Chapter 10

Crystallization of PTP Domains

Colin Levy, James Adams, and Lydia Tabernero

Abstract

Protein crystallography is the most powerful method to obtain atomic resolution information on the three-dimensional structure of proteins. An essential step towards determining the crystallographic structure of a protein is to produce good quality crystals from a concentrated sample of purified protein. These crystals are then used to obtain X-ray diffraction data necessary to determine the 3D structure by direct phasing or molecular replacement if the model of a homologous protein is available. Here, we describe the main approaches and techniques to obtain suitable crystals for X-ray diffraction. We include tools and guidance on how to evaluate and design the protein construct, how to prepare Se-methionine derivatized protein, how to assess the stability and quality of the sample, and how to crystallize and prepare crystals for diffraction experiments. While general strategies for protein crystallization are summarized, specific examples of the application of these strategies to the crystallization of PTP domains are discussed.

Key words Crystallogenesis, Se-methionine protein, Crystal seeding, Crystallization trials, Cryoprotectant

1 Introduction

Protein tyrosine phosphatases are important pharmacological targets [1–5] and therefore they have been extensively studied at the structural level [6–8]. A comprehensive list of available PTP structures, both of PTP domains and PTP extracellular regions, has been extracted from the Protein Data Bank (PDB) and can be found in the Appendix.

Protein crystallization has experienced a major evolution during the past 30 years with the introduction of faster cloning approaches, more sophisticated recombinant expression systems and the use of robotics to assist in setting up trials. Advances on all these fronts have transformed crystallogenesis from a "magic art"

The online version of this chapter (doi:10.1007/978-1-4939-3746-2_10) contains supplementary material, which is available to authorized users.

Rafael Pulido (ed.), *Protein Tyrosine Phosphatases: Methods and Protocols*, Methods in Molecular Biology, vol. 1447,
DOI 10.1007/978-1-4939-3746-2_10, © Springer Science+Business Media New York 2016

into a routine approach in structural biology laboratories around the world. These advances have evolved in parallel with major technical and engineering developments of X-ray generators, detectors and high-brilliance synchrotron beam lines, an increase in computing speed and the dissemination of new software packages that streamline data processing, structure determination and refinement. During this period, the number of protein structures deposited in the PDB [9] as grown from 186 in 1985 to more than 110,000 in 2015, 372 of which are PTPs.

However, the main challenge remains in the production of high quality crystals to yield good quality data from the X-ray diffraction experiments. A number of factors affect data quality, from the choice of the construct boundaries, to the freezing of the crystals in preparation for diffraction. A careful inspection and analysis of the protein sequence prior to cloning is essential to establish the correct sequence boundaries or to design mutations that may improve stability of the recombinant protein. The decisions taken at this stage will have a large impact on protein solubility, the purification process and its chances to crystallize. Even if crystals are readily obtained, the internal packing and therefore the diffraction quality and maximum resolution that can be achieved in determining the structure will depend in part, on the quality of the protein and its behavior in solution. If no crystals are obtained with the initial construct, we recommend a reinspection of the design to generate alternative boundaries, eliminate potentially disordered regions or to minimize surface entropy. PTP domains contain a compact and a well-conserved fold with an antiparallel β-sheet (five to eight strands) and surrounding helices (five to eight). Most variations in the constructs will affect the N-terminal region where more variability exists between PTPs. A sequence and structural alignment of several PTP structures related to your target PTP might help to define similar boundaries for your initial construct.

A main consideration for successful crystallization should be the solubility of the recombinant protein. Very soluble proteins can be concentrated to high levels (>10 mg/mL), which speeds up the initial nucleation process and facilitates crystallogenesis. Quality control of the protein sample is essential to ensure the purity (>95% pure) and homogeneity of the sample. For this, we recommend to characterize the behavior of your purified protein in solution by testing enzyme activity, if the wild-type PTP protein is expressed, as well as monodispersity of the sample to ensure that it contains only *one type* of particle, either monomer or dimer, or other oligomers. Polydispersity (i.e., a mix of monomer/oligomers) of the sample results inevitably in a lack of crystallogenesis or poorly diffracting crystals. Achieving monodispersity requires rigorous purification of the protein (*quality over quantity always!*) involving affinity purification, ion exchange and size exclusion chromatography. Often, an additional screening of different pH and buffer

conditions is necessary to prevent unspecific aggregation and precipitation of the protein. The use of additives such as EDTA/EGTA, reducing agents (DTT, TCEP, β-mercaptoethanol), metal ions and other ligands is sometimes needed to stabilize the protein or its oligomeric state, thus ensuring homogeneity.

PTPs are more stable in buffers that contain sulfonic groups such as HEPES or MES because these bind to the active site and maintain the P-loop in its extended open conformation. Often the active site in PTP crystals is occupied by either phosphate or sulfate that are captured from the culture media, contamination in the buffers (from glassware washing detergents), or the crystallization conditions (specially ammonium sulfate precipitant). These ligands will help crystallogenesis of the apo-enzyme, but may need to be removed if you are planning to do co-crystallization or soaking with other ligands (inhibitors, substrates or peptides).

Protein crystallogenesis consists of preliminary screens, using automation and robotics systems if available, testing a variety of preset precipitant conditions (mostly commercially available screens), followed by systematic optimization of the initial hits to produce larger or better quality crystals. Optimization approaches involve seeding with the initial crystals to grow larger crystals in larger drops and/or refining the crystallization conditions (precipitant concentration, pH, type of buffer, temperature, additives, or ligands). Once crystals are obtained, the next step is to identify a suitable cryoprotectant to allow efficient flash freezing of the crystals without compromising their integrity of the quality of X-ray diffraction.

The whole process, from protein purification, crystallogenesis to preparing crystals for X-ray diffraction, can take as little as a few days for high quality stable and well-behaved proteins, to several weeks if optimization is required. Most human PTPs will crystallize readily under standard and published conditions.

If no results are obtained, do not despair! There is always an alternative way for reluctant proteins. For example, co-crystallization with ligands or substrates, co-expression with chaperones or co-expression/purification with more biological protein partners to form stable complexes. Failing that, choosing an ortholog from another species may prove helpful, particularly if present in a thermophile.

2 Materials

All solutions should be prepared with ultra pure deionized water, filtered using a 0.22 μm filter and kept in glass containers. Chemical reagents should be of analytical grade with >98 % purity. We recommend storing all solutions at 4 °C except for high concentration salt buffers (>1 M), which should be stored at room temperature.

Polyethylene glycol solutions are susceptible to microbial contamination that can be prevented by adding 0.02% of sodium azide. Commercial solutions do not require further additives as they may already contain azide or other antimicrobials. Precautions should be taken when manipulating sodium azide or other toxic and hazardous chemicals; always follow manufacturer's instructions. Make sure that you adhere to good laboratory practice regulations and that you comply with health and safety guidelines for the disposal of chemicals and waste materials.

2.1 Se-Methionine Protein Production

1. LB media: Dissolve 10 g of tryptone, 5 g of yeast extract, and 10 g of NaCl in water and make up volume to 1 L. We recommend adding glucose up to 1% to stop any leaky expression. Sterilize by autoclaving and store at room temperature.
2. Minimal media: Prepare the M9 salts by dissolving 64.0 g $Na_2HPO_4{\cdot}7H_2O$, 15.0 g KH_2PO_4, 2.5 g NaCl, 5.0 g NH_4Cl in distilled deionized water and adjust to a final volume of 1 L. Sterilize by autoclaving. This is enough for 5 L of culture medium and can be split into aliquots and stored at 4 °C for future use. To prepare 1 L of minimal media, use 500 mL of sterile distilled deionized water, add 200 mL of the M9 salts stock (see above), 2 mL of 1 M $MgSO_4$ (sterile), 20 mL of 20% glucose or glycerol, 100 μL of 1 M $CaCl_2$ (sterile), 1 mL of the trace element mix and adjust pH to 7.4 with NaOH as needed. Adjust to a final volume of 1 L with sterile distilled deionized water (adapted from Sambrook [10]).
3. Trace element mix (1000×): use 800 mL of sterile distilled deionized water to dissolve 4.4 g of $ZnSO_4{\cdot}7H_2O$, 180 mg of $MnCl_2{\cdot}4H_2O$, 20 mg of $(NH_4)_6Mo_7O_{24}{\cdot}4H_2O$, 80 mg of $CuSO_4{\cdot}5H_2O$, 5.0 g of $FeSO_4{\cdot}7H_2O$, 250 mg of $CoCl_2$, and 1.0 g of H_3BO_3. Add 1 mL of concentrated H_2SO_4, adjust to a final volume of 1 L with distilled deionized water. Filter to sterilize (do not autoclave!).
4. Amino acid mix (powder): Phenylalanine (100 mg/L), threonine (100 mg/L), lysine (100 mg/L), isoleucine (50 mg/L), leucine (50 mg/L), valine (50 mg/L), and l-selenomethionine (80–120 mg/L).
5. Phosphate buffered saline (PBS): Dissolve 8.0 g of NaCl, 200 mg of KCl, 1.44 g of Na_2HPO_4, 240 mg of KH_2PO_4 in 800 mL distilled deionized water. Adjust pH to 7.4 with HCl and adjust to a final volume of 1 L. Sterilize by autoclaving.
6. Lysis buffer: use preferred lysis buffer. We recommend avoiding PBS if you plan to do co-crystallization with ligands. Phosphate binds to the active site and it may be difficult to displace by the ligand if used at low concentrations. Alternatively use a buffer containing 20 mM HEPES, 500 mM NaCl, 1 mM

