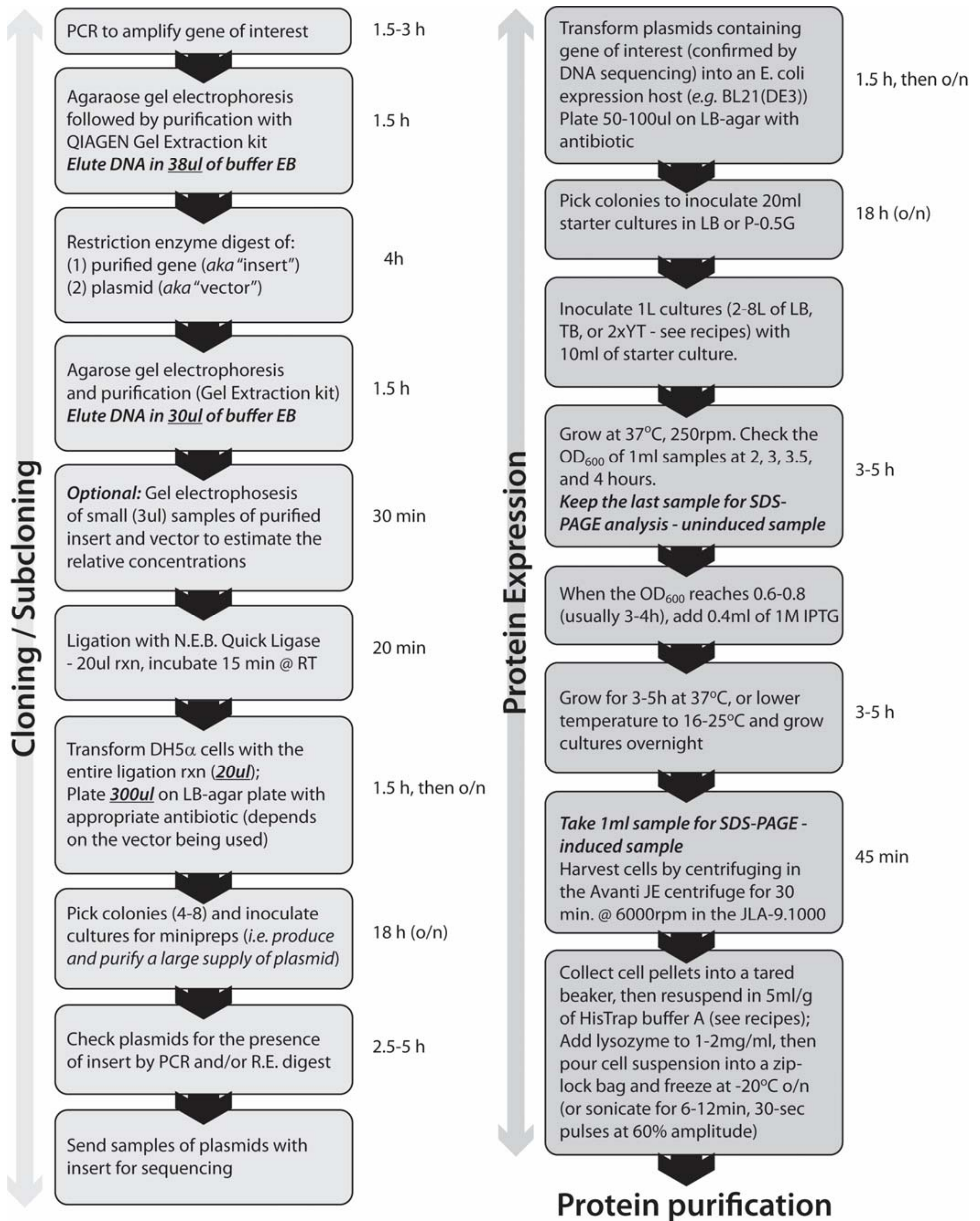
## Introduction: Overview

Success in any structural biology project depends on a consistent supply of pure, stable protein. Thus, most of the work done in this lab revolves around expressing and purifying proteins of interest. As the flow-chart below shows, the process begins with cloning or subcloning a gene into a suitable expression vector. Once the gene has been successfully incorporated into a plasmid, it can be ever-expressed in *E. coli* cells, and hopefully, purified to homogeneity (> 95% pure). This workflow can break down at any number of points, and so is necessarily an iterative process. Genes can be difficult to clone, protein might express in a misfolded, insoluble form, or not at all. It is not unusual to subclone a gene into 4 or 5 expression vectors with different properties to find one that gives high-level expression of the target protein.

The protocols in this guide represent but one possible way to accomplish a given task. Look at this guide as a starting point. It is more than likely that you will need to try multiple strategies before you are successful.


Cloning / Subcloning
PCR to amplify gene of interest
1.5-3 h
Agaraose gel electrophoresis followed by purification with QIAGEN Gel Extraction kit
Elute DNA in 38ul of buffer EB
1.5 h
Restriction enzyme digest of:
(1) purified gene (aka "insert")
(2) plasmid (aka "vector")
4h
Agarose gel electrophoresis and purification (Gel Extraction kit)
Elute DNA in 30ul of buffer EB
1.5 h
Optional: Gel electrophosesis of small (3ul) samples of purified insert and vector to estimate the relative concentrations
30 min
Ligation with N.E.B. Quick Ligase - 20ul rxn, incubate 15 min @ RT
20 min
Transform DH5α cells with the entire ligation rxn (20ul);
Plate 300ul on LB-agar plate with appropriate antibiotic (depends on the vector being used)
1.5 h, then o/n
Pick colonies (4-8) and inoculate cultures for minipreps (i.e. produce and purify a large supply of plasmid)
18 h (o/n)
Check plasmids for the presence of insert by PCR and/or R.E. digest
2.5-5 h
Send samples of plasmids with insert for sequencing
Protein Expression
Transform plasmids containing gene of interest (confirmed by DNA sequencing) into an E. coli expression host (e.g. BL21(DE3))
Plate 50-100ul on LB-agar with antibiotic
1.5 h, then o/n
Pick colonies to inoculate 20ml starter cultures in LB or P-0.5G
18 h (o/n)
Inoculate 1L cultures (2-8L of LB, TB, or 2xYT - see recipes) with 10ml of starter culture.
Grow at 37°C, 250rpm. Check the OD600 of 1ml samples at 2, 3, 3.5, and 4 hours.
Keep the last sample for SDS-PAGE analysis - uninduced sample
3-5 h
When the OD600 reaches 0.6-0.8 (usually 3-4h), add 0.4ml of 1M IPTG
Grow for 3-5h at 37°C, or lower temperature to 16-25°C and grow cultures overnight
3-5 h
Take 1ml sample for SDS-PAGE - induced sample
Harvest cells by centrifuging in the Avanti JE centrifuge for 30 min. @ 6000rpm in the JLA-9.1000
45 min
Collect cell pellets into a tared beaker, then resuspend in 5ml/g of HisTrap buffer A (see recipes);
Add lysozyme to 1-2mg/ml, then pour cell suspension into a zip-lock bag and freeze at -20°C o/n
(or sonicate for 6-12min, 30-sec pulses at 60% amplitude)
Protein purification

