Journal of Visualized Experiments

Optimizing the growth of endothiapepsin crystals for serial crystallography experiments.

--Manuscript Draft--

Article Type:	Invited Methods Article - JoVE Produced Video
Manuscript Number:	JoVE61896R2
Full Title:	Optimizing the growth of endothiapepsin crystals for serial crystallography experiments.
Corresponding Author:	John Henry Beale Paul Scherrer Institut Villigen PSI, Aargau SWITZERLAND
Corresponding Author's Institution:	Paul Scherrer Institut
Corresponding Author E-Mail:	john.beale@psi.ch
Order of Authors:	John Henry Beale
	May Marsh
Additional Information:	
Question	**Response**
Please indicate whether this article will be Standard Access or Open Access.	Open Access (US$4,200)
Please indicate the **city, state/province, and country** where this article will be **filmed**. Please do not use abbreviations.	Villigen-PSI, Aargau, Switzerland
Please confirm that you have read and agree to the terms and conditions of the author license agreement that applies below:	I agree to the Author License Agreement
Please specify the section of the submitted manuscript.	Biochemistry
Please provide any comments to the journal here.	

TITLE:

Optimizing the growth of endothiapepsin crystals for serial crystallography experiments

AUTHORS AND AFFILIATIONS:

John H. Beale[1]

May E. Marsh[1]

[1]Swiss Light Source, Paul Scherrer Institut, Villigen, Switzerland

may.sharpe@psi.ch

Corresponding Author:

John H. Beale

john.beale@psi.ch

KEYWORDS:

Serial crystallography, batch crystallization, micro-crystallization, XFEL, vapor diffusion, endothiapepsin

SUMMARY:

The aim of this article is to give the viewer a solid understanding of how to transform their small-volume, vapor-diffusion protocol, for growing large, single protein crystals, into a large-volume batch micro-crystallization method for serial crystallography.

ABSTRACT:

Here, a protocol is presented to facilitate the creation of large volumes (> 100 µL) of micro-crystalline slurries suitable for serial crystallography experiments at both synchrotrons and XFELs. The method is based upon an understanding of the protein crystal phase diagram, and how that knowledge can be utilized. The method is divided into three stages: (1) optimizing crystal morphology, (2) transitioning to batch, and (3) scaling. Stage 1 involves finding well diffracting, single crystals, hopefully but not necessarily, presenting in a cube-like morphology. In Stage 2, the Stage 1 condition is optimized by crystal growth time. This strategy can transform crystals grown by vapor diffusion to batch. Once crystal growth can occur within approximately 24 h, a morphogram of the protein and precipitant mixture can be plotted and used as the basis for scaling (Stage 3). When crystals can be grown in batch, scaling can be attempted, and the crystal size and concentration optimized as the volume is increased. Endothiapepsin has been used as a demonstration protein for this protocol. Some of the decisions presented are specific to endothiapepsin. However, it is hoped that the way they have been applied will inspire a way of thinking about this procedure that others can adapt to their own projects.

INTRODUCTION:

Room temperature (RT) macromolecular crystallography is now starting to become popular again within the structural biology community. The development of X-ray Free Electron Laser (XFEL) light sources has spurred the development of RT sample delivery approaches[1–4], and these methods have now been applied to synchrotrons[5–8]. Not only do RT methods open up the possibility of pump-probe experimental strageties[9–12], but there is also mounting evidence that they promote alternative conformational states[13–17].

However, the principal reason why cryo-methods gained traction over RT approaches in the late 1990s was the slowing of radiation damage by sub-zero crystal temperatures[18]. Cryo-methods[19] began to allow for the collection of a complete dataset from a single protein crystal. Modern RT methods at XFELs and synchrotrons solved the problem of single-crystal radiation damage by the development of rapid (> 100 Hz) crystal delivery strategies[1–4]. These methods allow for the collection of a complete dataset from thousands of individually exposed crystals. These RT delivery approaches therefore require the production of large quantities of solutions containing homogenous micro-crystals (> 100 µL of < 50 µm crystals). However, since cryo-methods only require single crystals, methods to create such micro-crystalline slurries are currently not ubiquitous across protein crystallography laboratories.

There are examples in the literature of how to do parts of the micro-crystallization optimization procedure for serial crystallography samples. Here, a distinction should be made between membrane and soluble proteins. Protocols to optimize the growth of micro-membrane protein crystals grown in monoolein (or some other lipid), for lipidic cubic phase (LCP), have been well described[20–22]. However, methods for the micro-crystallization of soluble proteins, including membrane proteins grown in non-lipid like conditions, are generally lacking. Previous studies have focused on specific parts of the process, such as micro-crystal screening[23,24], enhancing nucleation[24], and scaling using free-interface diffusion[25], but not a complete method.

However, a method was recently described[26] that attempts to offer a complete protocol. Like many aspects of protein crystallography, it is not new. Many of the ideas proposed were already described by Rayment (2002)[27]. The method aims to show crystallographers how to perform the conversion from a single, crystal grown using vapor diffusion, to a batch methodology to grow thousands of crystals. The method focuses on vapor diffusion as a common starting point as 95% of all Protein Data Bank (PDB) depositions come from crystals grown in vapor diffusion plates[26]. Vapor diffusion is, however, not the ideal method for micro-crystallization[26], so a methodology is described to convert vapor diffusion to batch crystallization. Once crystals can be grown in batch, scaling routes to larger volumes become more practicable. Given the vagaries of protein crystallization, the authors would stress that this method is not failsafe. However, the protocol should, at least, provide an insight into the 'crystallization space' of a protein.

This method relies on the protein crystallization phase diagram and how an understanding of that diagram can act as a guide during micro-crystallization optimization. A protein phase diagram is commonly depicted as an *x*/*y* plot with precipitant and protein concentrations on the *x* and *y* axes, respectively (**Figure 1A**). From the pure water point (bottom left corner - **Figure 1A**),

the concentration of both protein and precipitant increases until the solubility line is reached. The solubility line marks the point of supersaturation (purple line - **Figure 1A**). When a protein is supersaturated, the solution becomes thermodynamically unstable and will begin separating into two phases: 'protein-rich' and a stable saturated solution. This separation can occur anywhere beyond the solubility line and its kinetics are dependent upon the properties of the protein and the components of the solution.

When the protein and precipitant concentrations are too great, the protein will decompose unstably out of solution and result in amorphous precipitate (pink region - **Figure 1A**). However, ordered phase separation can occur in the nucleation region [see Garcia-Ruiz (2003)[28] for detailed description] and crystal nucleants have the propensity to form (green region - **Figure 1A**). Nucleation and growth removes protein from the solution and moves the drop into the metastable region where growth can continue until the solubility line is reached [see McPherson and Kuznetsov (2014)[29] for detailed discussion]. The diagram is, for the vast majority of crystallization conditions, a gross oversimplification[30]. Regardless of this however, the diagram is still of great utility for micro-crystallographers as the mapping of the diagram allows for the solubility line and the kinetics of nucleation to be determined.

In terms of creating micro-crystals, the two factors during crystallization that need to be optimized are the number of crystals (X_n) and their mean, longest dimension (X_s). X_n will be proportional to the number of nucleation events (n) (Eq. 1).

$$X_n \propto n \qquad \textit{Eq. 1}$$

X_s is proportional to the concentration of free protein above the solubility line (P_s) divided by X_n (Eq. 2).

$$X_s \propto \frac{P_s}{X_n} \qquad \textit{Eq. 2}$$

In a perfect situation, every nucleation event would yield a possible crystal and every one of these crystals would have equal access to the available protein in solution. **Figure 2** is a graphical representation from an ideal scenario of the relationship between X_n and X_s. Practically, the principal control a crystallographer has over X_n and X_s is by influencing the amount of nucleation or by the addition of seed crystals. The micro-crystallographer must judge how to increase X_n such that a suitable crystal concentration and crystal size can both be created.

The majority of crystallization techniques require a 'transitionary period' (**Figure 1B**). For example, in a vapor diffusion experiment, upon mixing the protein and precipitant solutions, the concentrations of each will change as the drop equilibrates with the well solution. One hopes that these changes will gradually transition the drop into the nucleation zone where the propensity for crystallization will increase. As crystals begin to nucleate and grow, the amount of protein in solution will begin to fall, decreasing the probability of further nucleation. The ultimate

amount of nucleation will be protein and condition specific, and also dependent upon the depth of penetration of into the nucleation zone. Given the limited nucleation zone penetration of methods that require a transitionary step, the level of nucleation will ultimately be limited to the rate of nucleation at the metastable-nucleation region boundary.

Due to the importance of being able to enhance the level of nucleation for a micro-crystallographer, it is important to move to a batch crystallization methodology. Batch can take greater advantage of the whole nucleation region (**Figure 1C**). In batch methods, the idea is to mix the protein and precipitant together such that a supersaturated solution is created without the need of any changes in component concentrations. Nucleation should be possible immediately upon mixing. Batch methods therefore allow for the entire nucleation zone to be theoretically reached. Any increase in nucleation kinetics beyond the metastable-nucleation boundary can then be utilized.

If the basal-level of crystal nucleation is not enough to generate a large X_n, micro-seeding methods can be used. In micro-seeding, pre-grown crystals are broken up to create a slurry of crystalline fragments which can act as a scaffold for fresh crystal growth[31, 32]. Micro-seeding has been widely used in serial crystallographic sample preparation as a way to increase X_n without the need of increasing crystal nucleation (**Figure 1C**).

The transition from vapor diffusion to batch can be visualized on a phase diagram as moving the experimental starting point from either the non-supersaturated or metastable regions to the nucleation zone. This can be done by increasing the protein and/or precipitant concentrations, and/or the ratio of the two within the drop (**Figure 1D**), and observing which conditions yield crystals appearing rapidly (< 24 h)[26]. Complete vapor diffusion drop equilibration can take days or weeks[33]. Therefore, by looking for conditions that show rapidly appearing crystals, batch conditions can be found without having to move to alternative crystallization screening formats such as micro-batch[34–37].

Once the nucleation zone has been found, a batch condition has been found and a morphogram – here, a rough phase diagram - can be created. The morphogram is of great utility when contemplating whether to use a seeded-batch or straight batch protocol. By plotting the X_n as a function of the protein and precipitant concentration, an assessment of the nucleation kinetics can be made[26]. If X_n remains low across the whole nucleation region, seeded-batch may be required to make X_n large enough to limit crystal growth. This assessment is the first step in the process of scaling to larger volumes (> 100 µL).

This method was designed such that it could be conducted in the majority of crystallization laboratories by using standard vapor diffusion crystallization equipment. Many studies have also been conducted which describe techniques to facilitate many parts of this process, should the equipment be available. These include, but are not limited to, dynamic light scattering (DLS)[25,27], non-linear imaging[20,24,25], powder diffraction[20,24,27], and electron microscopy[26] [see Cheng et al. (2020)[40] for a nice review].

The aim of this work is to provide a visual demonstration of the method to transition from small volume (< 500 nL) vapor diffusion crystallization to large volume (> 100 µL) batch crystallization. Endothiapepsin from *Cryphonectria parasitica* has been used as an example system for demonstrating this translation. The type of experiment and sample delivery method that the micro-crystals are required for will influence the ideal X_s output[26]. For mixing experiments requiring a millisecond time resolution[41] or gas-dynamic virtual nozzles[42], a final X_s of < 5 µm may be desirable. In this case, the goal was to produce protein crystals that diffract to approximately 1.5 Å, for a photon-activated pump-probe experiment, and using a fixed-target delivery approach.

To give an illustration of the sample requirements of such a serial crystallography experiment using endothiapepsin, **Table 1** shows the experimental parameters of a hypothetical experiment. The sample information was based upon the protocol described below. Given some conservative estimates on hit rates and data collection requirements, 50 mg is the total sample consumption estimate for the whole experiment.

Figure 3 shows a flow-chart of the complete optimization process from initial small volume vapor diffusion crystallization to large scale batch. For the majority of serial crystallography projects, this protocol will begin at Step 2: 'transitioning to batch', since the target protein will already have been crystallized. However, Step 1 has been included for completeness and to remind readers of its importance. Finding a condition that gives rise to a well diffracting, single, large crystal is the best starting point for micro-crystal optimization. In Step 2, this condition can then be optimized from vapor diffusion to batch, and a morphogram of the nucleation and metastable regions can be plotted. Once this has been done, scaling the batch condition to larger volumes can be performed in Step 3. By the end of the flow-chart, a crystallographer will have created a repeatable, large-volume (> 100 µL), micro-crystallization, batch protocol for endothiapepsin. This method can then be applied to their particular protein of interest.

PROTOCOL:

NOTE: All 96-well sitting-drop crystallization experiments were setup using either 2 or 3-drop plates. A liquid handling robot and a crystallization imager/hotel were used to facilitate the preparation and monitoring of all 96-well screens. All reagent concentrations for crystallization experiments are given at their starting concentrations prior to mixing.

1. **Optimizing crystal morphology**

NOTE: Steps 1.1.1. and 1.1.6. describe how endothiapepsin crystallization conditions were found, and how these conditions were optimized to find a single condition that yielded single, well-diffracting crystals.

1.1. Sparse-matrix optimization

1.1.1. Prepare fresh endothiapepsin solution.

NOTE: Endothiapepsin, when procured as Superan 600, must be buffer transferred out of its storage solution and concentrated.

1.1.1.1. Prepare 3 L of 0.1 M Na Acetate pH 4.6 at 4 °C.

1.1.1.2. Cut 20 cm of dialysis tubing and briefly wash in the buffer. Seal one end of the tubing using a clip, place 50 mL of the endothiapepsin solution into the tubing and then seal the other end.

1.1.1.3. Leave the solution to dialyze for at least 2 h at 4 °C in 1 L of the Na Acetate buffer. Due to the components of the storage buffer, the solution in the dialysis bag will now be approximately 100 mL.

1.1.1.4. Transfer the dialysis bag containing the endothiapepsin into a fresh liter of 4 °C, 0.1 M Na Acetate pH 4.6. Repeat this step once more such that the original buffer has been diluted 2000x against the Na Acetate.

1.1.1.5. The endothiapepsin will now be at approximately 10 mg/mL. Concentrate to 100 mg/mL using a 10 kDa centrifugal concentrator and a centrifuge.