TCEP, 0.1% Triton X-100, 10 mM imidazole, pH 7.4 for His-tagged proteins. For GST-tagged proteins use 20 mM HEPES, 150 mM NaCl, 3 mM TCEP, 0.1% Triton X-100, 2 mM EDTA, pH 7.4. Protease inhibitors can be also added if necessary (commercial cocktails are available). Follow with your established purification procedures.

2.2 Protein Crystallogenesis

1. Commercial screens for crystallogenesis. Several companies sell prefilled 96 deep well blocks screens ready for use with automated liquid-handling robotics systems or individual kits containing 24–96 reagents, each in 10 mL bulk that can be used to setup manual screens. A good starting screen is the JBScreen Phosphatase from MiTeGen, derived from data mining the PDB for successfully crystallized phosphatases [11]. Other popular general screens are the Morpheus (Molecular Dimensions [12]), Crystals Screen I and II (Hampton Research [13]), PEG/Ion [12, 13], PACT premier, JCSG$^+$, JCSG Core Suite (QIAGEN [14]).
2. Crystallization plates: MRC crystallization plates [12], 24-well "Linbro" plate [12], XTalQuest 24-well plates for sitting drops [11], Crystal Clear Sealing tape, siliconized glass cover slides for hanging drop crystallogenesis in "Linbro" style plates and Dow Corning high vacuum grease [12–14].
3. Seeding tools: Streak seeding tool [14], MicroSeed Beads for the preparation of seed stocks for use in microseeding and Matrix seeding protocols [12, 14].

2.3 Crystal Freezing and Mounting

1. Cryoproctectant oils: Parabar 10312 (formerly known as Paratone-N), Perfluoropolyether Cryo Oil, and Santovac Cryo Oil are all commonly used for cryoprotection of crystals [14]. Parabar 10312 is a viscous oil and it is recommended to mix with 50% mineral oil to lessen the mechanical stress on your crystals during the freezing procedure.
2. Two primary styles of cryo-loops are available, fiber-based loops [14] and Lithographic loops [11, 12]. A selection of loop sizes will be required, an ideal loop will match the crystal size, avoiding contact with the crystal whilst minimizing the surrounding liquor. Lithographic style loops are also available in a variety of shapes and meshes that can provide additional support when dealing with thin plate like crystals.
3. SPINE standard cryo-pins are compatible with the vast majority of synchrotron facilities worldwide. For a full list of compatibility *see* ref. 15.
4. Different Synchrotrons utilize different crystal positioning systems—it is important to check with your local synchrotron which system they are using as specialist equipment and tools

will be required. However basic cryo-cooling of crystal samples will require:

- Magnetic cryo-wand for spine caps [12, 13].
- Dewars, metal or foam [12, 13]. Modern foam-based Dewars are excellent for plunge freezing crystals. A shallow dewar holding no more than ~10 cm depth of liquid nitrogen is ideal for plunge freezing.
- Shipping Dewar. A popular option for shipping frozen crystals is the Taylor Wharton CX100 Dry Shipper and Shipping Case. This system can be matched with a variety of inserts to allow the safe transportation of cryogenically stored crystal samples internationally to your chosen synchrotron facility. The most appropriate dry shipper will vary depending upon the crystal positioning system utilized at your chosen synchrotron [12, 13].

3 Methods

3.1 Construct Design and Sequence Analysis

1. Use the XtalPred-RF website [16] to run a prediction on the crystallizability of your PTP construct. Paste the sequence of your protein construct in FASTA format and into the search box.
2. Tick boxes for the SERp analysis and the optional feature to find bacterial homologues.
3. Retrieve results. You will find information on the estimated isoelectric point (IP) of the protein (*see* **Note 1**), disorder prediction, crystallization classes and homologues in the PDB or other databases (Fig. 1).
4. Click on the "link to protein details", this will take you to a page containing the full analysis of your construct and its probability of crystallization in comparison to a pool of other proteins. It also contains links to homologues in either the PDB or a nonredundant database. It has a table with the protein features and predictions and a link to the construct design page (Fig. 2).
5. Click to the link "construct design" to access to the multiple sequence alignment and information on conservation, disorder, secondary structure, and others. When designing your construct it is important to consider a number of key factors and the bioinformatics analysis will help to guide you in choosing appropriate boundaries, avoiding regions of disorder whilst maintaining all highly conserved structural and catalytic elements (Fig. 2). It is wise to generate more than one construct with a series of termini varying by ~2 residues at a time, this will give you the greatest chance of hitting the ideal length for

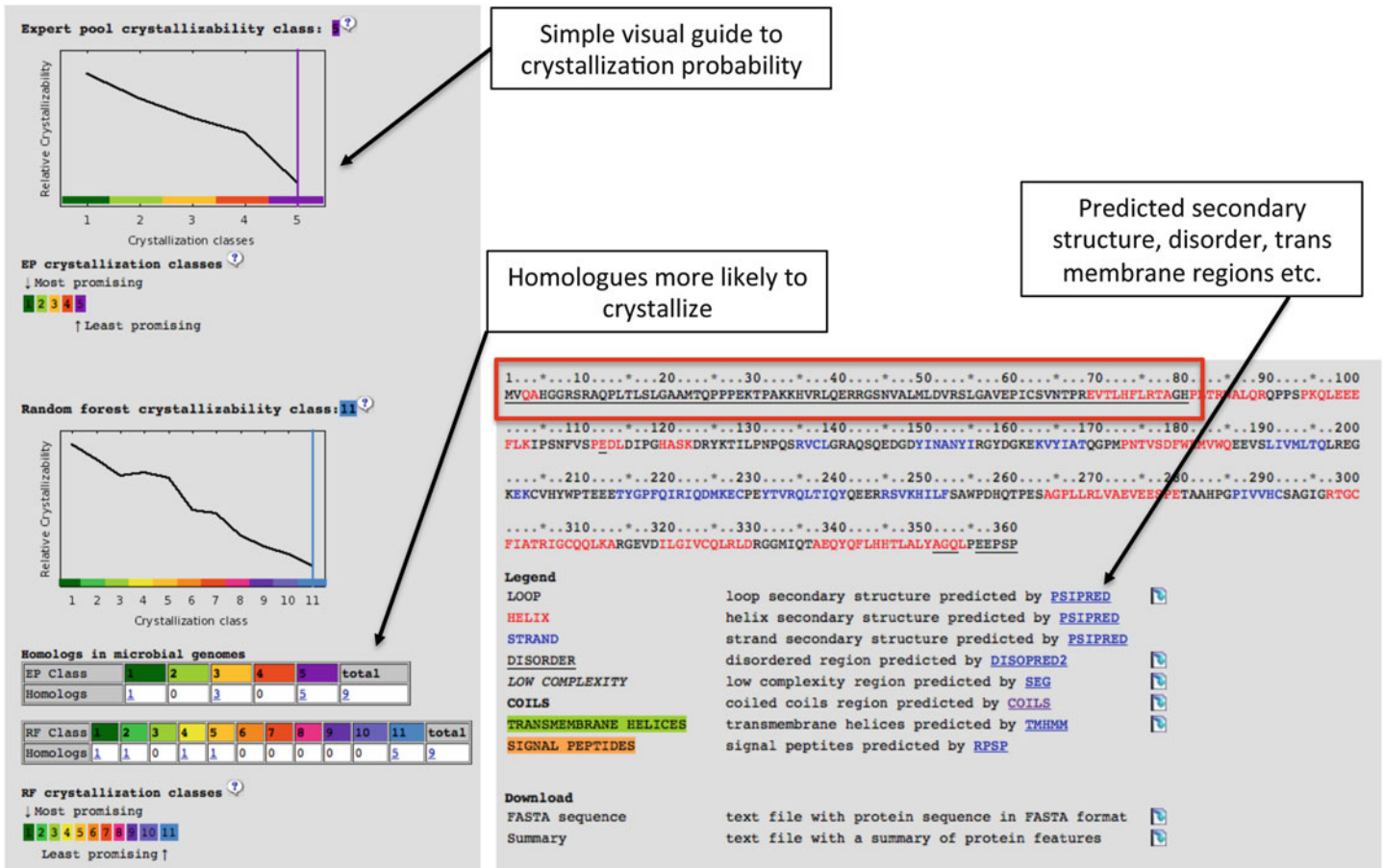

Fig. 1 Results obtained in a *XtalPred-SerP* query for PTPN7. Different parameters are provided: an estimated evaluation of the probability of successful crystallization of the construct, information of secondary structure predictions and homologues in microbial species that could be used as alternatives for crystallization. In this example, the full sequence has a very poor crystallizability score and the first 80 residues are predicted to be disordered, suggesting that a shorter version should be considered when designing the construct for protein expression