# Section 1: Cloning / Subcloning and General Molecular Biology Protocols

**Preparing and running agarose gels**

*For a wide gel (15cm x 10cm), prepare 100ml of agarose solution. The percentage will depend on the size of the DNA fragments to be resolved, but 1.0 or 1.2% should be appropriate for most applications. The narrow gel (8 x 10cm) requires only 50ml of agarose solution.*

1. Weigh out the appropriate amount of agarose into a 500ml flask (*e.g.* 1g for a wide gel at 1.0% agarose)
2. Add the corresponding volume (50 or 100ml) of SybrSafe DNA gel stain solution (0.5X TBE)
3. Microwave the agarose solution for 30 seconds on high power. Watch to make sure it does not boil over. Remove and swirl the mixture. *BE CAREFUL – superheated liquids can splash up when shaken!* Continue to alternate short bursts in the microwave with careful swirling until all of the agarose is dissolved. There should be no little flecks of undissolved agarose floating around (they can be hard to see because they are clear).
4. Lock the casting tray (the clear acrylic part) into the apparatus (white plastic with red rubber seals at each end), level the assembly on the bench* and pour in the melted agarose solution.
5. Insert a comb with enough wells for your samples[§] and let the agarose cool and set (~30-60 min depending on temperature). Once set, remove the comb from the gel and remove the tray from the casting apparatus. DO NOT rush it. If the comb is removed too early, then wells will collapse.
6. Put the gel (still in the casting tray) into the electrophoresis apparatus.
7. Pour in enough 0.5X TBE to cover the gel by a couple of millimeters
8. Load samples and a marker (1Kb or 100bp ladder) and run the gel for 30-60 minutes at 130V. Watch the dye front for an indication of when to stop the gel.

** The gel can be poured in the 4°C cooler if you are in a time crunch.*

[§] *There are 15 or 20-well combs for the wide gel and 8 or 15-well combs for the narrow gel. The combs with wide teeth (i.e. the wide 15-well comb) can hold ~20-30ul of sample. If you have larger samples, you can place small pieces of tape over the gap between two (or three) teeth to make wider wells.*

**Preparation of chemically-competent *E. coli***

***Day 1:***

1. Use glycerol stock or fresh plate to inoculate 10ml of LB and grow in shaker overnight at 37°C. *If using a cell line with accessory plasmids (e.g. pLysS/E or RILP), be sure to include appropriate antibiotic markers to maintain the plasmid. Otherwise the cells will revert to 'regular' BL21(DE3) cells.*

***Day 2:***

1. Inoculate 50ml of LB with 0.5ml of the overnight culture
2. Grow cells to an $OD^{600}$ of 0.5 to 0.6 (not beyond!)
3. Pour the cells into a sterile 50ml Corning tube (orange caps) and pellet the cells by centrifuging for 10 min @ 3,000rpm, 4°C (Eppendorf 5810r)
4. Resuspend the cell pellet in 25ml of **sterile** 50mM $CaCl_2$ **gently** by pipeting (no vortex!). This step should be done **on ice.**
5. Let the cells rest, on ice, for *at least* 1 hour
6. Collect the cells again by centrifuging for 10 min @ 3,000rpm, 4°C
7. Resuspend the pellet in 4ml of 50mM $CaCl_2$ and 1.5ml of **sterile** 50% glycerol (15% final)
8. Aliquot the cells between 50 eppendorf tubes (100ul each) and freeze the cells at –80°C or use directly for transformation (use 50-100µl and 10µl of ligation reaction or 1µl of plasmid)

## Transformation of *E. coli* competent cells

** This protocol is for transformation of "home made" chemically competent* E. coli*. Use the manufacturer's protocol for commercially prepared cells – see the black binder of product inserts/instructions.*

1. Thaw cells on ice (~10 min)
2. To 100µl of chemically competent cells, add 0.5-1.0µl of purified plasmid DNA **or** 10-20µl of a ligation reaction (There is much less intact plasmid in a ligation reaction)
3. Incubate on ice for 15-30 min
4. Heat shock for exactly 30 sec at 42°C
5. Incubate on ice for 2 min
6. Add 750µl of sterile LB or SOC medium
7. Grow out for 1 hour at 37°C in the incubator (no need for shaking)
8. Concentrate the cells by spinning them in the microfuge for 1 min at 6,000rpm
9. Discard 750µl of the supernatant and resuspend the cells in the remaining 100µl
10. Plate the entire 100µl on an LB-Agar plate with appropriate antibiotic(s) and incubate overnight at 37°C

## PCR amplification of "regular" sequences

Success of PCR reactions depends primarily on the quality of the primers, so spend time with OLIGO (or any on-line primer design program) to design primers that conform to these rules:

i. Length must be greater than 18 bp
ii. G/C content should be between 40-50%; Sometimes it just has to be higher (see difficult cases below)
iii. Keep G/C content and melting temperature (Tm) about equal for the primer pair
iv. End primers with a 'CG' doublet; Some restriction enzymes (e.g. NdeI) need a bit of DNA flanking the recognition sequence in order to cut the DNA well. So I put the sequence 'GGAATTC' before my NdeI cut sites. I am not sure the exact sequence matters, but I got this particular sequence from someone who does PCR for a living and have never had a problem with it.
**v. When you have finished designing primers, check that (1) the reading frame of the gene will be correct when you cut the PCR product and ligate it into the plasmid vector and (2) that there are no internal (i.e. accidental/natural) cleavage sites for the restriction enzymes you plan to use for cloning. Do this before you order the primers!**