1.1.1.6. Flash cool in the endothiapepsin solution in liquid nitrogen in 50 µL aliquots and store at -80 °C.

1.1.2. Prepare a PACT Premier 96-well sparse-matrix screen.

1.1.2.1. Using a liquid handling robot, dispense 100 nL of 70 mg/mL endothiapepsin and 100 nL of well solution into a single sub-well per well. Mix the protein and well solution 3 times upon addition of the crystallization buffer.

1.1.2.2. Seal the plate and leave for 28 days at 20 °C taking images every day for the first week and then every week thereafter for 4 weeks.

1.1.3. Sparse-matrix analysis

1.1.3.1. Identify hits that produce single endothiapepsin crystals. From the PACT screen, conditions that contained $MgCl_2$ grew as singletons rather than needle clusters.

1.1.4. Sparse-matrix optimization

1.1.4.1. From the $MgCl_2$ containing conditions identified in Step 1.1.3.1, create a 96-well screen randomly combining and varying the different well components.

1.1.4.2. Using a liquid handling robot, dispense 100 nL of 70 mg/mL endothiapepsin and 100 nL of well solution into a single sub-well per well. Mix the protein and well solution 3 times upon addition of the crystallization buffer.

1.1.4.3. Seal the plate and leave for 28 days at 20 °C taking images every day for the first week and then every week thereafter for 4 weeks.

1.1.5. Optimization analysis

1.1.5.1. Using a suitable spreadsheet software, rank the crystallization conditions that give rise to crystals based on the crystal quality and precipitation level, poor (0) to ideal (5) and low (0) to high (5), respectively. With respect to crystal quality, the broad criteria are single crystals with a box-like morphology.

1.1.5.2. Perform a Pearson's correlation analysis between the crystallization condition contents and the crystal quantity and precipitation level.

1.1.5.3. Plot these data as a heat map. Look for components and conditions that were correlated with the preferred outcomes.

1.1.6. **Diffraction analysis.**

1.1.6.1. Confirm that the crystals grown from the identified conditions in Step 1.1.5 are suitable for serial crystallography by performing an X-ray diffraction experiment.

1.1.6.2. Load a sample of the endothiapepsin crystals from each of the identified conditions onto supports that allow for data collection at either 100 or 293 K and perform an X-ray diffraction experiment. If working under cryo, use 25% ethylene glycol as the cryo-protectant.

1.1.6.3. Process these data *via* a suitable software suite. Endothiapepsin crystals should diffract to beyond 1.5 Å. Check for twinning, as twinned crystals can significantly complicate serial crystallographic data processing.

1.1.6.4. If crystals are singletons and diffract to 1.5 Å proceed to Step 2. If not, go back to Step 1.1.2 and try more sparse-matrix screens to identify promising conditions. After the analyses conducted in Steps 1.1.5. and 1.1.6., a crystallization condition of 25% (w/v) PEG 6,000, 0.1 M Tris-HCl pH 7.0 and 0.15 M $MgCl_2$ should have been found as the approximate ideal.

2. **Transitioning to batch**

2.1. Morphogram experiment

2.1.1. Create a micro-crystal seed stock.

NOTE: It is best practice when making seed-stocks, to make the seeds from crystals specifically grown for the task. This greatly helps with reproducibility. Other ideas presented in Steps 2.1.1.1 to 2.1.1.11 are to always use the crystals grown from a standard number of wells – here 5 – and aliquot the stocks once they are made to negate freeze-thaw cycles.

2.1.1.1. Prepare a 96-well crystallization plate with wells containing the crystallization buffer: 25% (w/v) PEG 6,000, 0.1 M Tris-HCl pH 7.0 and 0.15 M $MgCl_2$.

2.1.1.2. Using a liquid handling robot, dispense 200 nL of defrosted 70 mg/mL endothiapepsin and 200 nL of well solution into a single sub-well per well. Mix the protein and well solution 3 times upon addition of the crystallization buffer.

2.1.1.3. Seal the plate and leave for 24 h.

2.1.1.4. Fill a 1.5 mL centrifuge tube with 250 µL of crystallization buffer and 10-15 1 mm glass beads. Leave the centrifuge tube on ice to cool for 5-10 min.

2.1.1.5. Select 5 wells with crystals, open the wells with a scalpel and, using a pipette tip, crush the crystals in the wells.

2.1.1.6. Aspirate 1 µL of buffer from the iced centrifuge tube and use to homogenize the crushed crystal slurry. Once homogeneous, aspirate the entire slurry and collect in the cooled centrifuge tube.

2.1.1.7. Repeat Step 2.1.2.6 for each of the 5 sub-wells.

2.1.1.8. Vortex the centrifuge tube containing the buffer, pooled slurries and beads at 1000 rpm for 30 s.

2.1.1.9. Return the centrifuge tube to ice for 30 s.

2.1.1.10. Repeat Steps 2.1.2.8 and 2.1.2.9 two more times.

2.1.1.11. The seed-stock is now ready and can be aliquoted into 10 µL batches and stored at -20 °C.

2.1.2. Perform morphogram experiment.

2.1.2.1. Prepare a 2-drop 96-well grid screen. Vary the concentration of PEG 6,000 from 5 to 40% (w/v) along the plate columns, keeping the buffer and salt at 0.1 M Tris-HCl pH 7.0 and 0.15 M $MgCl_2$, respectively.

2.1.2.2. Prepare a sequential dilution of endothiapepsin in 0.1 M Na Acetate pH 4.6 from 100 to 12.5 mg/mL over 8 steps. A different concentration of endothiapepsin will be used for each row of the plate.

2.1.2.3. Using a liquid handling robot, dispense 150 nL of endothiapepsin into both sub-wells 1 and 2. In sub-well 1, dispense 150 nL of the well solution. In sub-well 2, multi-aspirate 50 nL of defrosted seed-stock and 100 nL of well solution, and then dispense both into the protein solution. Mix the solutions 3 times upon addition of the crystallization buffer.

2.1.2.4. Seal the plate and leave at 20 °C taking images every 0, 3, 6, 12, 18, 24 h, then every day for the first week, and every week for the next four. If automatic imaging is not possible, do not worry about the hourly imaging on day 1.

2.2. Morphogram analysis

2.2.1. Looking at the images taken after 24 h, estimate the number of crystals that are present in each well and record these estimates in the “morphogram generator” worksheet provided. These estimates do not have to be precise; individually counting thousands of micro-crystals, if present, is not practical or necessary. Principally try to ensure the estimates are consistent over the whole plate.

NOTE: The 24 h rule was based upon the observations made in Beale et al. (2019)[26]. Vapor diffusion crystallization conditions can take days or weeks to equilibrate. Crystals that appear rapidly are more likely to have grown *via* a batch process rather than by the gradual equilibration of the drop components. The 24 h criterion is, therefore, somewhat arbitrary and an exact cut-off time between a batch and vapor diffusion experiment will depend on the specific mixture of the condition [see Beale et al. (2019)[26] for full details].

2.2.2. Input the starting concentrations of endothiapepsin and PEG 6,000 in the boxes indicated.

2.2.3. The worksheet will automatically plot the results in the traditional phase diagram format with precipitant and protein concentration on the *x* and *y* axes, respectively. Well conditions that only give rise to crystals in their seeded drops indicate the metastable region of the diagram (transparent blue), whereas conditions that have crystals in both the seeded and non-seeded drops indicate the nucleation zone (solid green).

NOTE: Ideally, the majority of the nucleation zone should be present on the diagram (i.e., there are some clear wells on the bottom of the diagram and some precipitate should be visible at high protein and precipitant concentrations. If this is not the case, perhaps, repeat the experiment but increase the protein and/or precipitant concentration (if possible).

2.2.4. If crystal have appeared in less than 24 h, proceed to Step 2.3.1. If not, proceed to Step 2.4 and continue optimizing towards batch.

2.3. Crystal analysis

2.3.1. As said at the end of Step 1, before moving to the next step, ensure these crystals have the desired morphology and diffraction quality. With regard to morphology, are the crystals untwinned (i.e., are mainly observably single crystals forming, rather than needle-ball-like or fan-like structures)? With regard to diffraction, collect diffraction data from the crystals if possible. If these crystals do not diffract, it is improbable that the crystals grown in a larger volume will diffract.

2.3.2. Load a sample of the endothiapepsin crystals from the morphogram experiment onto supports that allow for data collection at either 100 or 293 K and perform an X-ray diffraction experiment. If working under cryo, use 25% ethylene glycol as the cryo-protectant.

2.3.3. Process these data *via* a suitable software suite. Endothiapepsin crystals should diffract to beyond 1.5 Å. Across the sample of crystals, also observe the cell size, the total number of observations, and the mosaicity; these measures will give an indication as to the homogeneity of the diffracting crystals.

2.3.4. If the crystal morphology and diffraction quality is sufficient, proceed to Step 3.

2.4. Optimize crystal growth time.

NOTE: The morphogram analysis (Step 2.2) will have given an indication of the crystallization starting point: in which region of the phase diagram is the drop located when the precipitant and protein solutions were mixed. Is the drop in the metastable region or below the solubility line? Batch crystallization begins in the nucleation zone (**Figure 1C**). The goal of this step is to move this starting point from either below the solubility line or metastable region, into the nucleation

zone (**Figure 1D**). If the seeded-drops from Step 2.2. have yielded crystals rapidly, this is an indication that the drop mixture is already in the metastable region, if not, then it is likely the drop is not supersaturated.

2.4.1. Optimizing crystal growth time.

2.4.1.1. Using the same screen as in Step 2.1.3, prepare a 96-well vapor diffusion crystallization experiment in a 3-drop plate.

2.4.1.2. Increase the starting protein concentration of endothiapepsin on the *y* axis (i.e., concentrate the protein further, perhaps 120 mg/mL for endothiapepsin).

2.4.1.3. Perform a serial dilution, as in Step 2.1.3.2, such that each row of the plate contains a sequentially lower protein concentration.

2.4.1.4. Use different drop ratios in each of the three drops on the plate: 1:1, 1:2, and 2:1, protein:precipitant.

2.4.1.5. View or image the plate on the first day at 0, 3, 6, 12, 18, 24 h and then every day for the first week, and every week for the next four. If automatic imaging is not possible, do not worry about the hourly imaging on day 1.

2.4.1.6. Identify drops that produce the most rapidly appearing crystals and makes these the starting points of repeated optimizations until crystal growth occurs with 24 h.

2.4.1.7. When a rapidly appearing crystal-condition has been identified return to Step 2.1 to replot the morphogram as a prelude to begin scaling.

3. **Scaling**

3.1. Rank scaling routes. At this stage, it is not necessary to decide on a single scaling route, only to identify and rank the options so that they can be explored in turn. As the volume of the batch mixture is increased during the scaling procedure, changes will occur in the rate of nucleation and the range of crystal sizes. However, these can be overcome by careful tweaking of component concentrations as the scaled volume is increased.

NOTE: Steps 3.1.1 and 3.1.2 describe how to discern, from the morphogram, whether a batch or seeded-batch protocol is more appropriate.

3.1.1. Straight batch protocol

3.1.1.1. Is the X_n in the nucleation zone proportional to the protein and/or precipitant concentration? *i.e.* the X_n **does** increase as a function of either precipitant and/or protein concentration – Yes – go to Step 3.1.1.2. – No – go to Step 3.1.2.

3.1.1.2. Locate conditions that yield crystals of the required size and go to Step 3.2.

3.1.2. **Seeded-batch protocol**

3.1.2.1. Is the X_n flat across the nucleation zone? *i.e.* the X_n **does not** increase as a function of either precipitant and/or protein concentration.

3.1.2.2. Locate seeded conditions that yield crystals of the required size and go to Step 3.2. If all crystals are too large – go to Step 3.1.2.3.

3.1.2.3. Repeat the morphogram experiment (Step 2.1) but this time increase the concentration of the seed-stock used in the seeded wells. The seed stock can be increased by using more crystals in its creation. For example, instead of 5 wells in Step 2.1.1.5, use 10 wells.

3.1.2.4. View or image the plate over the first 0, 3, 6, 12, 18, 24 h.

3.1.2.5. The X_n should have increased and the X_s decreased in the seeded-drops. Repeat this cycle if smaller crystals are needed and then follow a seeded-batch protocol.

3.2. Gradually scaling

3.2.1. Scaling in 96-well plates. From the endothiapepsin morphogram, a straight batch method using the crystallization condition 0.1 M Tris-HCl pH 7.0, 0.15 M $MgCl_2$, and 30% (w/v) PEG 6,000, was initially selected for scaling. 100 mg/mL endothiapepsin mixed with the crystallization buffer in a 1:1 ratio.

3.2.1.1. Prepare 2-3 wells in a 2-well 96-well sitting-drop plate with 100 µL of 0.1 M Tris-HCl pH 7.0, 0.15 M $MgCl_2$, and 30% (w/v) PEG 6,000.

3.2.1.2. Using freshly defrosted 100 mg/mL endothiapepsin solution, dispense 0.5 µL of protein and 0.5 µL precipitant per well, and seal.

3.2.1.3. View or image the plate over the first 0, 3, 6, 12, 18, 24 h. Note any changes in the range of X_s and X_n.

3.2.1.4. If changes have occurred, repeat Steps 3.2.1.1 to 3.2.1.2 but increase or decrease the protein, precipitant, and/or seed concentration to restore any changes to the range of X_s and X_n.

3.2.1.5. When/if the range of X_s and X_n are acceptable, proceed to Step 3.2.2.

3.2.2. Scaling in 24-well hanging-drop plates

3.2.2.1. Prepare a single well of a 24-well hanging-drop plate by greasing the edges of the well with vacuum grease.

3.2.2.2. Prepare 0.5 mL of 0.1 M Tris-HCl pH 7.0, 0.15 M $MgCl_2$, and 30% (w/v) PEG 6,000 and fill the greased well.

3.2.2.3. Using freshly defrosted endothiapepsin solution, pipette 1 µL of protein onto the surface of a glass coverslip. Pipette 1 µL of crystallization buffer onto the protein drop and mix using the pipette.

3.2.2.4. View or image the plate over the first 0, 3, 6, 12, 24 h. Note any changes in the range of X_s and X_n.

3.2.2.5. If changes have occurred, repeat Steps 3.2.2.1 to 3.2.2.4 but increase or decrease the protein, precipitant, and/or seed concentration to restore any changes to the range of X_s and X_n.