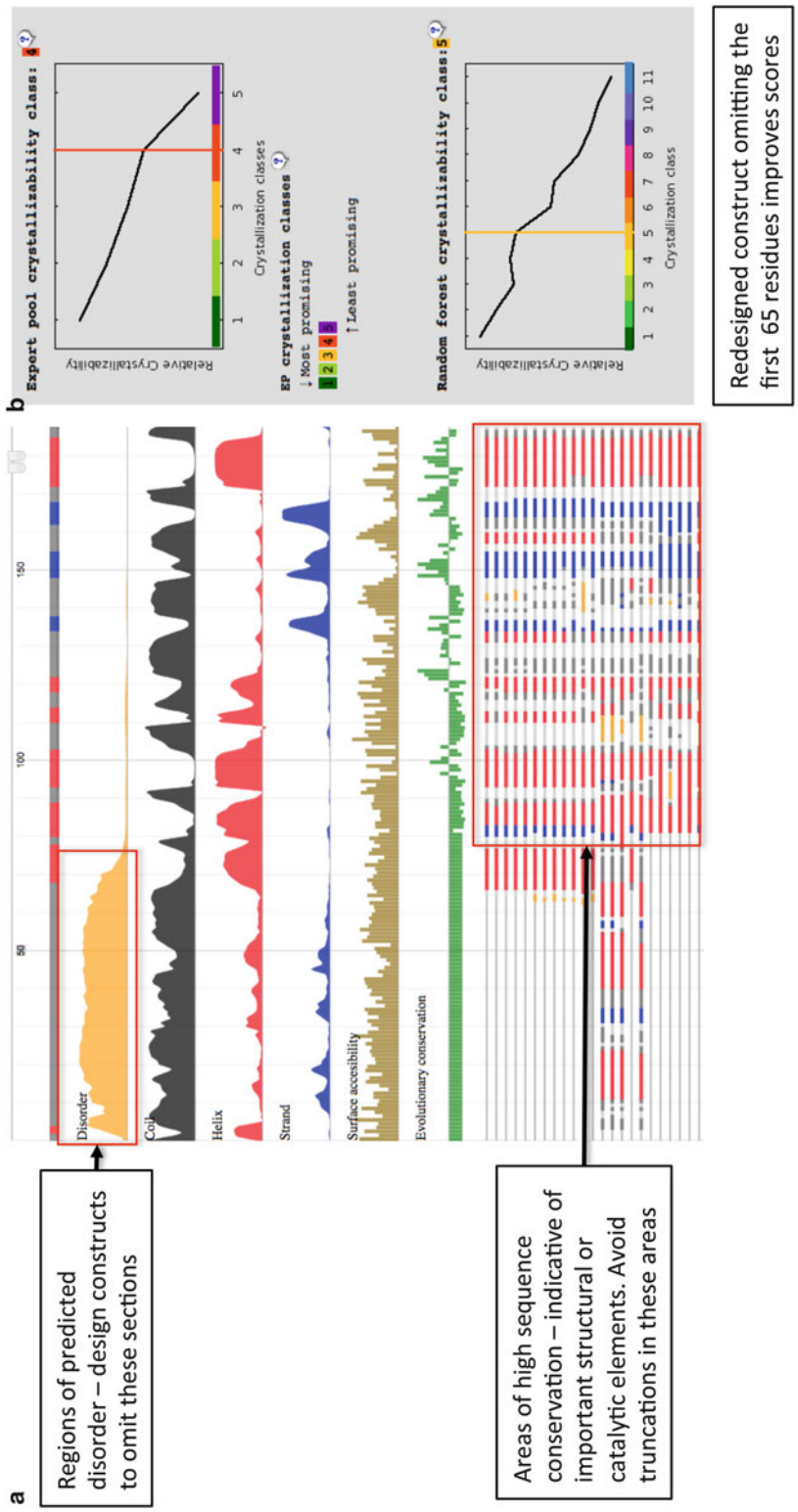

Fig. 2 (**a**) Example of results obtained for PTPN7 from the *XtalPred-SerP* "construct design" section. The prediction results highlight a disordered region at the N-terminus of the input construct sequence. Evolutionary conservation highlights areas to maintain in the construct, while others not conserved can be removed. A suggested strategy is to generate a series of constructs truncated in a stepwise manner, perhaps two residues at a time, to enhance the chances of obtaining a stable protein [17]. (**b**) Omision of the first 65 residues substantially improves the chances of crystallization as seen in the better scores. The prediction is confirmed in this case as the N-terminal truncation construct has yielded 15 structures (wt and mutants) deposited in the PDB

crystallogenesis [17] In addition it is useful when affinity tagging your protein for purification to design constructs with both N and C terminal cleavable tags. Many PTPs have been crystallized with their affinity His-tag in place, however as the tag can impede crystallogenesis or crystal packing in some instances (particularly with larger tags such as GST or MBP), designing the cleavage into your construct from the outset can save a lot of time later in the process.

3.2 Expression Systems Used to Produce PTP Proteins for Crystallographic Analyses

1. The preferred vectors for recombinant expression systems used for PTPs are those that provide His-tag fusions with the target PTP gene (pET, pNIC and their variations, *see* Appendix). Less common now is the use of GST-tag fusions using pGEX vectors, as these require protease cleavage of the GST tag prior to crystallization. GST-tags may also interfere with phosphatase activity and induce dimerization if left uncleaved. The use of proteases such as thrombin has additional liabilities because they may also cleave the target protein, therefore these should be avoided when possible. We recommend choosing a vector that contains a His-tag, a small linker and a very specific protease cleavage site (i.e., TEV protease) to minimize chances of proteolytic cleavage during preparation of the sample and generating undesirable heterogeneity.
2. The preferred host for expression of PTP catalytic domains is *E. coli* (Appendix). Bacterial hosts with T7 RNA polymerase systems are the most commonly used for large-scale expression because they are easy and relatively inexpensive to culture and produce high yield of recombinant protein. Common strains used are BL21 and all its derivatives (BL21 (DE3), BL21 pLysS, B834, C41 (DE3), CodonPlus, Rosetta, Shuffle, or Tuner). If you have difficulties optimizing expression on your chosen system, we recommend to try different hosts first (various bacterial strains) and then change your vector. Some strains such as C41, Rosetta, or Shuffle give better results with proteins that are toxic to bacterial growth.
3. Extracellular domains of PTPs are normally expressed in mammalian cells lines (CHO, HEK-293) (Appendix) or using cell-free systems (less common as yields are low). Mammalian cell expression allows for posttranslational modifications such a glycosylation, which may be essential for native protein folding. These expression systems are more expensive because of the medium used and the protein yield is sometimes lower than bacterial expression systems.

3.3 Preparing Se-Methionine Derivatized Protein

1. Prepare minimal media the day before the experiment as per materials instructions.
2. Grow a starter culture (20–50 mL), from your glycerol stocks or from plated colonies, in LB media supplemented with the

appropriate antibiotics and glucose (1%), at 37 °C overnight with agitation.