When you receive your primers, they will be lyophilized in plastic tubes. To reconstitute the primer, spin the tube in the microfuge for 1 minute at 13,000rpm to make sure that the lyophilized pellet (which will be tiny and may be invisible) does not go flying out the top of the tube when you take the cap off. Then figure out how much mq$H_2O$ you will need to add to reach a concentration of 30 pmole/µl (the tube will have the number of nanomoles written on the side, just multiply this by 1000 to convert to picomoles and then divide by 30. For example: 33.4nmol x 1000 = 33,400pmol / 30pmol/ul = 1113ul of mq$H_2O$).

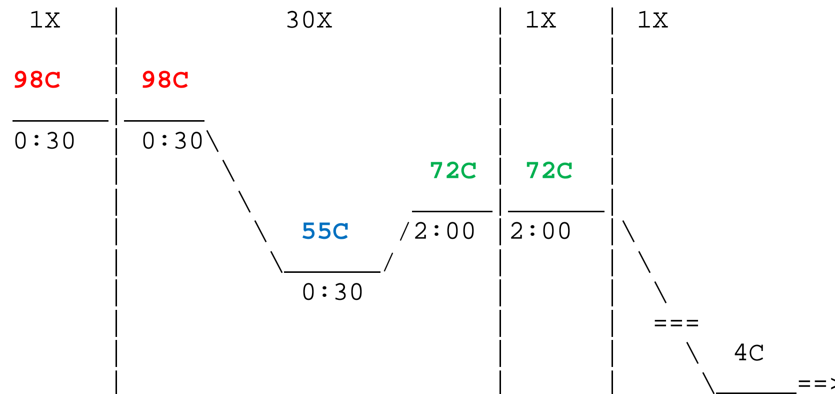
| Component | C1 | C2 | Vol/rxn | Vol/3rxn |
|---|---|---|---|---|
| Vector | ?ug/ml | 50ng | 1ul | 3ul |
| Forward Primer | 30pmol/ul | 25pmol | 1ul | 3ul |
| Reverse Primer | 30pmol/ul | 25pmol | 1ul | 3ul |
| dNTPs* | 10mM | 250uM | 1ul | 3ul |
| 10X DynaZyme Bfr** | 10X | 1X | 5ul | 15ul |
| mqH2O | - | - | 40ul | 120ul |
| Phusion DNA Pol.*** | | | 1ul | 3ul |
| | | | 50ul | 150ul |

*If using the 2.5mM dNTPs from Invitrogen, adjust this volume accordingly! (e.g. Use 4ul per reaction)

** Each brand/type of DNA polymerase has its own dedicated buffer. Some are 10X, some are 5X – make sure to adjust this volume as necessary.

*** The choice of polymerase depends on what type of gene you are working on. For most applications both Phusion (aka Pfu) or Dynazyme are appropriate choices (high fidelity). For very long sequences (see the section on difficult sequences below) or for special cloning strategies (e.g. T/A cloning) there are better choices than Pfu (e.g. LongAmp for long PCR products).

Use a PCR program like this one:




The temperature and duration of the melting phase (red) will change depending on the polymerase used. The temperature of the annealing phase (blue) should be set 5°C below the melting temperature of the primer pair. I like to use the computer program "OLIGO" because it suggests an optimum $T_A$ that I have come to trust. The temperature and duration of the elongation phase (green) should be set according to the polymerase instructions and will increase with the length of the sequence to be amplified. A general rule is 2 minutes minimum, then add 1 min per 1000 bp.

**PCR amplification of challenging sequences**

Challenging sequences are those that are very long (more than 2-3 kbp), have very high G/C content, form secondary structures, or are very repetitive (e.g. most NRPS genes!). These features cause problems by (1) knocking the polymerase off of the DNA before it finishes amplifying the strand or (2) causing amplification of a number of non-specific fragments.

There is no easy way to deal with difficult template sequences. Simple things that can help are:

i. Use a polymerase engineered for long and/or high G/C sequences (Herculase from Agilent/ Stratagene or Pfu Ultra II from Invitrogen)
ii. Optimize your PCR program – especially the annealing temperature (Ta)
iii. Try additives – DMSO, glycerol, PCRx Enhancer (Qiagen), **Phusion GC buffer**
iv. Do the PCR in two or more steps: (1) design primers complimentary to some unique sequences flanking your gene to get a non-repetitive template to work from, then (2) use a second set of primers complimentary to the start and end of the gene to amplify the region of interest .
v. Very large sequences can also be cloned in two separate pieces: clone the first half (for example) of the gene, then once that is successfully cloned into a suitable vector, clone the second half of the gene into that newly-created plasmid to generate the complete sequence.

**Restriction digest of plasmids and PCR products**

Once the gene has been amplified by PCR, the PCR reactions are run out on a gel to visualize the results. Bands on the gel with the correct size (in bp) are cut out and the DNA extracted from the agarose using the QIAGEN Gel Extraction kit. The final step in the procedure is to elute the DNA from the spin column(s). Use 38ul of buffer EB for this step, since it will facilitate the restriction digest procedure (below).

| Component | Insert | Vector |
|---|---|---|
| Plasmid (e.g. pET-15b) | -- | 22.0ul |
| Inserts | 36ul | -- |
| 10X BSA (1mg/ml)[a] | 5.0ul | 3.0ul |
| 10X N.E.B. Buffer (1X Final)[b] | 5.0ul | 3.0ul |
| Enzyme 1[c] | 2.0ul (40U) | 1.0ul (20U) |
| Enzyme 2[c] | 2.0ul (40U) | 1.0ul (20U) |
| | 50.0ul | 30.0ul |

[a] Not all restriction enzymes require BSA – check the product insert or the New England Biolabs web site.