3.2.2.6. When/if the range of X_s and X_n are acceptable, proceed to Step 3.2.2.7.

3.2.2.7. Repeat Steps 3.2.2.1 to 3.2.2.5, increasing the total volume of the experiment gradually to 10 µL.

3.2.2.8. Once at a volume of 10 µL or larger, proceed to centrifuge tubes in Step 3.2.3.

3.2.3. Scaling in centrifuge tubes

NOTE: The refinement of the endothiapepsin batch condition principally happened at the point of 200 µL volumes (see Scaling). The process began with a crystallization condition of 0.1 M Tris-HCl pH 7.0, 0.15 M $MgCl_2$, and 30% (w/v) PEG 6,000. However, the PEG concentration ultimately changed to 40% (w/v). Seeds were also required to control the X_n, and to prevent crystals growing too large, crystal growth had to be quenched. Steps 3.2.3.1 to 3.2.3.7 detail the process of condition optimization. Step 3.2.4. describe the final batch protocol.

3.2.3.1. Prepare crystallization buffer: 0.1 M Tris-HCl pH 7.0, 0.15 M $MgCl_2$, and 30% (w/v) PEG 6,000.

3.2.3.2. Using freshly defrosted 100 mg/mL endothiapepsin add 25 µL of protein to a 1.5 mL centrifuge tube.

3.2.3.3. Thoroughly mix the crystallization buffer with the protein solution in a 1:1 ratio with a pipette tip. Place the tube in a revolver/rotator with high agitation at 20 °C.

3.2.3.4. Take regular (5, 10, 30, 60 min, 2, 5, 10, 24 h) 2.5 µL aliquots and view in a hemocytometer. Record the X_n and the X_s range.

3.2.3.5. If changes have occurred, repeat Steps 3.2.3.1. to 3.2.3.4. but increase or decrease the protein, precipitant, and/or seed concentration to restore any changes to the range of X_s and X_n

3.2.3.6. When/if the range of X_s and X_n are acceptable, proceed to Step 3.2.3.7.

3.2.3.7. Repeat Steps 3.2.2.1 to 3.2.2.5, increasing the total volume of the experiment gradually to 200 µL or larger, as required.

3.2.4. Final seeded-batch protocol

3.2.4.1. Prepare seed-stock.

3.2.4.1.1. Prepare crystallization buffer: 0.1 M Tris-HCl pH 7.0, 0.15 M $MgCl_2$, and 40% (w/v) PEG 6,000.

3.2.4.1.2. Using freshly defrosted 100 mg/mL endothiapepsin add 100 µL of protein to a 1.5 mL centrifuge tube.

3.2.4.1.3. Thoroughly mix the crystallization buffer with the protein solution in a 1:1 ratio with a pipette tip. Place the tube in a revolver/rotator with high agitation at 20 °C for 24 h to allow 50 µm crystals to grow.

3.2.4.1.4. Add 10-15 1 mm glass beads to the 50 µm crystal slurry.

3.2.4.1.5. Vortex the centrifuge tube containing the slurry and beads at 1000 rpm for 30 s.

3.2.4.1.6. Return the centrifuge tube to ice for 30 s.

3.2.4.1.7. Repeat Steps 3.2.4.1.5 and 3.2.4.1.6 10 more times.

3.2.4.1.8. This is now a 200 µL of a 1x seed-stock. Dilute the seed-stock 10x by the addition of 1.8 mL of crystallization buffer. Aliquot the 10x seed-stock in 50 µL batches and store at -20 °C.

3.2.4.2. Seeded-batch protocol.

3.2.4.2.1. Prepare crystallization buffer: 0.1 M Tris-HCl pH 7.0, 0.15 M $MgCl_2$, and 40% (w/v) PEG 6,000.

3.2.4.2.2. In a centrifuge tube, mix 100 µL of crystallization buffer with the 50 µL of freshly defrosted 10x seed-stock.

3.2.4.2.3. Using freshly defrosted 100 mg/mL endothiapepsin add 150 µL of protein to a 1.5 mL centrifuge tube.

3.2.4.2.4. Thoroughly mix the crystallization buffer/seed mixture with the endothiapepsin solution with a pipette tip and place the tube in a revolver/rotator with high agitation at 20 °C.

3.2.4.2.5. Monitor the crystallization by taking regular 2.5 µL aliquots and view the crystals in a hemocytometer. Record the X_n and the X_s range.

3.2.4.2.6. After approximately 80 min, when the crystals have reached an X_s of 15 µm, quench the reaction by the addition of 150 µL of 0.5 M Na Acetate pH 4.6, 0.5 M Tris-HCl pH 7.0, 0.075 M $MgCl_2$, and 20% (w/v) PEG 6,000 (a solution composed of endothiapepsin buffer and crystallization buffer, mixed 1:1).

3.2.4.2.7. Store the crystals at 20 °C.

3.3. Has the protocol produced an acceptable crystal size range and number for the intended experiment? Yes – **DONE** – No – return to Step 3.1. and attempt an alternative scaling option. For example, a different protein:precipitant ratio may be possible or adding seeds if this was not done previously. When these are all exhausted, it might be necessary to find a new condition at Step 1.

REPRESENTATIVE RESULTS:

Optimizing crystal morphology

Step 1, optimizing crystal morphology, has been included to remind reader of its importance. It may be possible to create perfect micro-crystals from poorly diffraction needle-balls, however, the authors would suggest that it is better to optimize the two separately. First, find conditions that give rise to well-diffracting, single crystal *via* vapor diffusion, and then convert these conditions into batch rather than trying to do combine the two steps together. Discovering highly nucleating conditions, at this stage, is not necessary; morphology and diffraction quality are the principal goals.

Before beginning the micro-crystallization of endothiapepsin, an analysis of deposited structure crystallization conditions from the PDB was conducted. Crystallization conditions and approximate protocols could be obtained for 47 of the 48 depositions of enthothiapepsin. These were broadly all based upon the first crystallization of endothiapepsin conducted by Moews and Bunn (1970)[46]. Given the similarities of these conditions and their 'classical' origin, a 96-well, vapor diffusion, sparse-matrix screen was performed to explore a wider variety of crystallization conditions. Endothiapepsin was concentrated to 70 mg/mL and a PACT sparse-matrix screen[47] was performed in a 96-well sitting-drop plate at 20 °C mixing 100 nL of protein with 100 nL of well solution. Every condition from this experiment after 36 h gave rise to crystals. However, an analysis of the crystal morphology indicated that some conditions might prove better for micro-crystallization optimization.

Figure 4A shows a drop from the PACT screen that was broadly representative of those observed in the majority of the plate. At first glance, it may be tempting to think that these crystals might be worth optimizing further for micro-crystallization. The crystals are large and there appears to be significant nucleation. However, the overall crystal morphology is not ideal. Firstly, the crystals are not observably singletons as it appears that multiple crystals are growing from single nucleation points. Secondly, the crystal size is highly asymmetric with growth principally occurring down a single axis. Such crystals are theoretically more likely to preferentially align when delivered to the X-ray beam. Both characteristics present problems during collection and processing of serial crystallographic data.

Figure 4B, however, shows endothiapepsin crystals grown in the presence of $MgCl_2$. This morphology was consistent across all conditions that contained $MgCl_2$ and therefore suggested that their morphology was due to $MgCl_2$. The $MgCl_2$ conditions produced single, more box-like crystals that represented a better target for the ultimate serial experiments.

There were four conditions within the PACT screen that contained $MgCl_2$. To better understand the influence of all the different components of these conditions on endothiapepsin crystallization, a random optimization was performed. A screen was created containing a random combination of the buffers and precipitants at a range of concentrations and pHs. The $MgCl_2$ concentration was also varied and then the resulting drops were arbitrarily graded from 1-5 (1 being poor) in terms of their visual crystal quality and precipitation level.

Figure 5A shows a heatmap of the results from a Pearson's correlation analysis between the precipitation level and crystal quality, and the screen variables (examples of the drops from this experiment are shown in **Figure 5B**, **C** and **D**). The results indicated that the pH of the solution was highly correlated to the level of precipitation with alkaline buffers resulting in more precipitation. $MgCl_2$ concentration was slightly correlated to the level of precipitation, as was the pH and precipitant concentration to crystal quality.

Based on these results, the decision was taken to take the crystals grown in 0.1 M Tris-HCl pH 7.0, 0.15 M $MgCl_2$, 20% (w/v) PEG 6,000 to the next step of the protocol – **Transitioning to batch**. The morphology of crystals was acceptable and an analysis of the X-ray diffraction and data quality metrics from these crystals suggested that there was no significant difference between the crystals grown in and out of the presence of Mg^{2+} (**Figure 9**).

Transitioning to batch

For many serial crystallography micro-crystallization optimizations, Step 2 will be the starting point. The protein of interest will have already been crystallized for cryo-crystallography and the crystallization protocol will now need transforming to create micro-crystal slurries. This protocol has only used 96-well vapor diffusion plates to perform the transformation to batch since vapor diffusion is the crystallization method used by 95% of PDB entries[26]. The protocol has avoided moving into microbatch[34,35,37] since this transition might still incur a similar optimization. This is not to say that this protocol can only be done in vapor diffusion plates. All of the steps presented, would also work in microbatch if this was the original crystallization method.

To assess the crystallization of endothiapepsin in the chosen condition, a morphogram – or a rough phase diagram – was created. The purpose of the morphogram experiment is threefold. Firstly, an analysis of the morphogram is of great utility when assessing scaling routes in Step 3 – **Scaling**. Secondly, the morphogram acts as an optimization tool, helping to discover vapor diffusion conditions that give rise to crystals *via* batch (i.e., rapidly appearing crystals (< 24 h). Thirdly, if crystals have not appeared rapidly, an analysis of the seeded drops can give the crystallographer an idea of the approximate location of the current condition on the phase diagram. For example, if the seeded conditions give crystals but the unseeded do not, those conditions are likely to be in the metastable region.

The morphogram experiment of endothiapepsin was performed based on the 0.1 M Tris-HCl pH 7.0, 0.15 M $MgCl_2$, 20% (w/v) PEG 6,000 condition. The protein and PEG concentrations were varied from 100 to 12.5 mg/mL and 5 to 40% (w/v), respectively. The drops were analyzed and results plotted using the worksheet provided (**Figure 6A**).

It was also already clear from the **Optimizing crystal morphology** stage that endothiapepsin crystal growth in this condition, and at these protein concentrations, would result in crystals grown in under 24 h. This indicated that crystallization was occurring *via* a batch rather than a vapor diffusion driven process. The crystal grown in these conditions were, therefore, suitable for scaling to larger volumes.

If crystals had not been visible in the unseeded-drops after 24 h, then it would have been likely that crystallization was still dependent upon a transition (**Figure 1B**) and, therefore, not batch. In this case, the results from the morphogram experiment are still of interest. They give an indication of the probable starting point for crystallization on the phase diagram and hence, how the subsequent optimization should proceed. Look at the seeded drops. The seeds will allow for crystal growth in the metastable region regardless of nucleation. For example, if crystals appear within 24 h in the seeded-drops but not the unseeded drops, this indicates part of the metastable region can be observed. If no crystals are observed in either the seeded or unseeded-drops, all wells remain undersaturated.

Scaling

Looking at the morphogram (**Figure 6A**), a number of observations could be made. The amount of nucleation did appear to be affected by both the protein and precipitant concentrations. There was also a very clear demarcation of drops that lead to protein precipitation, with drops either containing: nothing, crystals or precipitate (**Figure 6B**). The addition of seeds (**Figure 6D**) also greatly increased X_n when compared to the drops without seeds (**Figure 6C**). Taking all these results together, it was decided to attempt to scale both a batch and seeded-batch protocol at 30% (w/v) PEG 6,000 and 100 mg/mL endothiapepsin.

The initial test scaling was done in 24-well hanging drop plates. The drop volumes were gradually increased so that any changes in crystallization behavior could be observed (**Figure 7**). As can be seen, in both the unseeded and seeded drops crystal growth has occurred. All the unseeded

drops grew a range of crystal sizes, but predominantly large crystals (100-200 µm – longest dimension). The seeded drops, however, produced smaller crystals (5 – 50 µm – longest dimension). These initial tests suggested that seeds would be required to decrease X_s, but also, that this condition should be suitable for larger volumes.

When the volume was increased in 200 µL, the crystallization volume was continually agitated during crystal growth. The principal reason for this agitation was to ensure that the crystallization solution remains homogenous and that growing crystals do not settle on the bottom or sides of the tubes. Settling of crystals can lead to a heterogenous crystal population with both very large and small crystals. Agitating the crystallization solution can also promote nucleation[44, 45].

Unfortunately, the unseeded 30% (w/v) PEG 6,000 produced no crystals, so the PEG concentration was increased to 35% (w/v). This increase improved the crystallization markedly, with a final X_n and X_s range of 3.6 ± 1.2 x 10^6 crystals·mL^{-1} and 42 ± 4.1 µm, respectively (**Figure 8A** and **B** - black). Although a significant improvement and an acceptable crystal concentration, the final crystals were too large for the planned experiment, so further optimizations were undertaken. To reduce the size of the final crystals two avenues were explored (**Figure 1E**): decreasing the protein concentration to try and limit the final crystal growth (**Figure 8A** and **B** – hot pink), and increasing the PEG concentration to try and increase nucleation (**Figure 8A** and **B** – green).

The reduction of the protein concentration unfortunately also dramatically reduced the X_n, which ultimately produced even larger crystals. Increasing the PEG concentration to 40% yielded a final X_n and X_s range of 3.1 ± 0.7 x 10^6 crystals·mL^{-1} and 39 ± 2.3 µm, respectively. These were not significantly different to the 35%, but since the final crystal size was reduced, this condition was continued with the further optimizations.

To increase the X_n, seeds were added. This dramatically increased the X_n (1.1 ± 1.8 x 10^8 crystals·mL^{-1}) and lead to a smaller X_s (4.2 ± 4.0 µm) (**Figure 8A** and **B** – dashed purple). These crystals, although very suitable for some experiments, were deemed too small so the concentration of the added seeds was changed.

This tuning of the added seed stock, however, proved difficult to reliably repeat; therefore, quenching was attempted. After the addition of a seed stock, the crystal size was monitored and once a suitable crystal size was achieved (approximately 10 – 20 µm), the batch crystallization was quenched (**Figure 8C** and **D**). Quenching was proposed, with regard to micro-crystallization, in Kupitz et al. (2014)[25]. Although perhaps not an ideal method, as protein solution will ultimately be wasted[26], the technique was very useful in this situation as crystal growth was difficult to control. The idea behind quenching is to rapidly return the crystallization mixture to a point just above the solubility line (**Figure 1F**). Once the solution has returned to the solubility line, the solution has returned to a stable saturated solution and no further crystal growth will occur.