3. Next morning, transfer some (or all) of the overnight culture into 1 L of prewarmed unlabelled LB media, supplemented with antibiotic and glucose. Grow this culture at 37 °C with good aeration to promote rapid proliferation to reach an OD_{600} of 6–10 in about 2 h.
4. Harvest the grown culture at 3500 × *g* for 30 min in a refrigerated centrifuge, discard the supernatant medium and wash the pellet with autoclaved PBS buffer three to four times to remove completely the LB medium.
5. Prepare the Se-Met medium by dissolving the amino acid mix (without methionine!!) powder into the minimal media. Resuspend the pellet in warm Se-medium and incubate the culture for 2 h at 37 °C. The mix is designed to inhibit methionine biosynthesis and to allow the incorporation of Se-Met instead (*see* **Note 2**).
6. Induce your culture with IPTG according to the preestablished conditions (temperature, and length of time) for the unlabeled protein. In the new medium it may require a bit of extra time for the bacteria to adjust to the new growth conditions.
7. Harvest the cells in the usual manner 3500 × *g*, 4 °C, 15–30 min and flash-freeze the pellets in LN_2 and store at –80 °C.
8. When ready to proceed with the purification, thaw the tubes with the pellet in a bucket with ice, resuspend the pellet in enough lysis buffer to obtain a density of approximately 0.2 g/mL (e.g., 10 g topped to 50 mL), and then transfer to a beaker. You can vortex the tubes to help homogenizing the suspension. Make sure that the suspension is smooth with no clumps of pellet left.
9. Disrupt the cells by sonication or other means (French press, homogenizer) and separate the supernatant containing your overexpressed protein by centrifugation (12,000 × *g* 30 min, repeat twice to clarify the supernatant until it looks transparent). Proceed with your purification protocol the same as used for unlabeled protein.

3.4 Assessing the Quality of Your Sample

Critical to successful crystallization is the quality of the protein sample. This should be in a native conformation, preferably active, monodisperse, homogeneous, and soluble at high concentrations if possible. A number of biophysical techniques are now routine in many laboratories and biomolecular facilities that provide a quality profile of a given protein sample. These include circular dichroism, analytical ultracentrifugation (sedimentation velocity and equilibrium sedimentation), multi-angle laser light scattering, dynamic

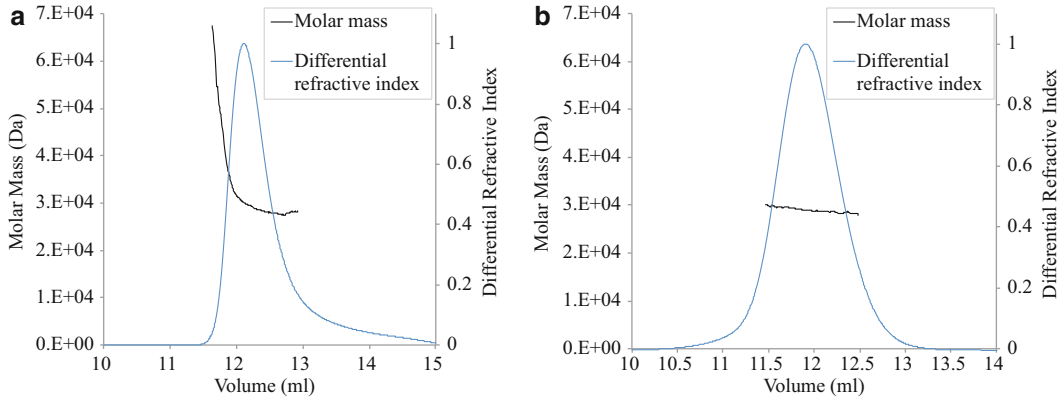

Fig. 3 Example of SEC-MALS (size-exclusion chromatography and multiangle light scattering) analysis of two different PTPs samples (**a**) protein sample, in 20 mM MES, 150 mM NaCl, 1 mM EDTA, pH 6.0, showing a polydisperse distribution of molecular mass values although it elutes in a single peak (11.5–13.3 mL). (**b**) Protein sample, in 20 mM HEPES, 150 mM NaCl, 1 mM EDTA, pH 7.4, showing a monodisperse distribution of molecular mass throughout the elution peak (11.3–12.7). Data courtesy of Dr. Andrew Currin

light scattering, 1-D NMR, and fluorescence spectroscopy (Thermofluor) [18]. These techniques allow the investigation of the size and shape of proteins particles, to calculate the hydrodynamic parameters, to estimate secondary structure content and the stability of your protein under different buffer and pH conditions [18]. This information is very valuable to design the best crystallization strategy and complements the structural analyses. Most of these techniques are conservative, do not destroy the protein and require small amounts of sample. An example of a monodisperse and polydisperse protein sample is shown in Fig. 3.

3.5 Assessing the Stability of Your Sample by Thermofluor

Thermofluor or Differential Scanning Fluorimetry (DSF) provides a fluorescence readout measurement of thermally induced melting in proteins. Knowledge about the temperature at which a given protein melts, T_m, is a good indicator of protein stability. It has been clearly shown that optimizing the stability of protein samples for crystallogenesis can aid the formation of crystals [19]. The three dimensional lattice is favored when all protein molecules present within the solution adopt a single structurally identical conformation. Thermofluor allows us to rapidly identify conditions that maximize the thermal stability of a protein and thus reduce the conformational heterogeneity of the protein dynamic states, leading to improved crystal formation [19]. Full protocols and extensive details available from refs. [19, 20].

3.6 Crystallogenesis

3.6.1 Pre-crystallization Test

1. Your sample will contain the purified protein at a concentration as high as it is reasonably achievable. We recommend that you aim to reach at least 10 mg/mL or higher. The sample buffer should contain some salt (100–150 mM minimum) to

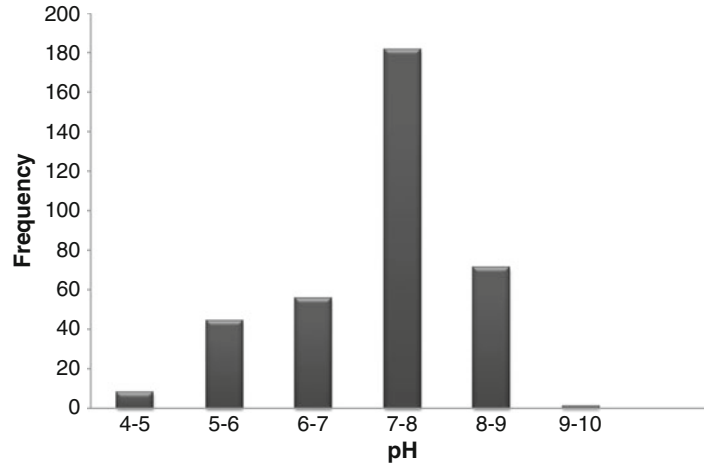

Fig. 4 Graph showing the frequency distribution of pH values in the crystallization conditions for PTPs. Most conditions are in the range of pH 7–8 although a large proportion of conditions are at pH 8–9, probably due to the increase in the p*K* of the fusion protein in the presence of the His-tag. Extracted from the PDB entries and published data as per list in Appendix

avoid premature precipitation and be at a pH that is compatible with activity and solubility of the protein (pH 6–8 is reasonable, although optimal pH for activity of most PTPs is pH 5–6). However, the pH of the crystallization conditions may be different than that of the sample buffer, as shown in a graph of the most common pH values from the crystallization conditions of PTPs (Fig. 4) extracted from the Appendix.

2. Mix 0.5 μL of sample with 30% PEG 5000 and check that the drop precipitates. If the drop remains clear then further concentrating of the sample will be required prior to crystallogenesis.

3.6.2 Crystallization Screening

Current automation favors sitting drop vapor diffusion as the means of crystallogenesis; however, alternative methods such as hanging drop, free interface diffusion and microfluidic techniques are all valid approaches and can be used according to sample availability and type of experiment. Crystals grow once the protein solution reaches super saturation. This can be achieved by either using highly concentrated protein samples or by increasing the precipitant concentration until the conditions in the drop have reached the metastable region of the saturation curve for the particular protein (Fig. 5). Factors such as buffer pH, salt concentration and temperature will all affect the saturation profile of your protein. A good screen should allow you to establish where the metastable region is for your sample and to identify conditions to drive nucleation and crystallogenesis [21, 22].