[b] Check the N.E.B. double digest chart to make sure that the two enzymes to be used in the digest are compatible. If not, then a sequential digest will be necessary (cut with first enzyme, gel purify, then cut with second enzyme and gel purify again).

*Some common double digest combinations:*

| | |
|---|---|
| **NdeI/BamHI-HF:** | NEB buffer **3** + BSA |
| **NdeI/XhoI:** | NEB buffer **4** + BSA |
| **NcoI/XhoI:** | NEB buffer **4** + BSA |
| **BsaI-HF/XbaI:** | NEB buffer **4** + BSA |

[c] Double-check the concentrations of the enzyme stocks and adjust these volumes accordingly.

Incubate the reactions for 1-4h in the 37°C water bath, then run them out on a 1-1.2% agarose/TBE gel. Be sure to run a small sample of uncut plasmid as well (e.g. 3ul uncut pET-15b + 3ul loading dye) to confirm that the digest worked well.

**N.E.B. Quick Ligation**

The restriction digest reactions are run out on a gel and the bands with the correct size are cut out and purified using the QIAGEN Gel Extraction kit. Use a final elution volume of 30μl to get the most concentrated solution possible. Run both the digested insert (PCR product) and vector (plasmid) on a gel just long enough to see the two bands (~20 min). Compare the relative intensities by eye in order to judge how much of each solution (insert and vector) should be added to the ligation reaction. Or, if time is short, just to both 1:1 and 1:3 ligation reactions to cover your bases.

| Component: | 1:1 | 1:3 | Control |
|---|---|---|---|
| Plasmid digest | 1ul | 1ul | 1ul |
| Insert digest | 1ul | 3ul | ~ |
| Quick Lig bfr (2X) | 10ul | 10ul | 10ul |
| T4 Quick ligase | 1ul | 1ul | 1ul |
| mqH2O | 7ul | 5ul | 8ul |
| | 20ul | 20ul | |

Incubate the ligase reactions at RT for 15 minutes and then use 10-20μl of each one to transform aliquots of high efficiency (i.e. store-bought) DH5α cells, like NEB5α. Pick several colonies from the plate(s) – normally 4-8 colonies are sufficient – and inoculate 10-20ml cultures for mini-preps. The mini-prepped plasmid DNA is then checked for the presence of the insert by sequencing or by digesting 10-20μl of the mini-prepped

plasmid with the same enzymes used for the cloning. When run out on a gel, the digested mini-prep plasmid will give a band with the same size as the insert. Obviously, no band means that either the digest reaction or the ligation reaction didn't work well enough. Alternatively, one can also do PCR reactions with the same primers used to do the original cloning. In this case, you need to run two extra reactions: one with a plasmid known to carry the gene of interest (pisitive control) and one known to lack the gene (negative control). Normally the templates for these reactions are the plasmid used in the original PCR reaction and the empty vector.

**Keep all of the clones that have the insert—Some may not have the right sequence (mutations, for example) or you may need to test them all for expression! Send at least one clone for sequencing to confirm that the gene is actually there and does not include any unwanted mutations.**

## Section 2: Protein Expression Protocols

### Selection of *E. coli* strains and expression vectors

The simplest, first-attempt approach to protein expression is outlined in the flow-chart on page 2. After subcloning a gene into a vector like pET-15b, the new plasmid (gene sequence confirmed by DNA sequencing) is put into an expression strain, usually BL21(DE3). A single colony is picked off of an LB-agar plate using a sterile pipet tip and used to inoculate a 20ml starter culture (eject the tip into 20ml of sterile LB + antibiotic). After growing to stationary phase overnight, this starter culture is used to inoculate larger (0.5 or 1L) expression cultures. The density of the culture is monitored by following the optical density at 600nm. An $OD_{600}$ of 0.6-0.8 indicates that the culture is in mid-log phase, and it is at this point that the inducer IPTG is added, normally at a final concentration of 1mM (1ml of 1M IPTG stock in a 1L culture). The cells are allowed to grow for 3 to 5 hours after induction at 37°C, after which they are harvested by centrifugation, lysed, and the protein purified.

**Once you find a clone that expresses the protein well, prepare a glycerol stock for long-term storage at -80°C! Simply mix 1ml of culture (stationary phase culture in P-0.5G medium is best) with 0.5ml of sterile 50% glycerol and put the labeled tube (include vector, insert, medium, antibiotic, and date) in the -80°C freezer.**

The most common expression vector, pET-15b, appends six histidine residues ($His_6$-tag) to the N-terminus of the target protein. This $His_6$-tag allows for convenient purification of the target protein by immobilized metal affinity chromatography. For many proteins, a nickel-containing column (*e.g.* HisTrap, GE LifeSciences) and a size exclusion column are enough to take the target protein from crude cell extract to >95% purity.

### Reduced-Temperature Induction Protocol

....

### Autoinduction Protocol

....

### Osmotic Stress Protocol

....

### Expression and Purification of Selenomethionine-Substituted Protein for SAD/MAD Phasing:

*The detailed instructions from Molecular Dimensions, who produce and distribute SelenoMet™ Medium, are in the binder with the rest of the product inserts and can be found electronically in the protocols folder on the W drive or on the Molecular Dimensions web site at (http://www.moleculardimensions.com/applications/upload/SeMet.pdf). What follows here are brief, practical hints from my notes—if you have questions, refer to the "official" instructions.*

Once you have (1) sequenced an expression plasmid and found no mutations, (2) expressed the protein in soluble form, and (3) crystallized the protein and demonstrated at least 2.5-3.0Å diffraction, you can try to express selenomethionine (SeMet)-labeled protein for phasing by MAD/SAD. The procedure begins with transformation of the confirmed expression vector into a methionine auxotroph strain like *E. coli*834(DE3), or T7 Express Crystal from NEB. The cells are transformed exactly like normal *E. coli* strains, and plated on plain LB-agar plates.

