, X. Y. (2010). Extracellular accumulation of recombinant protein by Escherichia coli in a defined medium. *Applied Microbiology and Biotechnology*, **88**, 75-86.

21. Leibly, D. J., Nguyen, T. N., Kao, L. T., Hewitt, S. N., Barrett, L. K., and Van Voorhis, W. C. (2012). Stabilizing additives added during cell lysis aid in the solubilization of recombinant proteins. *PLoS One*, **7**, e52482.

22. French, C.; Keshavarz-Moore, E., Ward, J.M. (1996) Development of a simple method for the recovery of recombinant proteins from the *E.coli* periplasm. *Enzyme Microb. Technol.* **19**, 332–8.

23. Caldwell R.B., and Lattemann C.T. (2004) Simple and reliable method to precipitate proteins from bacterial culture supernatant. *Appl. Environ. Microb*. **70**, 610-12.

24. Islam, M. N., Zhang, M., and Adhikari, B. (2014). The inactivation of enzymes by ultrasound: a review of potential mechanisms. *Food Reviews International*, **30**, 1-21.

25. Özbek, B., and Ülgen, K. Ö. (2000). The stability of enzymes after sonication. *Process Biochemistry*, **35**, 1037-1043.

26. Feliu, J. X., Cubarsi, R., and Villaverde, A. (1998). Optimized release of recombinant proteins by ultrasonication of *E. coli* cells. *Biotechnology and Bioengineering*, **58**, 536-540.

27. Doonan, S. (2004) Bulk Purification by Fractional Precipitation, in *Protein Purification Protocols, Methods in Molecular Biology* (Cutler, P., ed.), Humana, Totowa, NJ, **244**, pp 117-125.

28. Balasundaram, B., Sachdeva, S., and Bracewell, D. G. (2011). Dual salt precipitation for the recovery of a recombinant protein from *Escherichia coli*. *Biotechnology Progress*, **27**, 1306-1314.

29. Burgess, R. R. (2009). Protein precipitation techniques. *Methods in Enzymology*, **463**, 331-342.

30. Hekmat, D., Maslak, D., von Roman, M. F., Breitschwerdt, P., Ströhle, C., Vogt, A., Berensmeier, S., and Weuster-Botz, D. (2015). Non-chromatographic preparative purification of enhanced green fluorescent protein. *Journal of Biotechnology,* **194**, 84-90.

31. Hekmat, D., Breitschwerdt, P., and Weuster-Botz, D. (2015). Purification of proteins from solutions containing residual host cell proteins via preparative crystallization. *Biotechnology Letters*, 1-11.

32. Leibly, D. J., Nguyen, T. N., Kao, L. T., Hewitt, S. N., Barrett, L. K., and Van Voorhis, W. C. (2012). Stabilizing additives added during cell lysis aid in the solubilization of recombinant proteins. *PLoS ONE*, **7**, e52482.

33. Rothstein F. (1994) Differential Precipitation of Proteins, in *Protein Purification Process Engineering* (Harrison, R. G. ed), Marcel Dekker Inc., NY, pp 115-16.

34. Lindwall, G., Chau, M.-F., Gardner, S.R., and Kohlstaedt, L.A. (2000) A sparse matrix approach to the solubilisation of overexpressed proteins. *Protein Eng.* **13**, 67-71.

35. Luche S., Santoni, V., and Rabilloud, T. (2003). Evaluation of nonionic and zwitterionic detergents as membrane protein solubilizers in two-dimensional electrophoresis. *Proteomics* **3**, 249–253.

36. Cabrita, L. D., and Bottomley, S. P. (2004). Protein expression and refolding–a practical guide to getting the most out of inclusion bodies. *Biotechnology Annual Review*, **10**, 31-50.

37. Yamaguchi, H., and Miyazaki, M. (2014). Refolding techniques for recovering biologically active recombinant proteins from inclusion bodies. *Biomolecules*, **4**, 235-251.

38. Basu, A., Li, X., and Leong, S. S. J. (2011). Refolding of proteins from inclusion bodies: rational design and recipes. *Applied Microbiology and Biotechnology*, **92***, 241-251.*

39. Phan, J., Yamout, N., Schmidberger, J., Bottomley, S. P., and Buckle, A. M. (2011). Refolding your protein with a little help from REFOLD. *Methods in Molecular Biology*,**752**, 45-57.