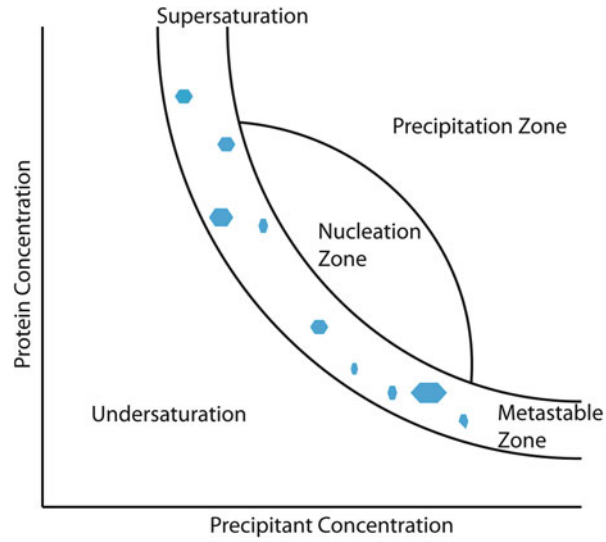

Fig. 5 Diagram showing a typical protein precipitation curve. Protein crystallization depends on reaching a metastable zone within the precipitation curve by either increasing the protein concentration or the precipitant concentration. The curve shape will vary for every protein according to its solubility and stability properties. Knowing the solubility of your protein under certain pH and salt concentration is critical to reach the metastable zone quicker, avoiding irreversible precipitation

1. Prior to setting drops it is good practice to spin down the sample (13,000 × *g* for 15 min). This will help to minimize the presence of detritus arising during the purification procedure including chromatography resin, plastic flakes, denatured and aggregated proteins (and fabric fibers if you are not wearing a lab coat!).
2. Protein samples should be stored on ice to minimize any degradation prior to screening.
3. Prepare the crystallization plates by adding the screen solutions (24-kit or 96-kit from commercial screens or homemade screens) to the reservoir wells (*see* **Note 3**). This can be done manually with a multichannel pipette or using a liquid handling automated system. The volume to add will vary depending on the plate (Appendix).
4. Protein sample and crystallization solution are mixed in a 1:1 ratio to form a drop, mixed and then sealed and incubated to allow vapor diffusion to equilibrate the precipitant concentration between the drop and the well.
5. Setting the drops manually: a total volume of 2 μL (1 μL of sample + 1 μL of reservoir solution) will be suitable for either hanging drops or sitting drops. Larger drops can be setup during optimization (up to 10 μL in hanging and 25 μL in sitting

drops) to grow bigger crystals, but during the initial screen you want to use minimal amounts of sample until you identify some hits.

6. Using a robotic platform to set the drops: use a total volume of 400 nL (200 nL of sample + 200 nL of reservoir). This is a good balance between minimizing sample consumption whilst still allowing for the growth of useable crystals.
7. Manual hanging drops: use 24-well Linbro Plates and fill reservoirs well with 0.8 mL of each screen condition. Cover the rim of the well with silicone grease (dispense from a 5 mL syringe previously filled with the grease), or vacuum oil. Take a round siliconized coverslip (22 mm), clean with wipe or spray with compressed air. Aliquot 1 μL or sample onto the coverslip. Aliquot 1 μL of reservoir solution and add to the sample drop, mixing by pipetting up and down three times. Invert the coverslip and seal over the siliconized rim of the well. Push gently to remove trapped air bubbles. Incubate (4 or 21 °C) for 24 h before checking for signs of crystal growth (*see* **Notes 4** and **5**) (Fig. 6).
8. Manual sitting drops using 24-well XTalQuest plates and fill well 0.8 mL. Aliquot 1–5 μL sample onto the central well. Aliquot 1–5 μL of reservoir solution and add to the sample drop, mixing by pipetting up and down three times. Seal the individual wells with a coverslip or with sealing tape over the entire plate. Incubate (4 or 21 °C) for 24 h before checking for

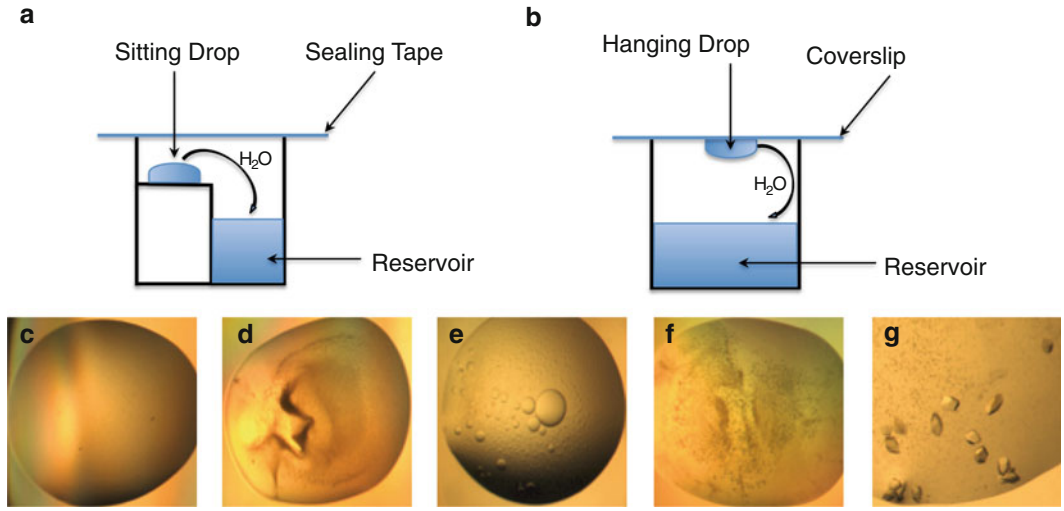

Fig. 6 Diagram showing vapor diffusion crystallization setup trials using either sitting drops (**a**) or hanging drops (**b**). Vapor diffusion allows equilibration of the concentration of precipitant between the reservoir and the drop to slowly reach the nucleation state. (**c**–**g**) Usual results from crystallization screens ranging: (**c**) clear drop, (**d**) precipitate, (**e**) phase separation, (**f**) microcrystals and (**g**) single crystals suitable for X-ray diffraction

signs of crystal growth (*see* **Notes 4** and **5**). Alternatively use MRC 48-well optimization plate [12]. This option maintains the Society for Biomolecular Sciences (SBS) standard footprint, allowing integration with high throughput screening automated systems.

9. Manual sitting drops using MRC 48-well plates (SBS footprint) or other 96-well format trays (MRC Maxi Optimization plates). Aliquot 200–400 μL of screen solution into the reservoir using a multichannel pipette. Aliquot 1–5 μL of protein into each of the drop well positions (some plates have two or three drop wells that allow testing different samples with the same precipitant condition) on the plate using a multichannel pipette or robotics systems if available. Aliquot 1–5 μL of the reservoir conditions into the respective protein drops already dispensed onto the plate. Seal the plate (sealing tape) and incubate (4 or 21 °C) for 24 h before checking for signs of crystal growth (*see* **Note 6**).
10. Requirements for drop size, well volume, protein sample amount and average crystal growth time, vary depending on the approach used (manual or automated). As guidance we have summarized the different parameters in Table 1.

3.6.3 Rapid Hit Optimization

Preliminary screening will routinely produce crystals that require some further optimization prior to data collection (Fig. 7). Grid based screening has been the method of choice to optimize these initial hits (Fig. 8). However, a powerful and rapid method of optimization is to utilize Matrix seeding as described by Darcy et al. [23]. Decoupling nucleation from crystal growth can very quickly optimize crystal growth and highlight areas of crystal space able to sustain growth but no nucleation of crystals.

1. Locate a preliminary hit containing some crystals or crystalline material.

Table 1
Typical values for crystallogenesis screens including both manual and automated systems

	Manual setup	Robotic setup
Drop volume	0.5–5 μL	100–500 nl
Reservoir volume	0.5–0.8 mL	30–50 μL
Growth time for crystals	3–5 days	24–48 h
Total protein volume required for initial screening (480 drops)	240–2400 μL	48–240 μL
Time for setup	~2–3 h	~30 min

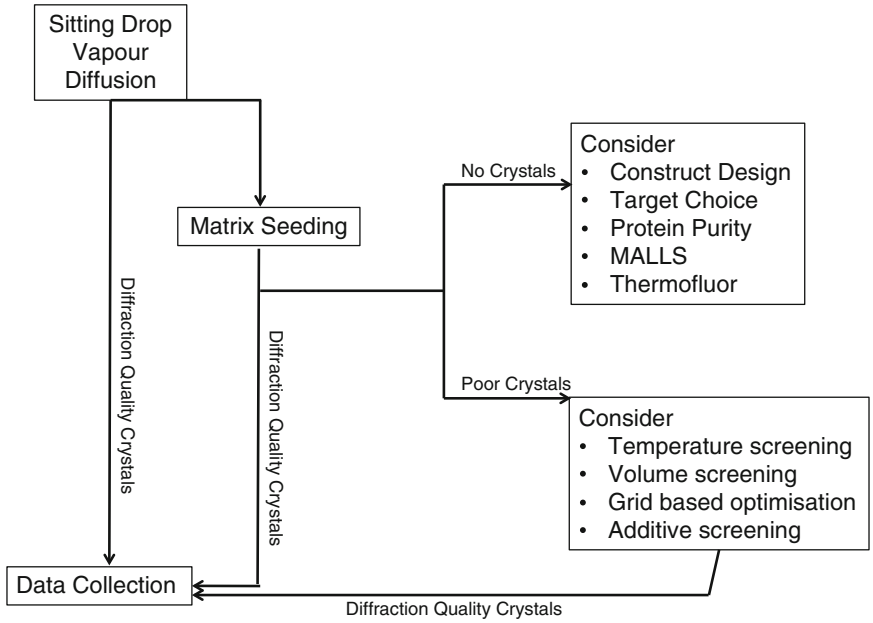

Fig. 7 Flowchart in a crystallization optimization protocol. Decision making points lead to either diffraction of suitable crystals or to reconsider new constructs or optimization of the initial crystals hit conditions to improve size/quality of the crystals. This is an iterative process that may require several runs of optimization depending on the difficulty of the project