Attempting to quench a crystallization reaction is not without risk. If too great a quenching solution is added, the protein in solution might be diluted so much that the solubility line is

765 passed. In this case, the solution will become undersaturated and the crystals might start
766 dissolving. To prevent this, it is possible to estimate the amount of required quenching solution
767 based on the morphogram results. At the point of quenching, take the concentration of the
768 protein solution. By comparing the protein concentration at the solubility line and the protein
769 concentration in solution, an estimate the required dilution can be made.
770
771 The quenched version of the 40% (w/v) PEG 6,000, 10 x diluted seed experiment gave a final
772 crystal concentration and size range of 2.6 ± 3.1 x 10^6 crystals·mL^{-1} and 15 ± 3.9 µm, respectively.
773
774 Throughout the entire process, test X-ray data collections of the endothiapepsin crystals were
775 collected at the Swiss Light Source PXII beamline using a 10 x 30 µm focus, an energy of 12.4 keV
776 attenuated by 80%, and under cryo-conditions. The data were processed using `dials` and
777 **Figure 9** shows a comparison of $CC_{½}$. No dramatic change in in $CC_{½}$ was observed over the course
778 of the optimization.
779

FIGURE AND TABLE LEGENDS:

Figure 1: An overview of transitionary and batch crystallization, and scaling methods mapped onto a phase diagram. A. The zones and limits of the archetypical protein crystallization phase diagram. The precipitant and protein concentrations are plotted on the *x* and *y* axes, respectively with the pure water point at the origin. The purple line indicates the protein supersaturation boundary, and the metastable, nucleation, and precipitation zones are shown in blue, green and pink, respectively. **B.** An example of the nucleation zone penetration limits of a transitionary phase crystallization method, such as vapor diffusion. In this theoretical experiment, the drop precipitant and protein concentrations begin just below the solubility line – not yet supersaturated. While the drop equilibrates, the drop component concentrations increase such that the drop becomes supersaturated, and continues to move – or transition – into the nucleation zone. Upon crystal nucleation, the protein concentration in solution begins to drop. The concentration continues to fall as crystals grow until finally stopping at the solubility line. The blue dotted line marks a theoretical limit of the transition into the nucleation zone. As soon as nucleation begins, the protein concentration will drop, preventing further penetration. **C.** Example batch and seeded-batch crystallization trajectories. In batch, the mixing of the protein and precipitant must create a solution within the nucleation zone so that crystal growth can occur. In seeded-batch, it is not strictly necessary to be in the nucleation zone due to the addition of micro-seeds, so locations in the metastable region can also be explored. **D.** A hypothetical optimization of the crystallization experiment shown in **B** from vapor diffusion to batch. The original vapor diffusion starting point has transitioned, *via* the resultant optimization vector, to the new start position; inside the nucleation zone. The resultant vector is the product of two optimizations: an increase in both protein and precipitant concentrations. **E.** Example optimizations when scaling batch conditions to tailor the final X_n and X_s. **F.** Quenching the crystallization experiment by the addition of crystallization buffer. It is essential that the quenching does not take the protein concentration out of the metastable region and, therefore, below the point of protein supersaturation. Otherwise, crystals will start to dissolve back into solution. **B.** and **C.** have been adapted from Beale et al. (2019)[26] with the permission of the authors.

Figure 2: Increasing X_n and decreasing X_s. The idealized relationship between the number of crystals produced from a crystallization experiment and their mean longest dimension. To create this graph, the crystallization of a hypothetical 10 kDa model protein was used. The protein crystallized at a concentration of 10 mg/mL and yielded $P2_12_12_1$ crystals with dimensions of 49x50x51 Å. Every nucleation event was assumed to yield a crystal. Crystal growth was assumed to be homogeneous from every face.

Figure 3: A flow-chart showing the steps to optimize a crystal grown in a small-volume (<500 nL), vapor diffusion experiment into a large-volume (> 100 µL) batch experiment. Crystal optimization is divided into three stages: 1. Optimizing crystal morphology. 2. Transitioning to batch. 3. Scaling. In Stage 1 it is important to identify suitable crystals for micro-crystallization. Some proteins only present in a single crystal morphology regardless of the crystallization condition. However, it is worth looking for conditions that give rise to single, cube-like crystals,

or as close to these as humanly possible. Single, cube-like crystals, hypothetically and anecdotally, will generally give rise to better outcomes from serial crystallography experiments. Once a crystal morphology has been selected and the diffraction confirmed, it is then necessary to move the crystallization experiment from vapor diffusion to batch (Stage 2). Here, crystals should be optimized by their nucleation time. The goal is to find conditions that yield rapidly appearing crystals (> 24 h) as these conditions are likely to hit the nucleation zone immediately, and are therefore batch. Once a condition in the nucleation zone has been found, a morphogram can be created. The morphogram allows for the majority of the nucleation zone to mapped and potential scaling routes identified for Stage 3. The volume of an identified batch condition can then be either gradually or rapidly scaled in size to yield a final volume of >100 µL.

Figure 4: An analysis of endothiapepsin crystallization conditions from a PACT sparse-matrix screen. A. and **B.** are photos after 24 h of wells A4 and C10, respectively from the PACT screen. The crystallization buffer components are highlighted on the figure. The SPG buffer is succinic acid, sodium dihydrogen phosphate, and glycine mixed in a 2:7:7 molar ratio.

Figure 5: An analysis of the endothiapepsin crystallization optimization from the PACT $MgCl_2$ conditions. A. A heat map of the results from a Pearson's correlation analysis between buffer pH, $MgCl_2$ concentration, and precipitant concentration and the precipitation level and crystal quality. The precipitation level and crystal quality were both assessed arbitrarily on a scale of 1-5 (with 1 being poor) after 24 h. **B. C.** and **D.** show examples of the crystallization and precipitation in three different drops. The crystallization condition and assessments of the precipitation level and crystal quality are also shown.

Figure 6: An endothiapepsin morphogram when crystallized in 0.1 M Tris-HCl pH 7.0, 0.15 M $MgCl_2$ and PEG 6,000. A. A morphogram created from the "phase-diagram-generator" spreadsheet provided. The relative number of crystals in each drop is denoted by the size of the circles, and the results from drop 1 (protein and precipitant) and drop 2 (protein, precipitant and seeds) are highlighted in green and blue, respectively. The values of the protein and precipitant concentrations, on the *x* and *y* axis, respectively, denote pre-mixed values of each rather than final volumes. Based on the results, black lines and a purple line have been drawn to show the boundaries of the nucleation zone and metastable zone, respectively. **B. C.** and **D.** show some example results from the experiment. The red and blue dots marked on **A.** indicate the locations of **B.**, and **C.** and **D.**, respectively.

Figure 7: Initial scaling trials of endothiapepsin in 24-well hanging drop plates. The same protein and precipitant concentrations were used for all trails: 100 mg/mL endothiapepsin in 0.1 M Na Acetate pH 4.6 and 0.1 M Tris-HCl pH 7.0, 0.15 M $MgCl_2$, and 30% (w/v) PEG 6,000, respectively. All of the displayed images were taken after 24 h and the final drop volumes are labelled on each image. The left panel (**A**, **D**, and **G**) are a 1:1 mix of protein and precipitant, the middle panel (**B**, **E**, and **H**) are a 1:2:3 mix of seeds, precipitant, and protein and the right panel (**C**, **F**, and **I**) are magnified images of the middle panel.

Figure 8: Analysis of the endothiapepsin micro-crystallization in 200-300 µL volumes. **A.** and **C.** show how X_n changed over the experiment time. **B.** and **D.** show how X_s (longest dimension) changed over time. The results of the experiments have been separated for clarity. The red dotted line on **C.** and **D.** show the point at which quenching was performed.

Figure 9: $CC_{½}$ results and images of crystals obtained at each stage of the micro-crystallization process to assess diffraction quality. A. $CC_{½}$ plotted against resolution from data collected from crystals grown: with and without Mg – part of the Stage 1 optimization, in a 200 nL volume, a 10 µL volume and the final 300 µL volume. **B. C.** and **D.** show the crystals from the 200 nL, 10 µL and 300 µL volume, respectively.

Table 1: An example of the sample requirements for a hypothetical optical pump-probe experiment performed using fixed-targets. The protein used in this theoretical experiment was endothiapepsin. The fixed-target parameters were based on experiments reported in Ebrahim et al. (2019)[48] and Davy et al. (2019)[49]. The sample information came from the protocol reported in this video article and the experimental variables were conservative estimates based on lived experience. The following sample requirements were subsequently calculated given the previous assumptions.

DISCUSSION:

The method presented shows how to optimize the crystallization of endothiapepsin from large crystals (≅ 100 µm longest dimension), grown in sparse-matrix 96-well screens, to micro-crystals, grown in centrifuge tubes (300 µL volume) *via* batch. The idea behind the protocol is that the steps taken to optimize endothiapepsin could also be used for other proteins. Ultimately, answering the problem of creating large volumes (>100 µL) of micro-crystals (10-20 µm) for serial crystallography experiments at XFELs and synchrotrons.

The protocol divides the task of large volume micro-crystallization into three steps: 1. Optimizing crystal morphology, 2. Transitioning to batch, and 3. Scaling. In Step 1, the range of crystal forms that a protein can create should be explored in vapor diffusion plates. Conditions that give rise to single, box-like crystals that diffract to the required resolution should be the goal. In Step 2, selected conditions can then be transformed from vapor diffusion into batch. Here, the optimization criterion is crystal growth time to find conditions that give rise to protein crystals within 24 h. A morphogram can also be plotted giving the experimenter an idea of the location of the solubility line and nucleation zone boundaries. This morphogram is of great utility in Step 3, scaling. The morphogram will give an indication of the whether nucleation alone can increase X_n and drive down X_s. As the volume of the experiment is increased, X_n and X_s can be continually assessed as the key criteria of scaling success.

In the case of Endothiapepsin, Step 1 unearthed what potentially was a previously unknown crystal morphology for endothiapepsin. This morphology had the same spacegroup as those previously reported but, importantly for serial crystallography, a more box-like shape. Single crystals also seemed to grow from single nucleation points, unlike the fans created from other conditions (**Figure 4**). For the selected condition, Step 2 was already partially satisfied as crystal

growth occurred in < 24 h. The morphogram indicated that both a straight or seeded-batch protocol might be successful when scaling in Step 3. Initial scaling in straight batch, created a condition that produced crystals with an X_n and X_s range of 3.6 ± 1.2 x 10^6 crystals·mL^{-1} and 42 ± 4.1 µm, respectively. These crystals, although acceptable for some serial crystallography experiments, were deemed too large. So additional optimizations were performed. The final protocol produced crystals with a concentration and size range of 3.1 x 10^6 crystals·mL^{-1} and 15 ± 3.9 µm, respectively. This was more than ideal for the planned experiments.

The method focuses on the transformation of 'soluble' protein crystals grown in vapor diffusion plates to batch. The reason for this focus is that the vast majority of soluble protein crystals are grown *via* vapor diffusion[26]. However, the concepts presented could also be applied to soluble protein crystals grown using other methods, such as micro-batch. The concepts may also be applicable to membrane protein crystals grown in LCP; as this too is a batch crystallization process.

A key aspect of the protocol is the process of transforming the conditions of crystals grown in vapor diffusion plates such that they can be grown in batch. For this transformation, the method uses the criterion proposed by Beale et al. (2019)[26]. Crystals grown *via* a batch process, even in vapor diffusion plates, will form rapidly (< 24 h). This criterion is an approximation based on the speed of vapor diffusion drop equilibration and is most true for PEG-based precipitant conditions. However, crystallization conditions will contain a wide variety of compounds that will influence the equilibration time. The equilibration of salt-based crystallization conditions, *e.g.* highly concentrated ammonium chloride, can happen in 1-2 days. Therefore, the 24 h criterion may not be true for salt-based conditions. Salt based conditions also can have more complex phase diagrams[26, 30] that may not conform to the archetype presented in this protocol. A reduction in the time criterion for salt-based conditions to 12 or 6 h may be necessary if scaling into larger volumes proves impossible.

Another limitation of this method is its apparent complexity. The protocol that was followed to optimize the micro-crystallization of endothiapepsin actually changed the original condition from the sparse-matrix screen relatively little. The first hit observed in the PACT screen was 0.1 HEPES pH 7.0, 0.2 M $MgCl_2$, and 20% (w/v) PEG 6,000. The final scaled crystallization buffer was 0.1 Tris-HCl pH 7.0, 0.15 M $MgCl_2$, and 40% (w/v) PEG 6,000. It is also very possible that the change in buffer from HEPES to Tris-HCl, and $MgCl_2$ concentration, contributed little to the success of the process. Leaving the increase in PEG 6,000 concentration the sole optimization, and one that could have been achieved quite simply.

This assessment, however, is also too simplistic. It not only discounts the problems encountered during scaling (i.e., the use of seeds and quenching, but also the fact that just because this protein proved straightforward, there is no guarantee the next will also prove to be. The steps advised in the protocol, were devised because optimizing the scaling of protein crystallization volumes can be very protein expensive. Over the seven endothiapepsin scaling trials that are shown, 100 mg of protein were consumed. Admittedly some of these steps were performed to show their consequences in the light of this protocol. Even so, 100 mg of a protein, plus potentially a further

50 mg for protein consumed during an experiment (**Table 1**), can be a significant investment in either time or money.

Fortunately, it is not clear that this mass of required sample is ubiquitous across all proteins. Endothiapepsin was highly soluble, and therefore required a large protein concentration to reach supersaturation. In others (currently under optimization), supersaturation can be reached at 10 or even 5 mg/mL. Such variables are protein specific and need to be embraced when they appear.

Other limitations of the method include its reliance on complex equipment such as liquid handling robots for screen and plate creation, and imagers to automatically image plates when required. Alternative routines have been offered to limit the need of some of these pieces of equipment, but the protocol will be more time consuming to follow without them. The protocol also suggests testing the diffraction of optimized crystals. For crystallographers without regular access to a synchrotron, these tests could prove challenging. Controls at every step may not be necessary, but these tests are strongly recommended once a hit has been identified, and pre-and post-scaling. Non-diffracting crystals at an XFEL are, unfortunately, not an uncommon occurrence. Given this, it is better to err on the side of caution regarding assumptions about crystal diffraction.