40. Tsumoto, K., Ejima, D., Kumagai, I., and Arakawa, T. (2003). Practical considerations in refolding proteins from inclusion bodies. *Protein Expression and Purification*, **28**, 1-8.

41. Rudolph, R., and Lilie, H. (1996) In vitro folding of inclusion body proteins. *FASEB* **10**, 49-56.

42. Jungbauer, A., and Kaar, W. (2007) Current status of technical protein refolding. *J. Biotechnol.* **128**, 587-96.

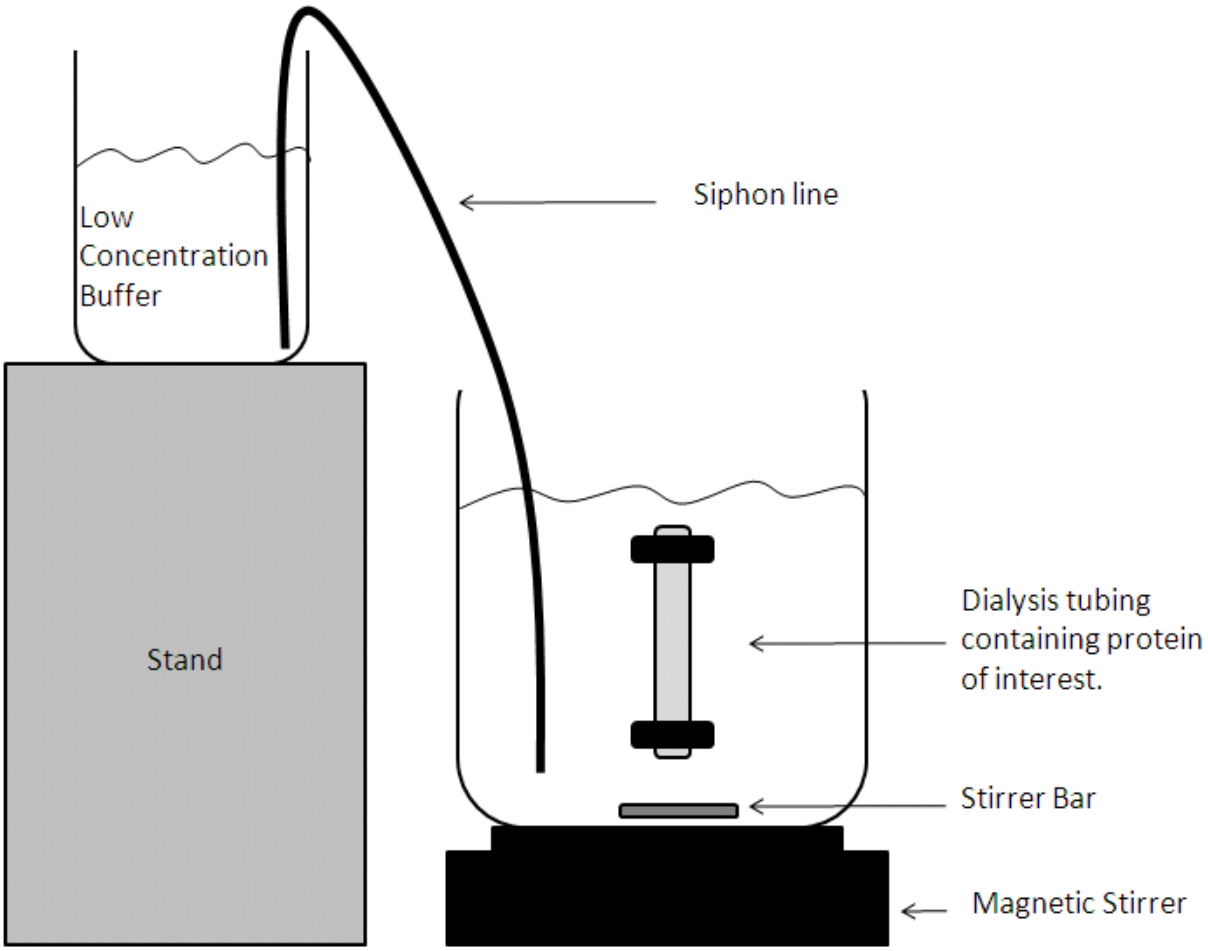

Figure legend

Figure 1: Schematic diagram of a simple gradient maker.