2. Aspirate the entire drop containing the crystals along with ~50 μL of the mother liquor from the reservoir and dispense this into a Micro seed bead [12, 13].
3. Vortex this mixture for 90 s to generate the seed stock (*see* **Note 7**).
4. Transfer to ice.
5. Setup repeat screens as carried out for the primary screening but reduce the protein volume by 10% and substitute this volume with seed stock (*see* **Note 8**). Aliquot 180 nL of protein. Multi-aspirate (*see* **Note 9**) 20 nL of seed stock and 200 nL of crystallization cocktail and dispense these together into the protein drop (Fig. 9).
6. Seal and incubate as previous.
7. Alternative Optimization Strategies. Despite the power of Matrix seeding as a means of rapid hit optimization, it may still be necessary to utilize alternative approaches for some systems (Fig. 8). Optimization routes to consider include:
   - Temperature screening. If after a period of initial incubation trays do not show signs of crystal formation or precipitation it can be useful to move them to an alternative incubation temperature. Some systems will readily crystallize over

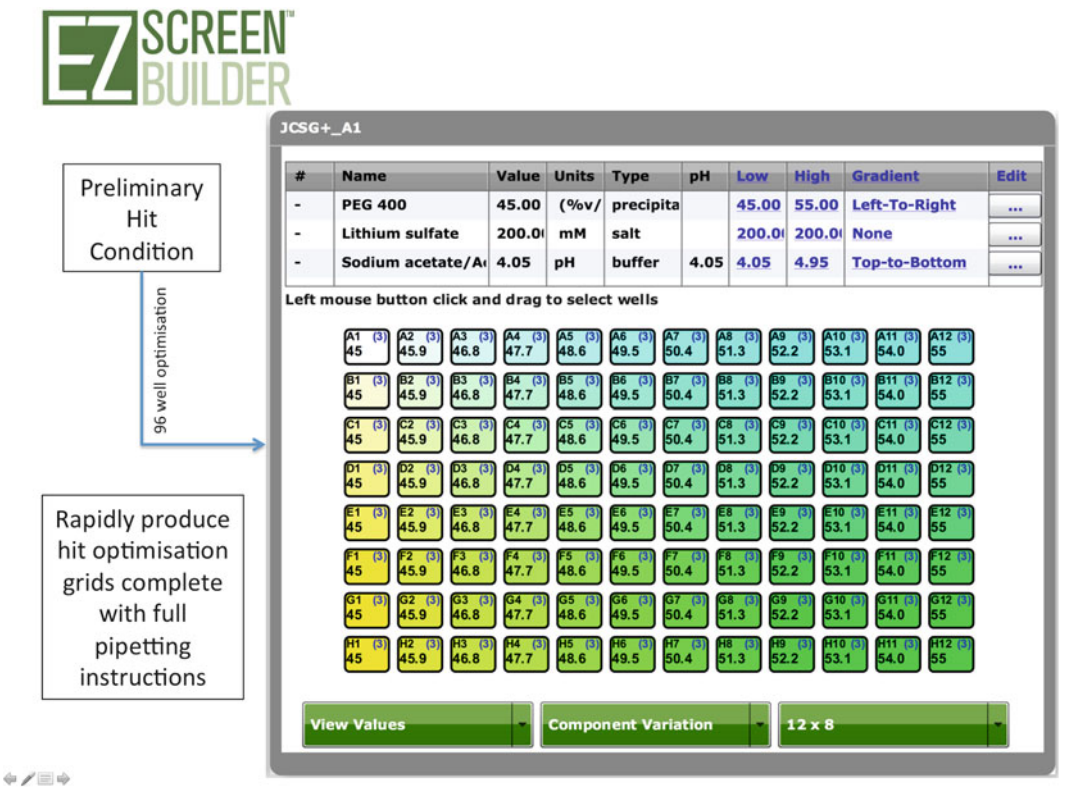

Fig. 8 EZ Screen builder allows designing a whole grid screen from the original hit conditions. It is a simple way to rapidly produce complete hit optimization grids with full pipetting instructions

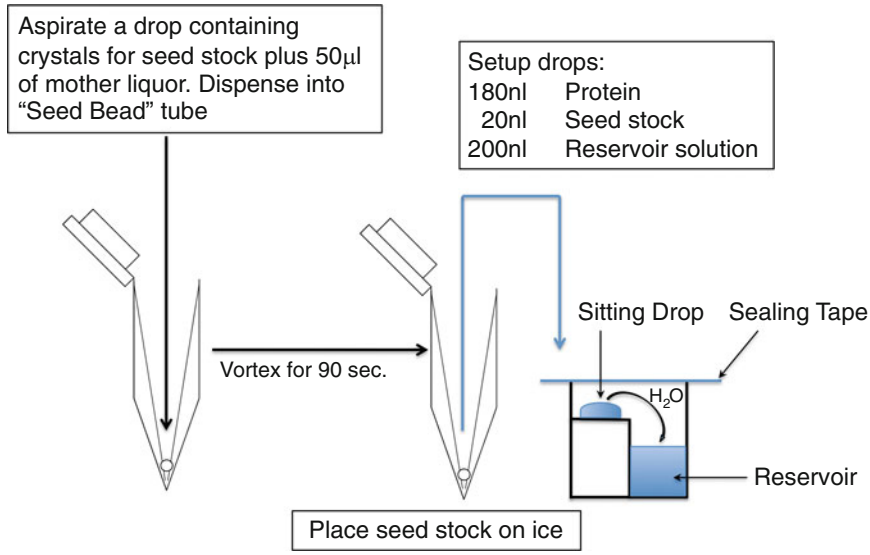

Fig. 9 Diagram showing how to prepare a seed stock from a drop containing initial crystals. Seed stocks can be used *as is* to rescreen using an optimization grid of conditions or to prepare serial dilutions to reseed into the original condition to grow larger crystals

a broad range of temperatures whilst others appear to only crystallize within a very narrow temperature window. Typical temperatures to test: 4, 15, 21, 25 °C.

- Varying drop volume. Drop volume is a much-underutilized screening variable and can have a marked impact on the equilibration rate, dynamics and on the crystal size. Larger drops will slow down the rate of equilibration whilst smaller volumes will lead to more rapid equilibration. Setting drops of total volume 400, 800 and 1200 nL on a single three well plate can be an excellent strategy to optimize initial hits obtained in 400 nL drops.
- Varying reservoir volume. This will similarly have an impact on the rate of equilibration of the drop. Reducing the reservoir volume will slow the rate of equilibration whilst increasing the reservoir volume will have the opposite effect.
- Once an initial hit condition has been located a grid screen expansion of this condition can help to optimize the crystals obtained (Fig. 8). Grid based optimization can be combined with seeding to further aid the process. An excellent free online tool for grid-based optimization is provided by Rigaku [24]. This tool will assist you in both designing and preparing your optimization grid screen. If funds are plenty you can simply have Rigaku make the optimized screen for you!

- Additive screening involves the introduction of additional components into an already located hit condition (ions, alcohols, organic solvents, etc.). Additive screens are available from a number of commercial suppliers [12, 13]. If your protein requires any ions or cofactors for activity these should be present in the sample buffer or added to the crystallization drop.
- Limited proteolysis screen. Some PTP constructs may have flexible or disordered regions that prevent crystallogenesis, as discussed in Subheading 3.1. One way to overcome this is to redesign your construct to eliminate those regions. Alternatively, you can use limited proteolysis during the optimization process by adding proteases into your crystallization drop. Protease kits are available commercially for example (Proti-Ace) [13], follow manufacturers recommendations to decide on the amount to add depending on your drop volume.

The ultimate goal is to obtain high quality diffracting crystals with size being dependent on the X-ray source to use (*see* **Note 10**).