Prepare 2-4L of the SeMet medium base (21.6g/L) in 2L Erlenmeyer flasks and 250ml of the medium base in a 500ml bottle (for starter cultures). Autoclave these. Once they have cooled, mix up one 50ml aliquot of the nutrient mix for each flask/bottle. Sterilize the nutrient mix by filtering through a **sterile** 0.2μm

syringe filter into a fresh, **sterile** 50ml tube. Add the sterile nutrient mix to the sterile SeMet base at a rate of 50ml/L. So, dump the whole tube into the 1L portions, and 12.5ml into the 250ml portion. Leave the 2L flasks in the 4°C overnight.

To the 250ml of SeMet medium, add 1ml of the 250X L-Met solution (stored in the -80°C, thaw in a water bath if necessary). Pick colonies from the T7 Express Crystal plate and use them to inoculate **two** 20ml starter cultures per liter of expression culture (e.g. for a 4L expression, you will have **8** starters). Use the SeMet medium containing L-Met. Grow these overnight at 37°C.

The next day, spin the starter cultures down at 4,000rpm for 10 minutes. While this is happening, add 4ml of 250X L-SeMet stock (also stored at -80°C) and 1ml of the appropriate antibiotic stock to each flask of SeMet medium. When the cells finish spinning, pour of the supernatant. **Be careful**—the pellets may be very loose and you may pour off some/all of the cells if you pour too fast. Resuspend the cells in 10ml of sterile water and spin them a second time. Repeat this procedure once more. The point of doing these three washes is to remove all traces of L-Met from cells so they are forced to use the SeMet during expression. After the third wash, use the cells (now in sterile $H_2O$) to inoculate the expression cultures and incubate at 37°C with 250rpm shaking. The cells will grow **very** slowly. The last time I did this, it took the cultures **9 hours** to reach an $OD_{600}$ of ~0.5! When the cultures reach an $OD_{600}$ of 0.5-0.6 (or the end of the day, whichever comes first), add 0.4ml of 1M IPTG and let the cells grow overnight. If the temperature is normally dropped during expression, do that here too. For example, MppR was expressed overnight in SeMet medium at 25°C. The rest of the process is exactly the same as for "normal" protein: Resuspend the cells in the usual buffer (e.g. HisTrap buffer A), lyse by sonication or freeze-thaw, purify, concentrate and (hopefully) crystallize. Keep in mind that (1) the cells are not very robust and (2) they are growing in minimal medium. So expect low cell masses (I last had 2.5g of cells per liter of culture) and low yields of protein. This is normal and it is fine since we only need **one** crystal to diffract well and give us a decent data set for phasing. Once the structure is solved, all subsequent crystallographic experiments will be done will "normal" protein.

**When Expression Fails:**

As alluded to above, there is a lot that can go wrong with the "first-attempt" protocol for general protein expression. The most common problems are:

1. The target protein is not expressed.
2. The target protein is expressed, but insoluble (inclusion bodies).
3. The target protein is soluble, but will not bind to the Ni column (*i.e.* the $His_6$-tag does not work for purification).
4. The target protein can be expressed in soluble form and purified, but the final protein is unstable and/or lacks its normal biological activity.

*Case 1: No detectable expression.*

When no protein is produced, it often means that the target protein is toxic to the cells, creating a powerful selective pressure for *E. coli* cells to "spit out" the offending gene. Lack of expression can also be due to mistakes in cloning (incorrect reading frame is a common one), rare codons, short mRNA lifetime or stable mRNA secondary structures, short protein lifetime (rapid proteolysis), and probably a host of other reasons.

The first thing to do when expression fails completely is to repeat the experiment with several more clones (different colonies from the original ligation plate) and take 10ml or so of the expression culture at the end to do a miniprep. Analysis of the plasmid thus obtained will likely indicate that the gene is no longer in the plasmid or that there was some mistake with the cloning (wrong reading frame, accidental

stop codon). The latter is just human error and is easily fixed. The former indicates that the gene is toxic, and is more difficult to deal with. Using *E. coli* strains that eliminate the basal expression (i.e. expression of the protein prior to induction) can help relieve the selective pressure against maintaining the plasmid (BL21 pLysS or pLysE). By the same token, switching from the notoriously leaky T7 promoter to something else (AraBad, cold-shock promoter) can also work. Changing the location of expression is another approach. For example, putting the gene into a vector that directs expression to the periplasmic space or purposely pushing the protein to express as insoluble inclusion bodies can both be effective solutions. As a last resort, one may need to change to a completely different system, like yeast (*e.g. Pichia pastoris* or *Kluyveromyces lactis*).

Assuming the cloning was done properly and the gene is not toxic, the next most likely problem is the presence of so-called "rare codons." Certain codons are rarely used by *E. coli*. The most problematic are AGG, AGA, and CGA (arginine), CTA (leucine), ATA (isoleucine), and CCC (proline). Having two or more consecutive rare codons, like AGA-AGA, almost invariably prevents over-expression. Sequences can be run through various servers to detect rare codons (http://nihserver.mbi.ucla.edu/RACC/ or http://www.genscript.com/cgi-bin/tools/rare_codon_analysis). If rare codons appear to be a problem, there are a few fixes. First, if there are only a few, then one can simply mutate them to match the *E. coli* codon bias. If the sequence is riddled with rare codons, as most Streptomyces genes are, then the only options are to (a) have the gene synthesized with a DNA sequence optimized for *E. coli* or (b) find a homolog from a different organism that doesn't have rare codons.

It is difficult to detect, let alone fix, other obstacles to expression. The best choices are to have the gene synthesized, optimizing GC content, secondary structure properties, etc or to find a homolog that behaves better in *E. coli*. Often finding a bacterial homolog of a eukaryotic protein works well. The problem with this approach is persistent criticism that there might be substantial structural and/or functional differences between the prokaryotic and eukaryotic versions of the protein.