Ultimately, this protocol and results presented here will offer a guide, ideas, and an example to those struggling with producing samples for serial crystallography experiments. Hopefully, as serial crystallography is further developed, the sample demands of the technique will be reduced such that the need for protocols like this will be reduced. However, even in this event, the strategies presented here will still be useful to those wishing to explore the crystallization space of their protein.

ACKNOWLEDGMENTS:
This project has received funding from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No 701647.

DISCLOSURES:
The authors have no conflicts of interest to disclose.

REFERENCES:
1. DePonte, D.P. et al. Gas dynamic virtual nozzle for generation of microscopic droplet streams. *Journal of Physics D: Applied Physics*. **41** (19), 195505 (2008).
2. Hunter, M.S. et al. Fixed-target protein serial microcrystallography with an x-ray free electron laser. *Scientific Reports.* **4** (1), 6026 (2014).
3. Weierstall, U. et al. Lipidic cubic phase injector facilitates membrane protein serial femtosecond crystallography. *Nature Communications*. **5** (1), 1–6, (2014).
4. Roessler, C.G.G. et al. Acoustic Injectors for Drop-On-Demand Serial Femtosecond Crystallography. *Structure*. **24** (4), 631–640 (2016).
5. Sherrell, D.A. et al. A modular and compact portable mini-endstation for high-precision, high-speed fixed target serial crystallography at FEL and synchrotron sources. *Journal of*

Synchrotron Radiation. **22** (6), 1372–1378 (2015).
6. Roedig, P. et al. A micro-patterned silicon chip as sample holder for macromolecular crystallography experiments with minimal background scattering. *Scientific Reports*. **5** (1), 1–11 (2015).
7. Botha, S. et al. Room-temperature serial crystallography at synchrotron X-ray sources using slowly flowing free-standing high-viscosity microstreams. *Acta Crystallographica Section D Biological Crystallography*. **71** (2), 387–397 (2015).
8. Weinert, T. et al. Serial millisecond crystallography for routine room-temperature structure determination at synchrotrons. *Nature Communications*. **8** (1), 542 (2017).
9. Tenboer, J. et al. Time-resolved serial crystallography captures high-resolution intermediates of photoactive yellow protein. *Science*. **346** (6214), 1242–1246 (2014).
10. Nango, E. et al. A three-dimensionalmovie of structural changes in bacteriorhodopsin. *Science*. **354** (6319), 1552–1557 (2016).
11. Suga, M. et al. Light-induced structural changes and the site of O=O bond formation in PSII caught by XFEL. *Nature*. **543** (7643), 131–135 (2017).
12. Mehrabi, P. et al. Liquid application method for time-resolved analyses by serial synchrotron crystallography. *Nature Methods*. **16** (10), 979–982 (2019).
13. Halle, B. Biomolecular cryocrystallography: structural changes during flash-cooling. *Proceedings of the National Academy of Sciences of the United States of America*. **101** (14), 4793–4798, doi: 10.1073/pnas.0308315101 (2004).
14. Fraser, J.S. et al. Hidden alternative structures of proline isomerase essential for catalysis. *Nature*. **462** (7273), 669–673 (2009).
15. Fenwick, R.B., van den Bedem, H., Fraser, J.S., Wright, P.E. Integrated description of protein dynamics from room-temperature X-ray crystallography and NMR. *Proceedings of the National Academy of Sciences of the United States of America*. **111** (4), E445-54 (2014).
16. Keedy, D.A. et al. Mapping the conformational landscape of a dynamic enzyme by multitemperature and XFEL crystallography. *Elife*. **4** (September2015) (2015).
17. Thomaston, J.L. et al. XFEL structures of the influenza M2 proton channel: Room temperature water networks and insights into proton conduction. *Proceedings of the National Academy of Sciences of the United States of America.* **114** (51), 13357–13362 (2017).
18. Haas, D.J., Rossmann, M.G. Crystallographic studies on lactate dehydrogenase at –75 °C. *Acta Crystallographica Section B Structural Crystallography and Crystal Chemistry*. **26** (7), 998–1004 (1970).
19. Hope, H. Cryocrystallography of biological macromolecules: a generally applicable method. *Acta Crystallographica Section B Structural Science*. **44** (1), 22–26 (1988).
20. Wu, W. et al. Batch crystallization of rhodopsin for structural dynamics using an X-ray free-electron laser. *Acta Crystallographica Section:F Structural Biology Communications*. **71** (7), 856–860 (2015).
21. Ishchenko, A., Cherezov, V., Liu, W. Preparation and delivery of protein microcrystals in lipidic cubic phase for serial femtosecond crystallography. *Journal of Visualized Experiments*. **2016** (115), e54463 (2016).
22. Andersson, R. et al. Well-based crystallization of lipidic cubic phase microcrystals for serial X-ray crystallography experiments. *Acta Crystallographica Section D: Structural Biology*. **75** (10), 937–946 (2019).

23. Luft, J.R. et al. The detection and subsequent volume optimization of biological nanocrystals. *Structural Dynamics*. **2** (4), 041710 (2015).
24. Lee, D.B. et al. Supersaturation-controlled microcrystallization and visualization analysis for serial femtosecond crystallography. *Scientific Reports*. **8** (1), 1–10 (2018).
25. Kupitz, C. et al. Microcrystallization techniques for serial femtosecond crystallography using photosystem II from Thermosynechococcus elongatus as a model system. *Philosophical transactions of the Royal Society of London. Series B, Biological Sciences*. **369** (1647), 20130316 (2014).
26. Beale, J.H. et al. Successful sample preparation for serial crystallography experiments. *Journal of Applied Crystallography.* **52** (Pt 6), 1385–1396 (2019).
27. Rayment, I. Small-scale batch crystallization of proteins revisited: An underutilized way to grow large protein crystals. *Structure*. **10** (2), 147–151 (2002).
28. García-Ruiz, J.M. Nucleation of protein crystals. *Journal of Structural Biology*. **142** (1), 22–31 (2003).
29. McPherson, A., Kuznetsov, Y.G. Mechanisms, kinetics, impurities and defects: Consequences in macromolecular crystallization. *Acta Crystallographica Section F:Structural Biology Communications*. **70** (4), 384–403 (2014).
30. Rupp, B. Origin and use of crystallization phase diagrams. *Acta crystallographica. Section F, Structural biology communications*. **71** (Pt 3), 247–60 (2015).
31. Luft, J.R., DeTitta, G.T. A method to produce microseed stock for use in the crystallization of biological macromolecules. *Acta Crystallographica Section D Biological Crystallography*. **55** (5), 988–993 (1999).
32. Ireton, G.C., Stoddard, B.L. Microseed matrix screening to improve crystals of yeast cytosine deaminase. *Acta crystallographica. Section D, Biological crystallography*. **60** (3), 601–605 (2004).
33. Forsythe, E.L., Maxwell, D.L., Pusey, M. Vapor diffusion, nucleation rates and the reservoir to crystallization volume ratio. *Acta Crystallographica Section D Biological Crystallography*. **58** (10), 1601–1605 (2002).
34. Chayen, N.E., Shaw Stewart, P.D., Maeder, D.L., Blow, D.M., IUCr An automated system for micro-batch protein crystallization and screening. *Journal of Applied Crystallography*. **23** (4), 297–302 (1990).
35. Chayen, N.E., Shaw Stewart, P.D., Blow, D.M. Microbatch crystallization under oil — a new technique allowing many small-volume crystallization trials. *Journal of Crystal Growth*. **122** (1–4), 176–180 (1992).
36. Chayen, N.E. Comparative studies of protein crystallization by vapour-diffusion and microbatch techniques. *Acta Crystallographica Section D: Biological Crystallography*. **54** (1), 8–15 (1998).
37. D'Arcy, A., Mac Sweeney, A., Stihle, M., Haber, A. The advantages of using a modified microbatch method for rapid screening of protein crystallization conditions. *Acta Crystallographica - Section D Biological Crystallography*. **59** (2), 396–399 (2003).
38. Darmanin, C. et al. Protein crystal screening and characterization for serial femtosecond nanocrystallography. *Scientific Reports*. **6**, 25345 (2016).
39. Gati, C. et al. Atomic structure of granulin determined from native nanocrystalline granulovirus using an X-ray free-electron laser. *Proceedings of the National Academy of Sciences*

of the United States of America. **114** (9), 2247–2252 (2017).
40. Cheng, R.K.Y. Towards an optimal sample delivery method for serial crystallography at XFEL. *Crystals*. **10** (3), 215 (2020).
41. Schmidt, M. Mix and Inject: Reaction Initiation by Diffusion for Time-Resolved Macromolecular Crystallography. *Advances in Condensed Matter Physics*. **2013**, 1–10 (2013).
42. Oberthuer, D. et al. Double-flow focused liquid injector for efficient serial femtosecond crystallography. *Scientific Reports*. **7** (1), 44628 (2017).
43. McPherson, A., Gavira, J.A. Introduction to protein crystallization. *Acta crystallographica. Section F, Structural biology communications*. **70** (Pt 1), 2–20 (2014).
44. Yaoi, M. et al. Effect of stirring method on protein crystallization. *Japanese Journal of Applied Physics, Part 2: Letters*. **43** (10 A), L1318 (2004).
45. Castro, F., Ferreira, A., Teixeira, J.J., Rocha, F. Influence of Mixing Intensity on Lysozyme Crystallization in a Meso Oscillatory Flow Reactor. *Crystal Growth & Design*. **18** (10), 5940-5946 (2018).
46. Moews, P.C., Bunn, C.W. An X-ray crystallographic study of the rennin-like enzyme of Endothia parasitica. *Journal of Molecular Biology*. **54** (2), 395–397 (1970).
47. Newman, J. et al. Towards rationalization of crystallization screening for small- to medium-sized academic laboratories: the PACT/JCSG+ strategy. *Acta crystallographica. Section D, Biological crystallography*. **61** (Pt 10), 1426–1431 (2005).
48. Ebrahim, A. et al. Resolving polymorphs and radiation-driven effects in microcrystals using fixed-target serial synchrotron crystallography. *Acta Crystallographica Section D Structural Biology*. **75** (Pt 2), 151–159 (2019).
49. Davy, B. et al. Reducing sample consumption for serial crystallography using acoustic drop ejection. *Journal of Synchrotron Radiation*. **26** (5) (2019).