3.7 Ligand Inhibitor Complexation

Preparation of PTP complexes with ligands (substrates, inhibitors, peptides) can be achieved in different ways: by incubating purified PTP with the ligand previous to crystallization, by co-crystallization mixing protein and ligand in the crystallization drop or by soaking apo-protein crystals in a solution containing the ligand. In all instances, ensuring that the active site is available for ligand binding is essential, as it will affect the efficiency of the binding as well as the final occupancy of the ligand in the crystal structure. We recommend minimizing the exposure of the PTP sample to phosphate and sulfate ions during purification by using Tris-based buffers. Another precaution is to rinse thoroughly your glassware with distilled water to wash off any phosphate traces in the detergents. In addition, you can include a desalting step at the end of your purification to exchange the purification buffer into a phosphate-free buffer. A simple and quick way to check the presence of phosphate in your buffer is to test it with malachite green reagent.

Small-molecule ligands may have very limited solubility in aqueous buffers and thus cosolvents (ethanol, DMSO, other alcohols or organic reagents) may be used to enhance their solubility. Unfortunately these solvents can have a negative impact upon crystallogenesis by preventing nucleation. One alternative method of overcoming this issue is to mix the ligand and protein at low concentration by directly dissolving the powder into the protein sample. Solvents can thus be minimized or avoided altogether, instead allowing the protein and ligand to complex over a protracted period.

1. Dissolve ligand compound (powder or stock solution) in the protein sample solution (before protein concentration). Gently mix on a roller shaker for 1–2 h if the ligand is very soluble or if used from a stock solution. Incubate for 16–24 h for ligands with limited solubility or if added as powder (*see* **Note 11**).
2. Following incubation, spin the mix to remove undissolved ligand or aggregates (13,000 × *g* for 15 min).
3. Concentrate the protein complex until the desired concentration for crystallogenesis is achieved (>10 mg/mL) (*see* **Note 12**).
4. Screening and Matrix optimization is carried out as described previously. Apo-protein crystals can often be used to assist in the nucleation of the complex.
5. Co-crystallization can also be achieved by adding a small volume of ligand stock directly into the crystallization drop already containing the protein sample, prior to the addition of the well solution. A good precaution is to test the solubility of the ligand solution in the presence of the precipitant as a control and to adjust the concentration of ligand in the drop. Set up a control drop containing sample buffer, add the ligand solution (<10% of total drop volume if it is in volatile solvents or DMSO) and then the well solution (volumes as per standard screening protocols). Incubate and check next day for precipitation or crystal growth from the ligand. Readjust the concentration of the ligand solution accordingly to avoid precipitation or change precipitant conditions if crystals of the ligand appear.
6. Crystal soaking: dissolve the ligand in the mother liquor (taken from the reservoir solution where the crystals had grown). Prepare a drop containing 5–10 μL of this solution in a sitting drop well. Transfer a few crystals from the crystallization drop into the ligand solution drop using a cryo-loop. Try to minimize the amount of liquid carried on from the original drop. Let the crystals soak for a couple of minutes and then fish them with the loop and freeze. Optimizing the soaking may take a few tests until the concentration of ligand is such that crystals do not crack, but that it is diffusing into the crystal in sufficient amounts to visualize it in the electron density maps.
7. Gradual soaking can be done using increasing amounts of ligand in the mother liquor. This is a more gentle approach and may prevent crystals from cracking or dissolving. If soaks are done with small concentrations of ligand we recommend extending the time of soaking to a few hours or even several days The soaks will be done in a sealed sitting drop plate for easy manipulation and the buffer changes can be done by simply aspirating 80% of the drop volume and then adding a drop with the new buffer. No need to mix, diffusion will take care of it. The gradual soaking can also include an initial step

of "soaking-out" phosphate or buffer components from the crystal by using a phosphate-free buffer.

3.8 Cryoprotection and Crystal Freezing

Prior to data collection crystals will require cryoprotection and freezing. Deciding on a suitable cryoprotectant can be difficult. However, as a general rule, crystals grown from polyethylene glycol (PEG) containing conditions will often tolerate the addition of low molecular weight PEGs (PEG 200, 400) at 10–20 % or glycerol (20–30 %), which will be sufficient to cryoprotect the crystal. For crystals of complexes with ligands, the cryoprotectant should contain a similar amount of ligand as used for crystallization to avoid soaking it out.

If crystals are grown in ammonium sulfate the best option would be to try an oil as the starting cryoprotectant; alternatively glycerol (20–30 %) or small PEGs may also work. Another option, longer and more risky for the integrity of the crystals, is to "desalt" the crystals by soaking them in a drop containing decreasing amounts of ammonium sulfate (decrease by 100 mM at a time) with increasing amounts of PEG (from 0 to 20 %). These soaks can be done changing the buffer every hour if the crystals are robust. If they start to crack, then leave them in the drop for a few hours and leave the last soak overnight (exchange the buffer as directed for the ligand soaking procedure).

3.8.1 Cryofreezing Using Glycerol or PEG

1. Use a large cryo-loop to check the status of your mother liquor (equilibrated reservoir) and determine if it is forming a vitreous glass. Using the largest (0.7 mm) loop available, pick up some of the liquor and plunge freeze it in LN_2 using a cryo-wand. Maintaining the wand and loop under the LN_2 bring the loop towards the surface of the LN_2 so you can observe the loop region itself. If the loop appears white then you will need to add additional components or find an alternative cryoprotectant, if, however, the loop is clear then you can be confident that the solution will not result in the formation of hexagonal ice.
2. If working in small volume drops, add 1 μL of mother liquor to the drop, this can make the mounting significantly easier. Add 1 μL of cryoprotectant solution (made with the reservoir solution), wash twice, use cryo-loops to lift the crystal from the drop and transfer to a second drop (1 μL) of cryoprotectant solution, then fish immediately and plunge freeze.

3.8.2 Cryofreezing in Oils

Oils can be excellent cryoprotectants; this is particularly the case when ligand complexes are being sought, as the oils will minimize any back soaking or competition effects from cryo-components that may decrease ligand occupancy.

1. Harvest the crystal as normal and then transfer the crystal to a drop containing 1 μL of the oil of choice. More viscous oils

such as Parabar 10312 will require more care to minimize mechanical damage during the crystal manipulation. Passing the crystal repeatedly through the air oil interface will cause the remaining mother liquor around the crystal to slowly be replaced, leaving the crystal surrounded by oil.

2. Prior to plunge freezing the crystal it is often useful to touch the tip of the loop on the surface of the plate. This will allow the excess oil to run off the surface and leave a loop containing only a very thin film of oil with the crystal suspended in it. Minimizing the oil in this way will not only enhance the subsequent rate of freezing but reduces the background noise during data collection and makes centering the sample significantly easier!

3.9 Online Tools and Resources

3.9.1 Online Links and Useful Websites

1. The Protein Data Bank (PDB) is a publicly available online database where all macromolecule structures are deposited (determined by X-ray crystallography and NMR) [9]: http://www.rcsb.org
2. XtalPred-RF is an online tool that provides a crystallizability classification and it is an useful aid in construct design [25]: http://ffas.burnham.org/XtalPred-cgi/xtal.pl
3. Series of online tutorials regarding crystallography techniques and many related links, news and others. Created by Bernhard Rupp: http://www.ruppweb.org/default.htm
4. Protein Crystal Structure Propensity Prediction Server [26] from the Northeast Structural Genomics Consortium: http://nmr.cabm.rutgers.edu:8080/PXS/
5. SERp, Surface Entropy Reduction prediction server that helps to identify sites that are most suitable for mutation, to aid in redesigning your construct and enhance crystallizability [17]: http://services.mbi.ucla.edu/SER/
6. Rigaku EZ-Screen Builder, grid based optimization tool to help building suitable crystallization grids to optimize initial hits: https://www.rigakureagents.com/escreentoolkit/escreen.html
7. Helpful tips and advice on standard procedures regarding crystallogenesis: https://hamptonresearch.com/growth_101_lit.aspx
8. Online tutorial covering the basic principles of X-ray crystallography: http://www.ruppweb.org/Xray/101index.html
9. Ray Salemme's website including links and references to protein structure and analysis, structure based drug discovery and many other useful links: www.beta-sheet.org
10. Videos on how to freeze crystals: https://www.youtube.com/watch?v=FENUWRYXMOM, https://www.youtube.com/watch?v=QcsaWowulDM

11. Demonstration on how to use the mosquito automated crystallization setup system https://www.youtube.com/watch?v=BZkyRu_UMw8
12. Structural Genomics Consortium, Oxford, UK: http://www.thesgc.org/

3.9.2 Databases Containing Information on Crystallization Conditions for Protein Phosphatases

1. Protein Data Bank (PDB) [9]: http://www.rcsb.org
2. Marseille Protein Crystallization Database (MPCD) [27]: www.cinam.univ-mrs.fr/mpcd/
3. Biological Macromolecule Crystallization Database (BMCD) [28]: http://xpdb.nist.gov:8060/BMCD4/index.faces
4. Large-Scale Structural Analysis of the Classical Human Protein Tyrosine Phosphatome [6] http://www.thesgc.org/resources/phosphatases

3.9.3 Crystallography Textbooks

1. Principles of Protein X-ray Crystallography [29].
2. Introduction to Macromolecular Crystallography [21].
3. Biomolecular Crystallography: Principles, Practice, and Application to Structural Biology [22].
4. Data mining crystallization databases: Knowledge-based approaches to optimize protein crystal screens [30].