## Section 3: Protein Purification Protocols

**Cell lysis by sonication**
**Cell lysis by conventional freeze-thaw**
**His-Trap (Ni affinity) Protocol**
**GS-Trap (glutathione affinity) Protocol**
**General Ion Exchange Protocol**
**Preparative Size Exclusion Protocol**
**Removing Tags with Proteases**
**SUMO Protease (Ulp1)**

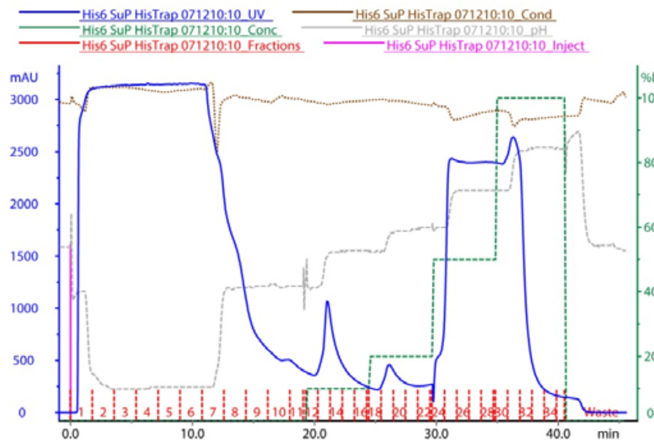
**I. SUMO-Fusion protein expression and purification**

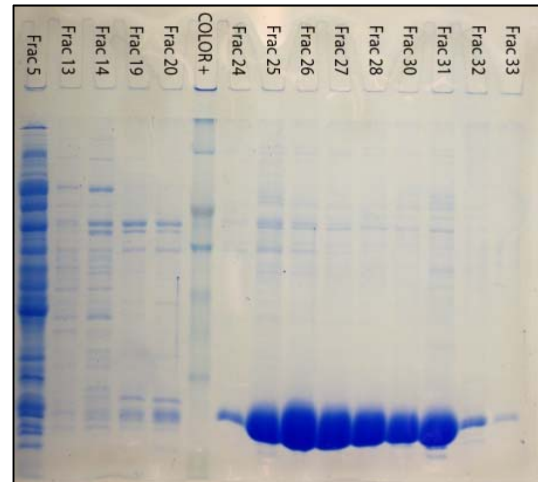
I use the pE-SUMO$_{kan}$ vector from LifeSensors, Inc. It comes as a "regular" plasmid (i.e. not linearized for TA/TOPO cloning like the Champion pET-SUMO system from Invitrogen) , so you can buy it once and make your own renewable supply. Expression of SUMO fusions from either of these vectors uses the standard pET-style protocol. I have found that (1) concentrations above 0.4mM IPTG don't lead to more protein (i.e. it's just wasteful) and (2) SUMO fusions often do best expressed overnight at 20-25°C. I don't think the latter has to do with protein stability, but rather increased cell mass leading to increased yield. Of course, some fusion partners will need the lower temperature for stability. The crude lysate from expression is purified over a HisTrap (or similar) column and the peak fractions are pooled and 1-3ml of SUMO protease stock (see below) at ~150µM are added directly to the pooled fractions. I just eyeball the amount of protease based on the amount of protein. The pooled fractions with SUMO protease are dialyzed overnight at 4°C or RT, depending on the protein. The next day I run the dialysate back over the Ni column to remove the protease and tag.

## II. SUMO Protease (Ulp1) expression and purification

The SUMO protease (SuP) is expressed from BL21star(DE3) cells containing the pET15b-SuP plasmid in the usual way (e.g. 2-4L LBamp grown at 37°C to $OD_{600}$ = ~0.8, induce with 0.4mM IPTG, grow o/n @ 25°C). Purify by HisTrap or similar. Normally the single column gets the protease pure enough. Dialyze or desalt the protein into 25mM TRIS pH 8.0, 150mM NaCl, 5mM DTT. We don't usually have to concentrate the protein at this stage—it is typically ~150-250 µM—we just dilute it 1:1 with 100% glycerol to snap-freeze in liquid $N_2$. When stored at -80°C, the protein will be active for at least 1 year.

```
Chromatogram Questions
No 1:   Sample Volume and Type
        55ml His6-SuP c.l. - lysis by sonication; H   isTrap A

No 2:   Column
        5ml HisTrap HP

No 3:   Eluent A
        25mM TRIS pH 8.0, 300mM NaCl, 10mM imidazol    e

No 4:   Eluent B
        A + 250mM imidazole

No 5:   Remarks
        2L BL21* in LB o/n @ 25C w/ 0.4mM IPTG
```




**TEV Protease**

**HRV3C Protease (aka Pre-Scission Protease)**

The simplest, first-attempt approach to prot

### Expression and Purification of Recombinant $His_6$-TEV Protease

The plasmid, pTH24-TEV was a gift of the Allen lab at BU and, I think, ultimately derives from the Waugh lab. The two critical parts of this protocol are (1) using the codon-plus cells (remember to use chloramphenicol to maintain the RILP plasmid) and (2) expressing at 16°C rather than 37° or 25°C. Other than these two considerations, this is a standard nickel affinity purification.

1. Inoculate 2 x 20ml $LB_{amp}$ starter cultures with BL21 CodonPlus cells (Stratagene) carrying the pTH24-TEV plasmid from either a fresh transformation or a glycerol stock and grow overnight at 37°C.
2. The following day, use the starter cultures to inoculate 2 x 1L $LB_{amp}$ expression cultures and grow at 37°C, 250 rpm to an $OD_{600}$ of 0.6-0.8.
3. Induce expression with 0.4ml of 1M IPTG (0.4mM final [IPTG]).
4. Reduce temperature to 16°C and grow cultures overnight.
5. Harvest cells by centrifugation at 6000 rpm for 20 min. Resuspend cell pellets in lysis buffer (25mM TRIS pH 8.0, 300mM NaCl, 10mM imidazole) at a rate of 5 ml of buffer per g of cells.
6. Lyse cells by sonication or freeze/thaw, treat with DNAse, and clarify the lysate by centrifugation at 18,000 rpm for 1h. This and all subsequent steps should be performed at 4°C (our FPLC is in a chromatography fridge).

7. Load the supernatant onto a 5ml HisTrap column and purify using a standard step-gradient program.
8. Check purity by SDS-PAGE and, if satisfactory, dialyze the protease into storage buffer (50mM TRIS pH 8.0, 1mM EDTA, 5mM DTT, 20% glycerol). Divide the dialyzed protein into 500 µl aliquots and snap-freeze in $lN_2$. Store at -80°C. The protease should remain active for at least a year (probably much longer).