Figure 1

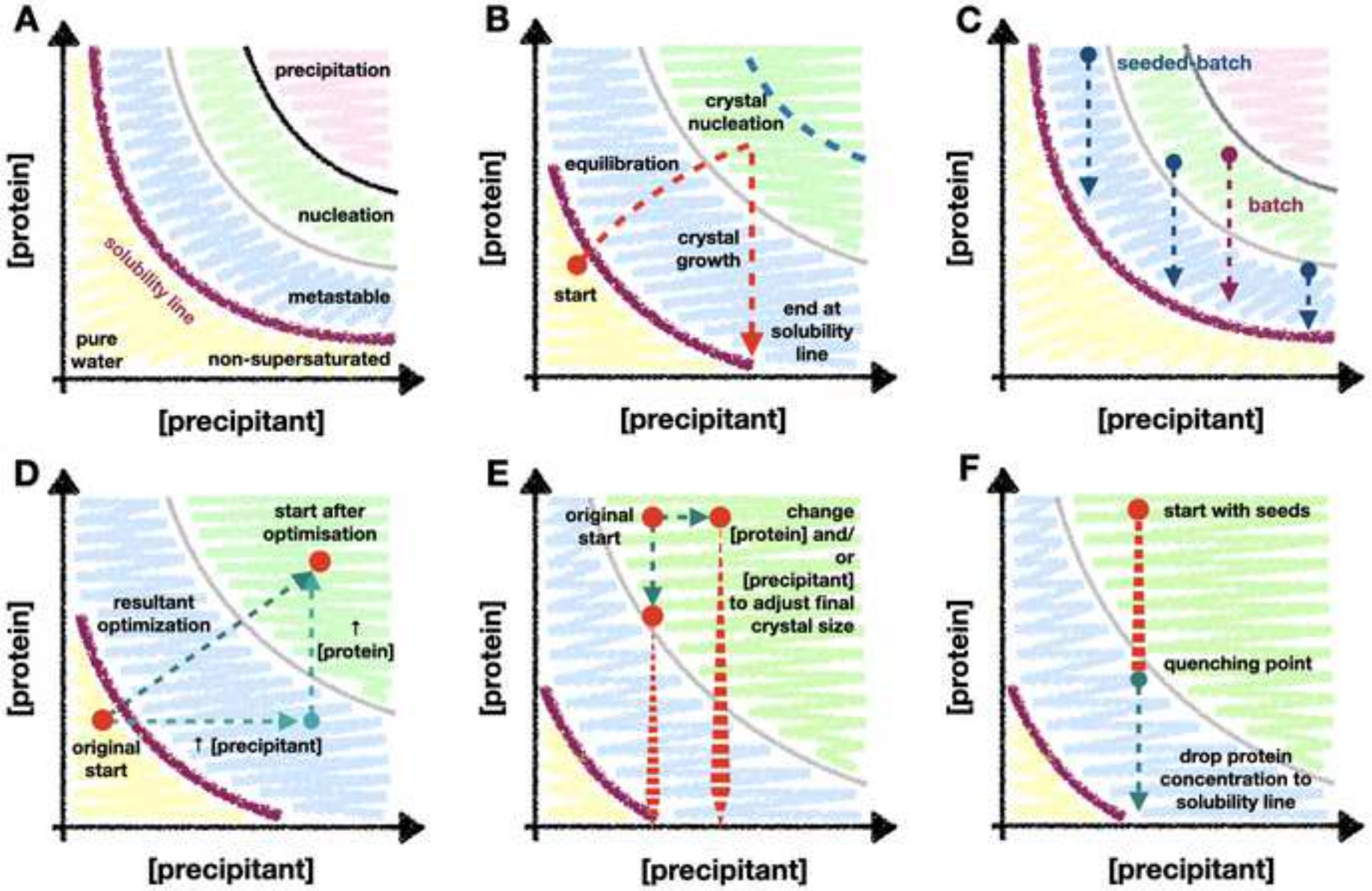


Figure 2

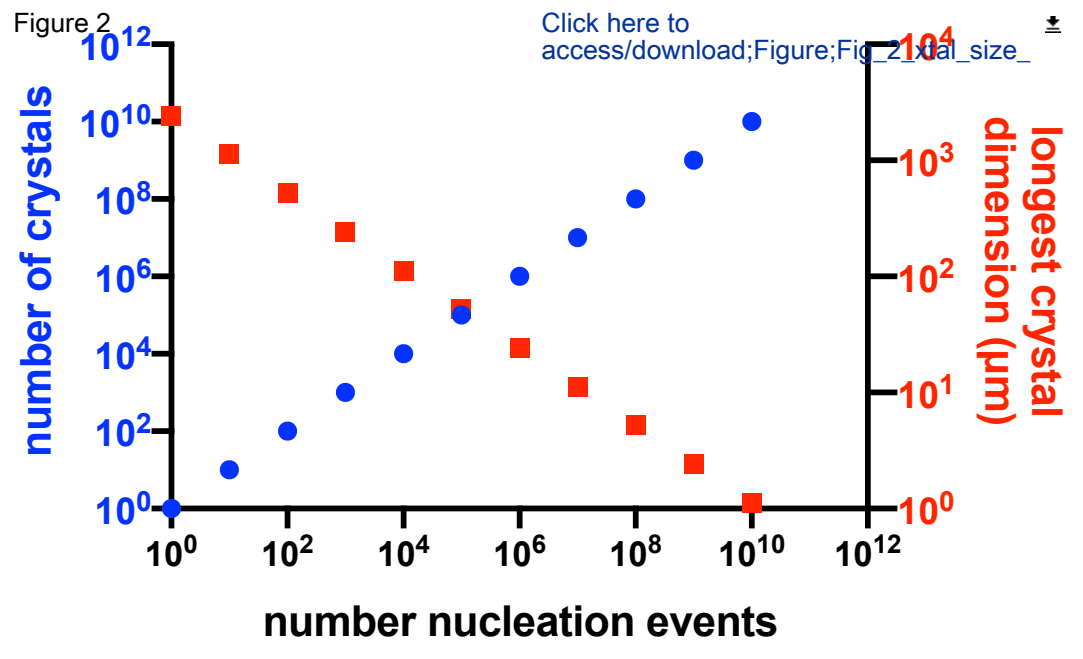


Figure 3

Click here to access/download;Figure;Fig_3_flow-chart-v5.001.jpeg

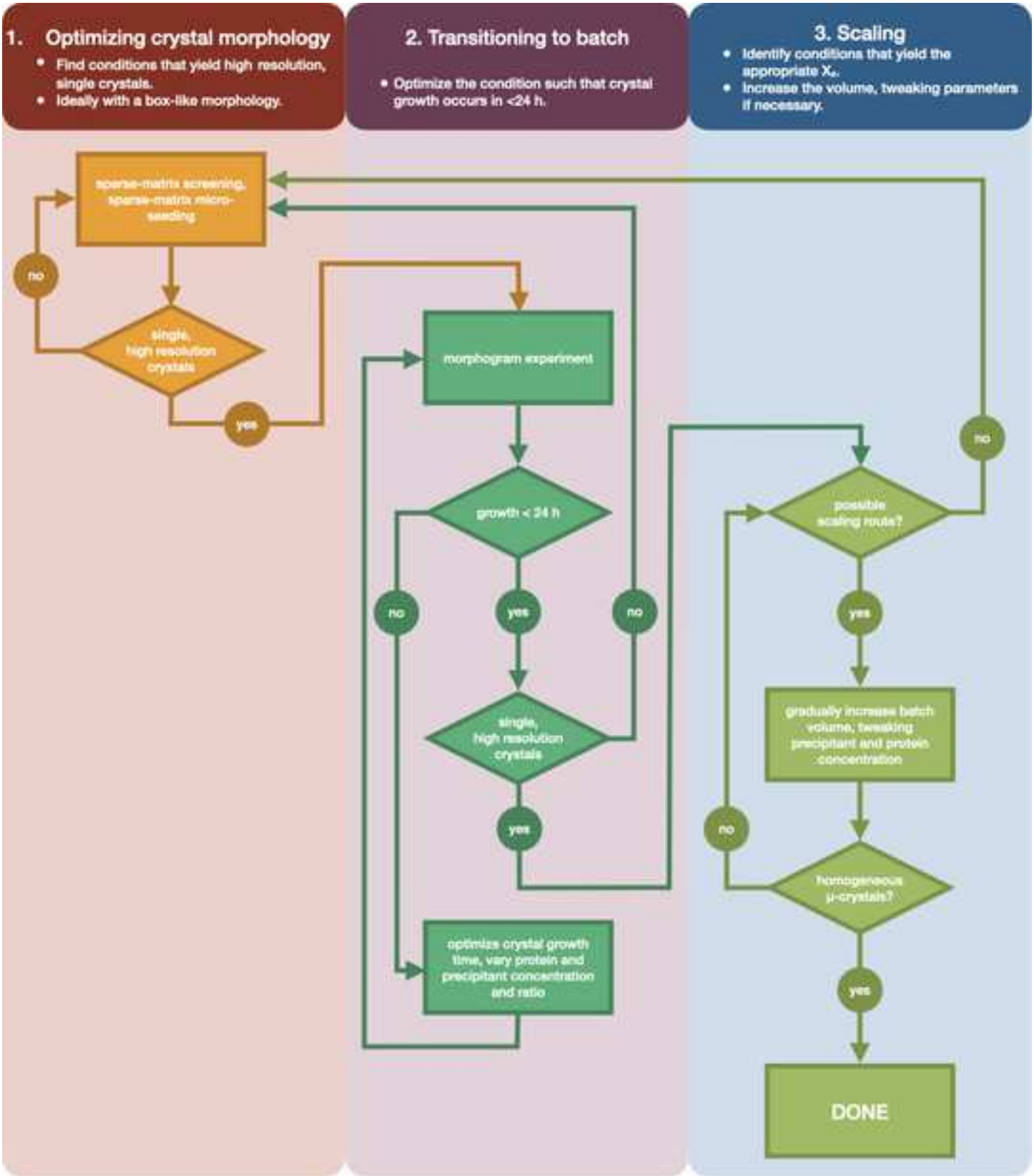


Figure 4

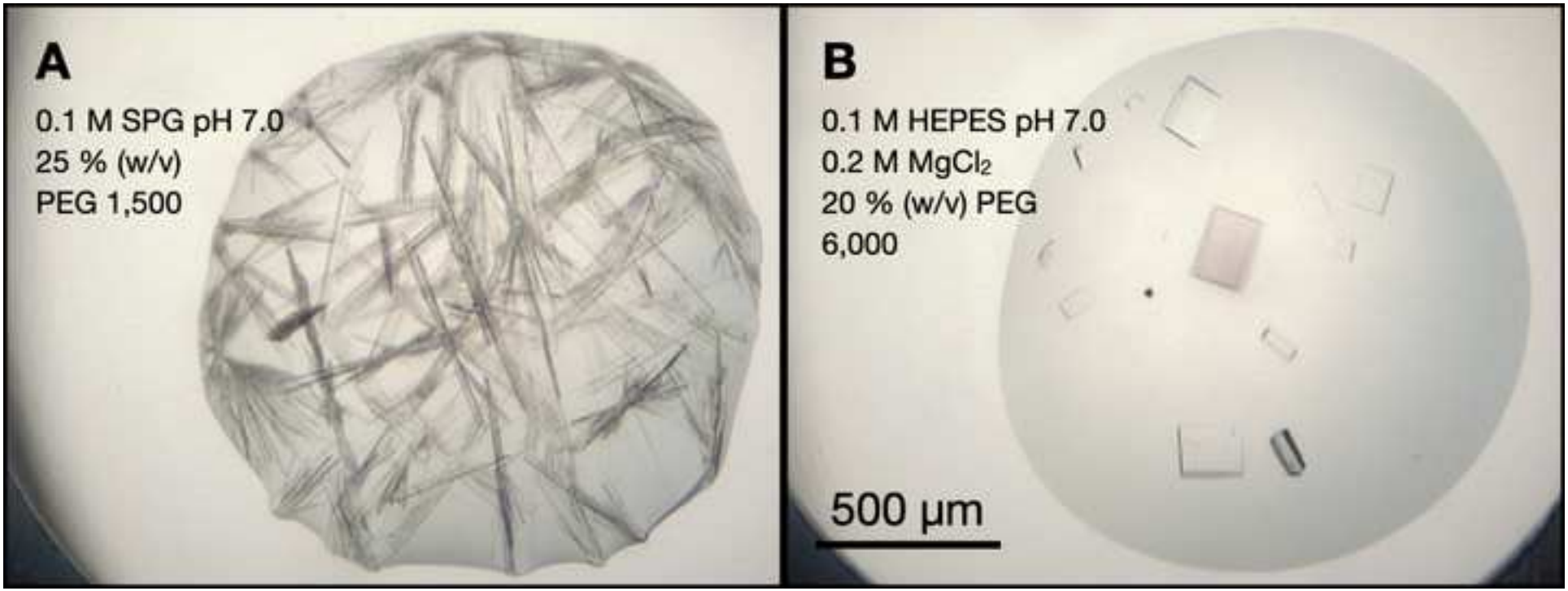


Figure 5

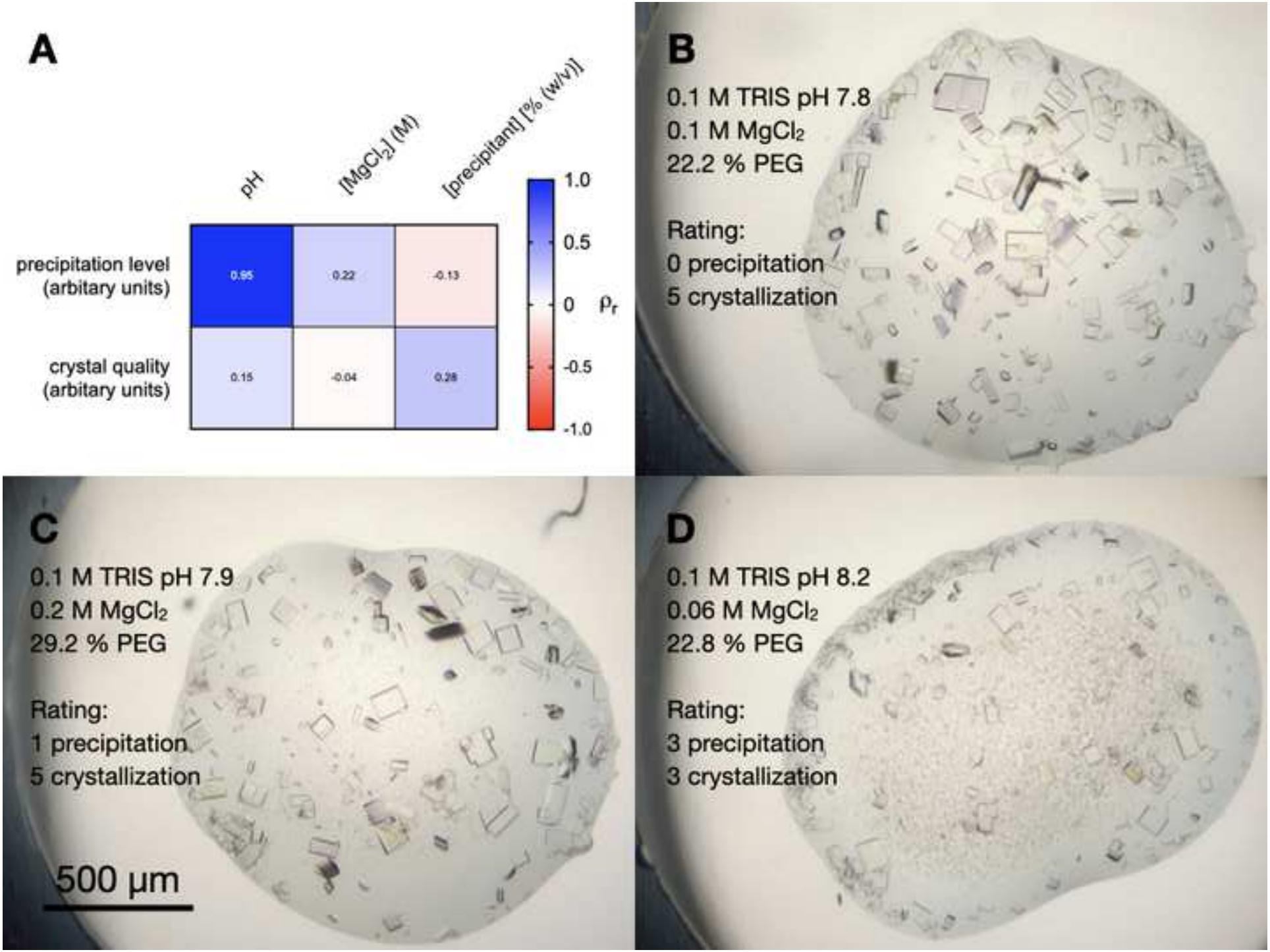


Figure 6

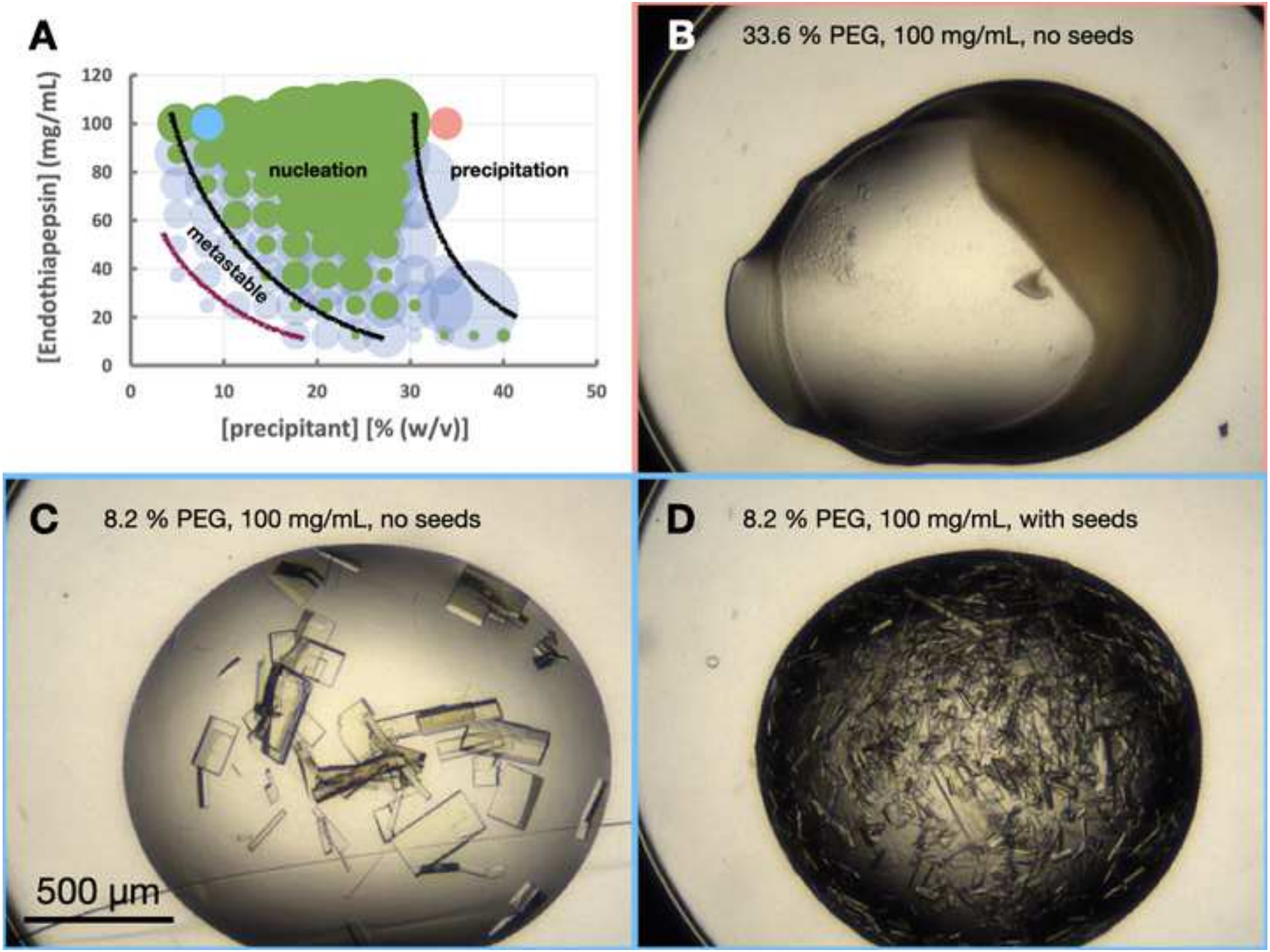


Figure 7

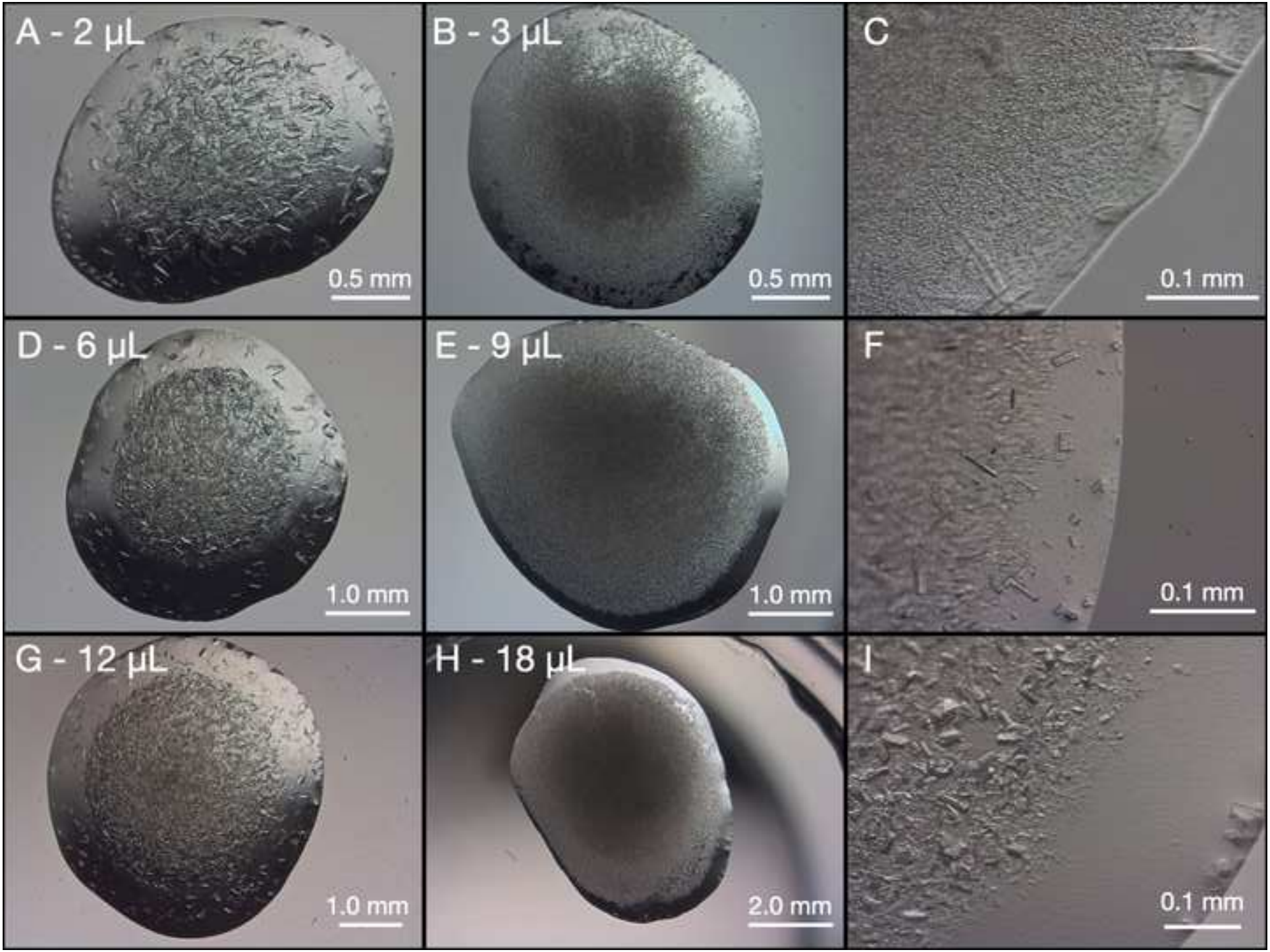


Figure 8

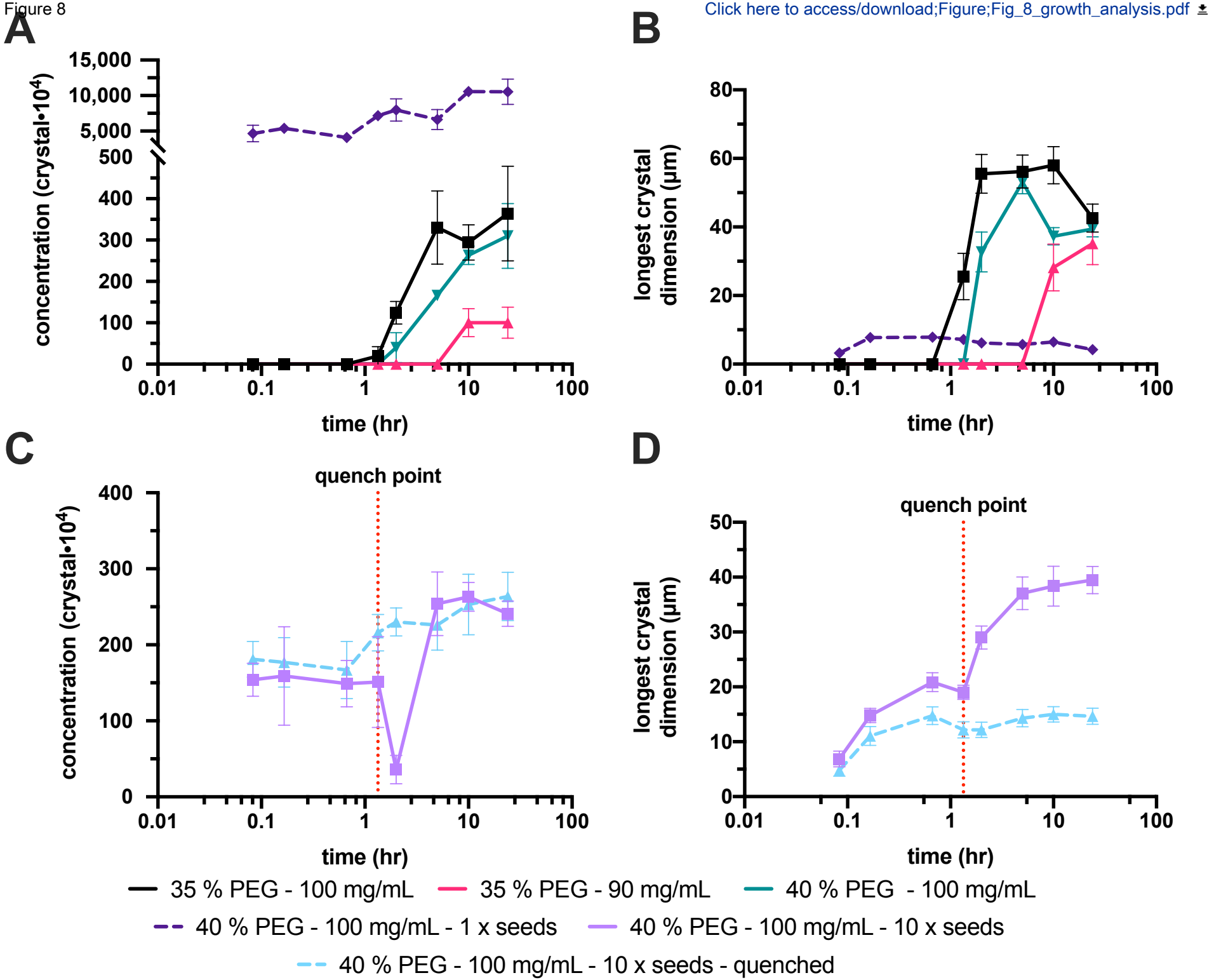


Figure 9

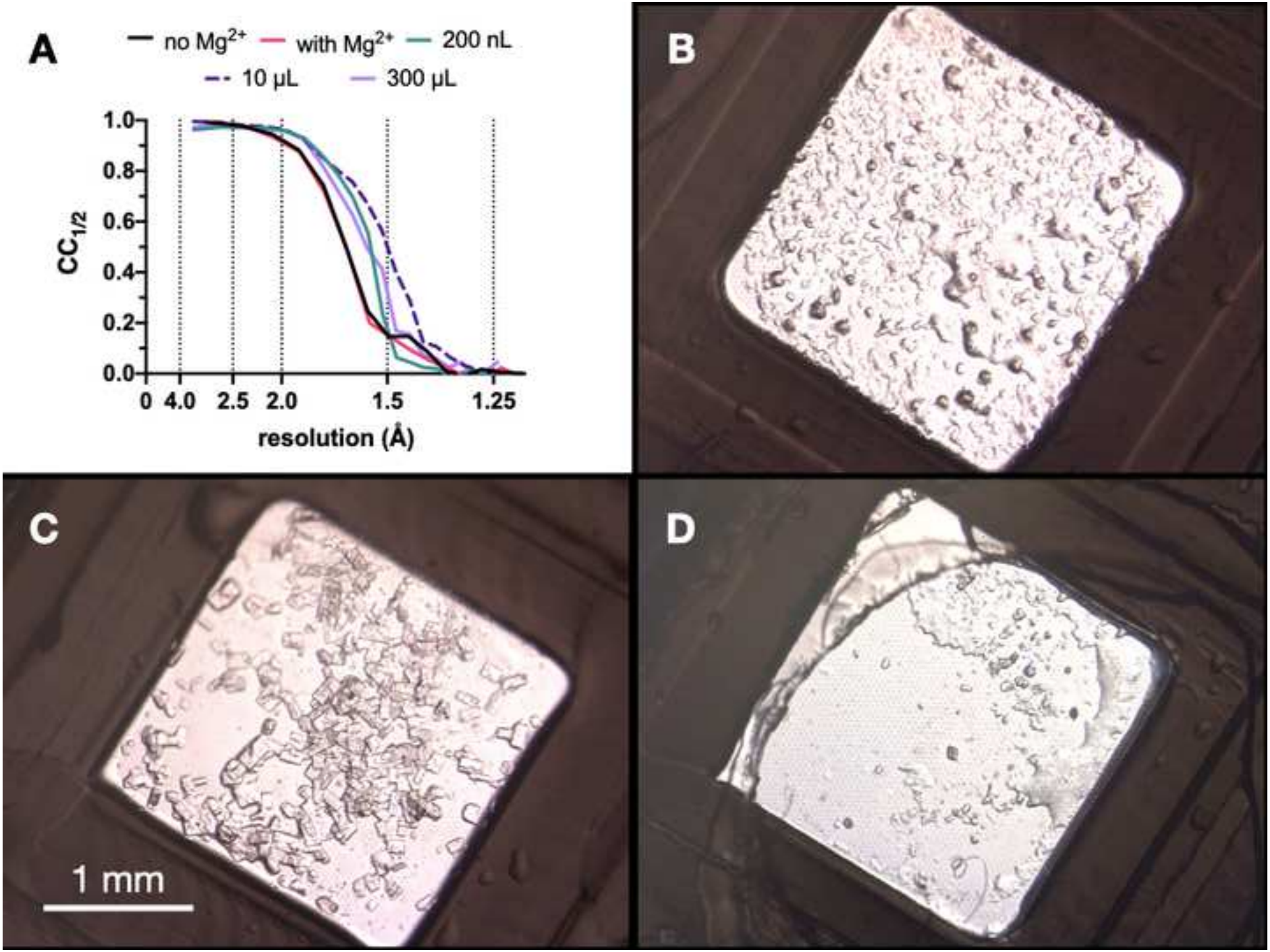


Table 1

Click here to access/download;Table;Tab_1_example-serial-sample-consumption.xlsx

Protein Information	
Protein	Endothiapepsin
Molecular Weight (kDa)	33.8
Spacegroup	P1211
a, b, c (Å)	45.2, 73.3, 52.7
α, β, ɣ (°)	90.0, 109.2, 90.0
Fixed-target parameters	
Volume loaded per chip (µL)	150
Aperatures per chip	25,600
Required crystal concentration (crystals/mL)	500,000
Sample Information	
Protein mass used to make 200 µL of sample (mg)	10
Crystal longest dimension (µm)	15
Crystal concentration (crystals/mL)	2,500,000
Experimental Variables	
Number of time points required	5
Number of images required per time point	50,000
Hit rate (integrated patterns/images collected)	0.3
Fixed-targets required per time point (rounded up)	7
Sample requirements	
Sample volume required per time point (µL)	1,050
Total sample volume required for experiment (mL)	5.25
Total mass of protein required (mg)	52.5