4 Notes

1. The information of the estimated PI of your construct (including the tag) is useful to decide on the best buffer conditions during purification of your protein, particularly if you are using ionic exchange chromatography. At a pH above the PI the protein will have a net negative charge. Below the PI it will have a net positive charge. However, it is better to determine the actual PI values using isoelectric-focusing electrophoresis, this will also show if your sample has multiple isoform species with slightly different PI. This is another source of heterogeneity that it is worth minimizing if crystals are difficult to grow. A change in salt concentration or pH conditions may alleviate the issue followed by effective ion-exchange chromatography to separate the isoforms.
2. Some of the solids may not dissolve completely right away but they may eventually during culture growth.
3. Plastic plates should be cleaned prior to filling with screen solutions. Use compressed air to spray and remove dust and plastic debris. Alternatively you can wash with distilled water and leave to dry inverted on absorbent paper.

4. It is good practice to check the drops right after setting up the plate. If precipitate has already formed in all drops, this is a sign that protein concentration may be too high. If drops remain clear then check after 24 h. Monitoring the dynamics of the crystallization is important to understand the behavior of your sample and to aid decisions as to how to change concentration or buffer conditions during subsequent rounds of purification and sample preparation.
5. Single crystals with a clean and sharply faceted appearance are most suitable for crystallographic study. Clusters, needles and crystalline precipitates are all excellent sources of seeds for optimization.
6. As drop volumes increase the equilibration time of the drop will also increase. When using 5 + 5 μL or larger drops it is not uncommon for crystals not to form for several weeks following initial setup.
7. Seed stocks can be flash frozen and stored for future use.
8. Seed stocks can be serially diluted to control the levels of nucleation achieved (10^{-9}, 10^{-12} dilutions are typically used).
9. In order to maximize the stability of our seed stock we utilize a multi-aspirate step when dispensing this component. The robot will first aspirate 20 nL of seed stock before moving to the crystal screen and aspirating a further 200 nL of screen into the same tip as the seed. The high concentration of precipitants in the screen aid the stability of the seeds and the increased volume of liquor in the tip helps to ensure a more accurate dispense. It is important to always have seeds and screen mixed before adding them to the protein.
10. Modern synchrotron beam lines allow excellent data to be collected from even relatively small crystals (10–50 μm). The size of crystal necessary for successful data collection will vary on a protein-by-protein basis. Home source X-ray systems require larger crystals (100–200 μm).
11. The final concentration of the ligand will depend upon its maximal solubility and the effective concentration for binding. As a rule-of-thumb we recommend using a final concentration 10–20 times the K_m or K_i value. If the binding affinity is not known, use a molar ratio of at least 2:1 ligand–protein or higher if it is a small molecule. For peptide ligands 1–2 mM usually is sufficient for co-crystallization, for small molecules use 0.5–1 mM. If feasible, we recommend adding an excess of ligand to overcome potential competition with any ions or buffer bound to the active site, particularly if crystal soaking is used.
12. The concentration achievable for the complex may significantly differ to that of the Apo protein.

Acknowledgements

This work was supported by MRC funding (MR/K011049/1) to LT and BBSRC Doctoral Training Studentship, Sir Kenneth Murray Scholarship, Manchester Presidents Doctoral Scholar Award and a SCI Scholarship to JA. CL is a senior experimental officer at the Manchester Protein Structure Facility, University of Manchester. We thank Darel Macdonald, Andrew Currin, Efrain Ceh Pavia, and Deepankar Gahloth for sharing their data and protocols.

References

1. Tonks NK (2006) Protein tyrosine phosphatases: from genes, to function, to disease. Nat Rev Mol Cell Biol 7(11):833–846
2. Tautz L, Pellecchia M, Mustelin T (2006) Targeting the PTPome in human disease. Expert Opin Ther Targets 10(1):157–177
3. Eswaran J et al (2006) Crystal structures and inhibitor identification for PTPN5, PTPRR and PTPN7: a family of human MAPK-specific protein tyrosine phosphatases. Biochem J 395:483–491
4. Nunes-Xavier C et al (2011) Dual-specificity MAP kinase phosphatases as targets of cancer treatment. Anticancer Agents Med Chem 11(1):109–132
5. Rios P et al (2014) Dual-specificity phosphatases as molecular targets for inhibition in human disease. Antioxid Redox Signal 20(14):2251–2273
6. Barr AJ et al (2009) Large-scale structural analysis of the classical human protein tyrosine phosphatome. Cell 136(2):352–363
7. Tabernero L et al (2008) Protein tyrosine phosphatases: structure-function relationships. FEBS J 275(5):867–882
8. Bohmer F et al (2013) Protein tyrosine phosphatase structure-function relationships in regulation and pathogenesis. FEBS J 280(2): 413–431
9. Berman HM et al (2000) The protein data bank. Nucleic Acids Res 28(1):235–242
10. Sambrook J, Russell DW (2001) Molecular cloning: a laboratory manual, 3rd edn. Cold Spring Harbor Laboratory Press, New York, p A2.2
11. Mitigen. Available from http://www.mitegen.com/
12. Molecular Dimensions. Available from http://www.moleculardimensions.com/
13. Hampton Research. Available from http://hamptonresearch.com
14. Qiagen. Available from https://www.qiagen.com/gb/
15. Hampton Research, product detail. Available from http://hamptonresearch.com/product_detail.aspx?sid=152&pid=445
16. XtalPred. Available from http://ffas.burnham.org/XtalPred-cgi/xtal.pl
17. Goldschmidt L et al (2007) Toward rational protein crystallization: a web server for the design of crystallizable protein variants. Protein Sci 16(8):1569–1576
18. Cantor CR, Schimmel PR (1980) Schimmel biophysical chemistry, Part 2: Techniques for the study of biological structure and function (Pt. 2). W. H. Freeman, New York
19. Phillips K, de la Peña AH (2011) The combined use of the Thermofluor assay and ThermoQ analytical software for the determination of protein stability and buffer optimization as an aid in protein crystallization. Curr Protoc Mol Biol Chapter 10:Unit 10.28
20. Beta-sheet. Available from http://www.beta-sheet.org/resources/T11-Crystallization-ericsson.pdf
21. McPherson A (2009) Introduction to macromolecular crystallography, 2nd edn. Wiley, Hoboken, NJ
22. Rupp B (2009) Biomolecular crystallography: principles, practice, and application to structural biology. Garland Science, New York
23. D'Arcy A, Villard F, Marsh M (2007) An automated microseed matrix-screening method for protein crystallization. Acta Crystallogr D Biol Crystallogr 63(Pt 4):550–554
24. Rigaku EZ screen. Available from https://www.rigakureagents.com/escreentoolkit/escreen.html

25. Slabinski L et al (2007) XtalPred: a web server for prediction of protein crystallizability. Bioinformatics 23(24):3403–3405
26. Price WN et al (2009) Understanding the physical properties that control protein crystallization by analysis of large-scale experimental data. Nat Biotechnol 27(1):51–57
27. Charles M, Veesler S, Bonnete F (2006) MPCD: a new interactive on-line crystallization data bank for screening strategies. Acta Crystallogr D Biol Crystallogr 62(Pt 11):1311–1318
28. Tung M, Gallagher DT (2009) The Biomolecular Crystallization Database Version 4: expanded content and new features. Acta Crystallogr D Biol Crystallogr 65:18–23
29. Mesters J (2007). In: Drenth J (ed) Principles of protein X-ray crystallography. Practical Protein Cristallization. Springer, New York. pp 297–304
30. Kimber MS et al (2003) Data mining crystallization databases: knowledge-based approaches to optimize protein crystal screens. Proteins 51(4):562–568