**Denaturing Gel Electrophoresis (SDS-PAGE)**

Denaturing gel electrophoresis (aka SDS-PAGE) is a relatively quick and convenient method for monitoring protein over-expression and purification processes. Proteins in the samples are denatured and forced through the cross-linked polymer gel (acrylamide) by an electric field. Individual proteins are separated based on size alone. The presence of one particular protein is confirmed by (1) comparison with molecular weight standards and (2) protein concentration. For example, over-expressed proteins should be present in large quantities. Care should be taken when using these criteria, since both can be "fooled." For example, there have been cases where a researcher over-expressing a ~27kDa protein has purified, crystallized, and solved the structure of chloramphenicol acetyltransferase – the protein responsible for chloramphenicol resistance. This was **not** the protein the researcher was after, but an element from the expression plasmid! Unequivocal identification of a particular protein requires a western blot using a specific antibody. Almost as good and much cheaper/more convenient are the poly-histidine specific fluorescent dyes sold by Invitrogen (InVision) and Thermo (GelCode).

Samples are prepared for SDS-PAGE by mixing equal volumes of protein solution (e.g. fractions from purification, aliquots of bacterial culture taken at various time points, etc) and loading dye (recipe below) and then boiling for 3-5 min. The gels are fitted into the electrophoresis apparatus (follow instructions in the white binder) and the standard(s) and samples loaded into the "wells" at the top. The amount of sample to load depends on the sample type. Whole-cell extracts (raw lysate) are very rich in protein, so only 2-4µl will be sufficient. Fractions from a purification run might be more dilute and would require more. The width of the well (10-well vs. 15-well combs) make a difference too. It might be necessary to re-run a gel if the original was badly over- or under-loaded. For best results, run gels at 100V for 10 minutes to line proteins up in the stacking gel, then resolve at 200V for 50 minutes. The first step can be omitted, though, so that the gel is run at 200V for 60 min.

## Native Gel Electrophoresis (non-denaturing)

Native (non-denaturing) gel electrophoresis is done in *exactly* the same way as SDS-PAGE, **except** that all of the denaturing agents (SDS, βME) are left out of the gels, loading dye, and running buffer **and** the sample are **not** boiled before loading on the gel. Native gels are used to check for sample homogeneity before crystallization. Since native gels separate based on size, surface properties, and shape, they can be used to identify improperly folded protein that would not show up on a denaturing gel.

## Differential Scanning Fluorimetry (DSF)—aka "Thermofluor Assay"

The Thermofluor assay can be used as a quick and easy method to (1) identify buffer conditions where a protein is most stable or (2) identify ligands that bind to the protein. The assay is predicated on the fact that a more stable protein will melt at a higher temperature. The assay is done in the presence of a fluorescent dye that interacts with the protein in such a way that there is a change in fluorescence when the protein unfolds. Thermofluor assays are simple to set up and do not require a huge amount of protein.

**Materials:**

1.1 ml **20-40µM protein solution** in dilute buffer (water if possible). The concentration is something that might need to be optimized, but to save sample we start at the low end (20µM).

| | |
|---|---|
| 2.2 µl | Sypro Orange Dye, 5000X concentrate in DMSO. |
| NA | A solution screen formulated to test pH or concentrations of various salts, detergents, etc. We have the "Slice pH" kit from Hampton research, but more in-depth screens must be made in-house. We have 96 x 1ml blocks for this purpose. |

**Procedure:**

Prepare the dilute protein solution and add the 2.2 µl of dye. Mix gently but thoroughly. Portion the protein—dye mixture into 8 PCR tubes (130 µl each) and use an 8-channel pipet to place 10 µl of protein—dye solution into each well of a 96-well, white PCR plate (e.g. Eppendorf TwinTec plates). Next, use the 8-channel pipet to add 10 µl of screening solution to the protein solution in the PCR plate. Mix by pipetting up and down once or twice. Cover the plate with an optically clear sealing film and centrifuge it (you will need empty 96-well blocks to serve as adapters for the plate rotor) for 30-60s at 1000 rpm just to get the solution to the bottoms of the wells. The real-time machine is located in the Frick lab on the 4th floor of the Chemistry Building.

## Recipes

**10X TBE for agarose gel electrophoresis (1L)**

108g Tris base
55g Boric acid
40mls 0.5M EDTA (pH 8.0)

Dissolve all three chemicals in 800ml of di$H_2O$. Once dissolved, adjust volume to 1L and filter sterilize (vacuum filter through a 0.22μm membrane) and store at 25°C.

**6X Sample loading dye (10ml) for DNA gels**

25 mg bromophenol blue
25 mg xylene cyanol FF
3 ml glycerol

Dissolve the dyes (xylene cyanol can be omitted) in a few ml of water. Add glycerol and add enough additional water to make 10ml. Mix vigorously and store at 4°C.

** This dye can also be made with only the xylene cyanol, since bromophenol blue can sometimes run at the same rate as interesting DNA fragments, making them difficult to see clearly.*

### Preparing SDS-PAGE (*DENATURING)* gels

***For 2 gels (approx)* – Only use this recipe if you need an "odd" acrylamide concentration (e.g. 15%). If you are using 10 or 12% gels, then make enough for everyone to use!**

| | 8% | 10% | 12% | 15% | Stacking (5%) | |
|---|---|---|---|---|---|---|
| Water | 4.6 | 4.0 | 3.3 | 2.3 | 3.4 | |
| 1.5M Tris pH 8.8 | 2.5 | 2.5 | 2.5 | 2.5 | 0.63 | 1.0M Tris pH 6.8 |
| Acrylamide | 2.7 | 3.3 | 4.0 | 5.0 | 0.83 | |
| 10% SDS | 0.10 | 0.10 | 0.10 | 0.10 | 0.05 | |
| 10% APS | 0.10 | 0.10 | 0.10 | 0.10 | 0.05 | |
| TEMED | 0.006 | 0.004 | 0.004 | 0.004 | 0.004 | |

***For 12 gels (approx)* – We keep a supply of 10% gels for everyone's use. Make a new batch of 12 if you use the $2^{nd}$-to-last gel!**

| | 8% | 10% | 12% | 15% | Stacking (5%) | |
|---|---|---|---|---|---|---|
| Water | 27.6 | 24.0 | 19.8 | 13.8 | 20.4 | |
| 1.5M Tris pH 8.8 | 15.0 | 15.0 | 15.0 | 15.0 | 3.8 | 1.0M Tris pH 6.8 |
| Acrylamide | 16.2 | 19.8 | 24.0 | 30.0 | 5.0 | |
| 10% SDS | 0.6 | 0.6 | 0.6 | 0.6 | 0.3 | |
| 10% APS | 0.6 | 0.6 | 0.6 | 0.6 | 0.3 | |
| TEMED | 0.036 | 0.024 | 0.024 | 0.024 | 0.024 | |

**10X SDS-PAGE running buffer**

30.3g Tris base
144.0g Glycine
10g SDS (sodium dodecyl sulfate) – avoid breathing this dust! Use a mask.
Add water to make the volume ~900ml, add a stir-bar, and stir the solution aggressively. It will take some time for all of the SDS to go into solution. When it does, adjust the volume to 1L with $dH_2O$.

**2X SDS-PAGE sample loading dye**

2.4ml of 1M Tris-Cl pH 6.8
3ml of 20% SDS
3ml of Glycerol (100%)
1.6ml β-mercaptoethanol (100%)
6mg Bromophenol blue – Be careful with this dye! It gets everywhere and a single fleck goes a **long** way!
Add $dH_2O$ to 10 ml and store at 4°C

## Preparing native (NON-DENATURING) gels

***For 2 gels (approx)* – Since native gels are not used very often, do not make 12 at a time.**

| | 7% | 9% | 11% | Stacking (5%) | |
|---|---|---|---|---|---|
| **Water** | 5.1 | 4.4 | 3.7 | 3.40 | |
| **1.5M Tris pH 8.8** | 2.5 | 2.5 | 2.5 | 0.75 | 1.0M Tris pH 6.8 |
| **Acrylamide** | 2.3 | 3.0 | 3.7 | 1.00 | |
| **10% APS** | 0.10 | 0.10 | 0.10 | 0.06 | |
| **TEMED** | 0.008 | 0.006 | 0.004 | 0.005 | |

**50X Native gel running buffer**

15.0g Tris base
72.0g Glycine
Adjust the volume to 500ml with $dH_2O$.

**3X Native gel sample loading dye**

0.6ml 50X running buffer
3ml of Glycerol (100%)
2.5mg Bromophenol blue – Be careful with this dye! It gets everywhere and a single fleck goes a **long** way!
6.4ml $dH_2O$
Store at 4°C

**10% Ammonium persulfate (APS, 10ml)**

1g APS
Fill to 10ml with $mqH_2O$
Store at 4°C; good for up to **2 weeks**

**EDTA 0.5M Stock (0.5L)**

73.1g EDTA (FW=292.25g/mol)
Add 300ml of mg$H_2O$
Adjust pH to 8.0 with 10N NaOH*
When all of the EDTA has dissolved, adjust volume to 500ml with mq$H_2O$

* The solution must be somewhat basic to get the EDTA to be soluble. Adjust the pH while stirring the solution. Add enough base to make the pH 8.0+ and, as more EDTA dissolves, the pH will drift down. Once most of the powder has dissolved, be very careful with the base.

**Isopropyl β-D-1-thiogalactopyranoside (IPTG) 1M Stock (10ml)**

2.38g IPTG
Add mq$H_2O$ to 10ml in a 15ml tube
Vortex or shake vigorously to dissolve IPTG, then filter with a sterile syringe filter into a new, sterile 15ml conical tube.

**Ampicillin 100mg/ml Stock (50ml)**

5.0g Ampicillin sodium salt (NOT the monohydrate!)
Add mq$H_2O$ to 50ml in a 50ml tube
Vortex or shake vigorously to dissolve ampicillin, then filter with a sterile syringe filter into a new, sterile 50ml conical tube.

**Kanamycin 50mg/ml Stock (50ml)**

2.5g Kanamycin sulfate monohydrate
Add mq$H_2O$ to 50ml in a 50ml tube
Vortex or shake vigorously to dissolve kanamycin, then filter with a sterile syringe filter into a new, sterile 50ml conical tube.

**Terrific Broth (TB)**

16g N-Z Amine (or Tryptone)
24g Yeast extract
4ml Glycerol
920ml $H_2O$

Autoclave at 121°C, then cool to room temperature and add 100ml of autoclaved phosphate buffer.
* *If the phosphate is added to the broth before it is autoclaved, then it precipitates as it cools.*

***For 1L of phosphate buffer***
XXXg $KH_2PO_4$ (0.17M)
XXXg $K_2HPO_4$ (0.72M)
1L mq$H_2O$

**2xYT Medium**

16g N-Z Amine (or Tryptone)
12g Yeast extract
5g NaCl
1L $H_2O$

**<u>HisTrap Buffer A (2L):</u> 25mM TRIS pH 8.0, 300mM NaCl, 10mM imidazole**

50ml of 1M TRIS, pH 8.0 (or pH 8.5; can also use HEPES, pH 7.5 if it improves the purification)
35.0g NaCl
1.36g imidazole
Add mq$H_2O$ to 2L, 0.2μM filter, and degas by bubbling $N_2$ gas through the solution for 10-15min.

**<u>HisTrap Buffer B (1L):</u> 25mM TRIS pH 8.0, 300mM NaCl, 250mM imidazole**

25ml of 1M TRIS, pH 8.0 (or pH 8.5; can also use HEPES, pH 7.5 if it improves the purification)
17.5g NaCl
17.0g imidazole
Add mq$H_2O$ to 1L, 0.2μM filter, and degas by bubbling $N_2$ gas through the solution for 10-15min.

**<u>PreScission Protease Buffer (1L):</u> 50mM TRIS pH 8.0, 100mM NaCl, 1mM EDTA, 1mM DTT**

50ml of 1M TRIS, pH 8.0
5.8g NaCl
0.154g DTT
2ml of 0.5M EDTA stock
Add mq$H_2O$ to 1L, 0.2μM filter, and degas by bubbling $N_2$ gas through the solution for 10-15min.