Name of Material/Equipment	Company
Swissci 96-well 2-Drop plates	Molecular Dimensions
Swissci 96-well 3-Drop plates	Molecular Dimensions
mosquito LCP liquid handling robot	sptlabtech
ClearVue Sheets	Molecular Dimensions
Safe-Tube 1.5 mL	Eppendorf
Scaple	Swan and Morton
MS 3 Vortex	IKA
24-well XRL Plate	Molecular Dimensions
Tube revolver/rotator	Thermo Fischer Scientific
Eppendorf Research plus pipettes	Eppendorf
Eppendorf pipette tips	Eppendorf
Suparen 600	Prochem AG
Sodium Acetate	Sigma-Aldrich
Tris	Merck
Magnesium Chloride	Sigma-Aldrich
PEG 6,000	Sigma-Aldrich
Ethelyene glycol	Sigma-Aldrich
PACT Premier HT screen	Molecular Dimensions
DOW CORNING high vacuum grease	Molecular Dimensions
Hirschmann 22 x 22 mm glaser cover slides	Hirschmann
Crystal pins	PSI
1-1.3 mm SiLibeads Type S	Faust
Macbook Pro	Apple
CCP4 software suite	CCP4
Excel	Microsoft
Hausser Scientific Bright-Line counting chamber	Thermo Fischer Scientific
PACT Premier	Molecular Dimensions
Rock Imager	Formulatrix
Rock MakerWeb	Formulatrix

Formulator	Formulatrix
Leica MZ16 Microscope	Leica
LAS V4.6	Leica
Spectra/Por 3.5 kDa dialysis tubing	Spectrumlabs
Dialysis tubing closures	Spectrumlabs
Amicon 10 kDa centrifugal concentrator	Merck-Millipore
5810 R swing bucket centrifuge	Eppendorf

Catalog Number	Comments/Description
MD11-002	96-well 2-drop crystallisation plate
MD11-003	96-well 3-drop crystallisation plate
mosquito LCP	Crystallisation robot
MD6-015	96-well crystallization plate seals
30120086	1.5 mL centrifuge tubes
No. 3 scalple and No. 3 handle	Scalple for cutting open plate seals
3319000	Vortex for mixing solution and making seed stocks
MD3-11	24-well hanging-drop plates
88881001	Tube revolver for mixing solution during scaling
	Range of manual pipettes, 0.5-10, 1-20, 10-100, 100-1000 µL
	Range of tip sizes for manual pipettes
Suparen 600	Endothiapepsin solution
241245-1KG	Sodium Acetate
8382T014	Tris
M2670-1kg	Magnesium Chloride
81255-1kg	PEG 6,000
324558-1L	Ethelyene glycol for cyro-protecting the crystals
MD1-36	PACT Premier 96-well crystal screen
MD6-02	Grease for sealing 24-well plates
8000104	Cover slides for sealing 24-well sitting drop plates
Manufactured inhouse	Thin-film supports for micro-crystals.
6239547	Glass beads for making mico-seed stocks
Macbook Pro	Computer for performing data analysis
	Diffraction pattern data processing software
Microsoft Office	Plotting tool for phase diagram
02-671-51B	Tool to calculate crystal concentration
MD1-29-ECO	Sparse-matrix crystallization screen
Rock Imager	Temperature controlled crystal plate storage and imager
Rock MakerWeb	Crystal plate creation and image storage stoftware

Formulator	96-well crystal screen creation liquid handling robot
Leica MZ16	Light microscope
LAS V4.6	Software for Leica microscopes
Spectra/Por 3 Dialysis Membrane	3.5 kDa dialysis membrane
Spectra/Por 3 Duniversal Closures	Clips to seal the dialysis tubing ends
Amicon Ultra-15 10 kDa centrifugal concentrator	10 kDa centrifugal filter
5810 R Centrifuge	Swing bucket centrifuge

Editorial comments:
Changes to be made by the Author(s):
1. Please take this opportunity to thoroughly proofread the manuscript to ensure that there are no spelling or grammar issues (Line 329: impropriate?).
2. Please add more details to your protocol steps. Please ensure you answer the “how” question, i.e., how is the step performed?
3. 2.1.2: As this is filmed, please specify the details on how to crush the crystals. If this is not filmed, a citation would suffice.
4. Please note that we need a specific protocol in order to film. Please highlight a specific experiment with specific values and experimental parameters for the filming. We cannot film a generalized protocol.
5. Please present some limitations of the technique in the discussion.

Responses to all reviewers

Many thanks for all your comments and the time you took to make them. Based on your comments and those from the editor, the resubmitted manuscript has been extensively revised. For the ease of understanding how the major concerns have collectively changed the manuscript, some specific notes regarding the introduction and protocol are given bellow.

Introduction:

The introduction has been reworked to include a section describing the protein crystallization phase diagram. We initially felt that this was not the paper to include such a description. The aim of this paper was to describe a crystallization methodology. Although, the method makes use of the diagram as a construct, it is not the ultimate goal. Since others have already provided detailed descriptions and discussions of the diagram, it was felt that this was enough.

However, this is clearly incorrect since the method asks its readers to create such a diagram. Therefore, a description has been given of the diagram, its regions, and the forces at work during crystallization.

We also further contemplated what is the purpose of the diagram in the context of this protocol. Why were we asking readers to create one? And the answer is to try to empirically determine how to increase the number of crystals generated from a crystallization experiment. Can the nucleation rate be increased by increasing the protein or precipitant concentration, and/or are seeds required?

A discussion has also been added to the introduction regarding the relationship between the crystal number and crystal size. This theme has then been carried through the rest of the protocol, and in particular in reference to the Step 2, where a ‘phase diagram’ is generated. We accept the criticism of the reviewers that the diagram generated in the protocol is not a true phase diagram and we accept the term ‘morphogram’ as a better descriptor.

Protocol

The protocol has also been heavily revised. This was principally due to comments from the editors. From the outset, the idea was to write a protocol that could be used for any protein and then, in the results section, describe how it was specifically applied to endothiapepsin. Although the protocol does advise the following of experimental methods, in some respects it is more of an intellectual framework on how to approach the problem of micro-crystal optimization.

This framework would not be readily applicable to a video protocol, as a specific case would be needed to demonstrate the protocol during filming. Therefore, the protocol has been rewritten for the optimization of endothiapepsin. To try and keep the sense that this protocol could be applied to other soluble proteins, notes have been added to describe the decision-making process.

Despite our initial reluctance, Step 1 – the optimization of the crystal morphology – has been added to the protocol. The reason for not including this step explicitly was that, in terms of a general protocol, it will be very protein specific. The path to obtain crystals can be long and may require optimization of the protein buffer, crystallization temperature or perhaps the construct and purification protocol, to name a few. We did not feel this could be generalized into a common protocol. However, since the protocol has now been rewritten for endothiapepsin, the inclusion of Step 1 now makes a lot more sense.

The flow chart has also been considerably simplified to make it easier to follow.

_____

Reviewers' comments:
Reviewer #1:
Manuscript Summary:

The paper by Beale & Marsh presents a protocol to obtain large volumes of protein micro-crystals suitable for serial crystallography and XFEL analyses. In particular, the optimization process aimed at meeting constraints on micro-crystals size and concentration imposed by their final application. The protocol is mainly based on one of the authors publications (Beale et al., 2029), following the same procedure explained in the article. The process of obtaining large volumes of micro-crystals is still mostly empirical, and a trial-and-error approach is often applied. In this scenario, the Authors provide a rational sequence of steps to move from vapor diffusion experiments to large volume batches referring to the protein phase diagram. A series of actions which can increase the probability of success in getting diffraction-quality micro-crystals and minimize protein wasting is discussed. Alternative routes based on seeding and quenching have also been proposed for proteins with difficult optimization. Without going into the chosen representative variable "induction time" the protocol will have to be clearly explained since there are some assumptions that are not common practice in a regular laboratory (see below). Still the community, mainly starting PhD students, may benefit from this JOVE protocol.

Below are the feedbacks from a PhD student, a Postdoc and a colleague and some notes from myself at the end. I have tried to somehow organized their comments, removed duplicities, etc. but other than that, it is a list of comments or suggested corrections/clarifications.
Many thanks for taking the time to seek multiple opinions from researchers at different stages of their careers.

- Overall, the paper provides a clear and detailed strategy which can support the crystallographers' community in the preparation of micro-crystals. Adjustments to specific proteins have, of course, to be considered as the protocol had been only validated for a single model protein, i.e., endothiapepsin."
We questioned whether to include more example proteins, as was done in Beale *et al.*, (2019). More proteins would obviously help to indicate that the protocol was applicable to more than just model proteins, such as endothiapepsin. However, we felt that actually a really detailed description of one protein might have more impact. Very rarely are such detailed descriptions of the crystallization journey of a single protein even published, and this seemed a nice opportunity to do so.

- The manuscript tries to generalize a protocol to obtain crystals suitable for XFEL experiments using Endothiapepsin as a model, based on information obtained from the phase diagram. The authors honestly describe that generalization of this procedure is an almost impossible task, but provide different approaches and strategies advising potential readers on the XFEL field on how to proceed among the different difficulties that might be found during crystallization setup. The strategy is mainly based on a manuscript previously presented by one of the authors (Beale et al., 2019, ref 26 in the manuscript), where they already suggested "(i) how vapour diffusion conditions can be converted into batch by optimizing the length of time crystals take to appear; (ii) how an understanding of the crystallization phase diagram can act as a guide when designing batch crystallization protocols; and (iii) an accessible methodology when attempting to scale batch conditions to larger volumes".
This comment highlights one of the chief concerns we had during the paper's conception, which was not to present the protocol as definitive, *i.e.* would always work. From our experience, the venn diagram of protein crystallography and certainty can be illusory. We would ideally present conceptual ideas rather than definitive steps.

- The article illustrates in a very systematic way how large volumes of Endothiapepsin microcrystals have been achieved. In my opinion, the article is aimed mainly at those researchers who are starting to obtain microcrystals.
We agree, although we would also say that some of the ideas presented could be applied by anyone wishing to manipulate the growth of crystals. For example, growing very large crystals – for neutron diffraction.

- In terms of clarity in reading, it was easy for me, really. Each point is extensively and correctly illustrated. The beauty of the article is the extremely systematic nature of the work, although I think everything they have done has a point of logic, that is why I think this work is indicated rather for people who start from scratch on the subject. I understand that the work has a protocol character and, therefore, what it intends is to clearly show the rationalization of the entire process.
We are grateful that this impression was received as this was the goal.

- As the authors clearly state in the manuscript (lines 61-73), it has been already shown how to move from hanging drop crystallization experiments to batch configuration. It would be advisable to mention the work by Rayment, 2002 (Small-scale batch crystallization of proteins revisited: an underutilized way to grow large protein crystals. Structure. 10(2):147-151) or that from Luft et al,. 2009 (Efficient optimization of crystallization conditions by manipulation of drop volume ratio and temperature. Protein Sci. 2007;16(4):715-722), as examples. In this sense, paragraph 61-73 should be rephrased to soften the novelties described in this manuscript.
We thank the reviewer for these suggestions. We were not aware – particularly of Rayment (2002) – which already present many of the ideas in the protocol.

Major Concerns:
My main concern, probably left-over from the original Beale article, is the use of the phase diagram argumentation when it is never determined, not for endothiapepsin neither for the

proteins used in the original article, although some figures were named "Phase diagrams". In any of those cases, neither in this article, it is determined the solubility curve, the main piece of data to work with a phase diagram. Besides this concern I understand that to simplify the discussion the term phase diagram is used but to the light of scientific precision perhaps it will be good to clearly state that they are using a hypothetical phase diagram (a morphogram, …). This may be an excellent opportunity to introduce the concept of phase diagram, induction time and even supersaturation rate, something used but not explained along the text.

I – John Beale – from the first paper, and we John and May from this one, fully agree with this assessment. This argument was voiced by other reviewers and we have been struck by the force of it. This point has principally been answered in the introduction, and we particularly thank you for suggesting the term morphogram as a way to square this circle.

Minor Concerns:

- I have doubts about whether it can be generalized by stating that nucleation time is a good metric to choose to go from diffusion to batch without first giving details about for example the droplet volumes that have been used in the initial experiments. In the event that the volumes of the drops in the 96 plates have, for example, a total of 100nL, it is hard to believe that the vapor diffusion does not have a relevant impact. Also, if the experiments are done with nano-drops, the crystallization robot loses precision and then a randomness factor comes into play that can become important. Do you use humidification when working with the robot?

Thank you for highlighting this point. In any experiment described in this paper, a total drop volume of < 200 nL was never used and neither was humidification. Volumes have now been added where missing for clarification. We agree that the drop volume will certainly influence the equilibration time of a vapor diffusion experiment.

- Also, the selected system, Endothiapepsin, seems to me a protein such as lysozyme since it crystallizes in many of the PACT screening conditions. This high tendency to crystallize invalidate "induction-time" as the best indicator for the transition from vapor diffusion to batch. Furthermore, if we are going to consider only crystallization conditions that lead to crystals in less than 24 hours, why not then consider creating the phase diagram from a microbatch experiment? The scaling would then be much easier, and the conclusions much more precise.

We agree endothiapepsin is essentially a ‘test protein’. One of the reasons for its use was to really explore lots of different large volume (>100 µL at 100 mg/mL) experiments without the need to purify lots of protein. However, we do not follow the reasoning that this invalidates induction time as the transition metric. The propensity of a protein to crystallize is determined by how easily the thermodynamics of crystallization can be satisfied? This could be in one or many conditions. Not all of the PACT conditions showed crystals at the same time, suggesting that there were differences in the individual conditions. Differences that could be optimized away if one was inclined to do so.

Regarding the use of micro-batch. This was also one of the comments from the reviewers of the original paper, and we think the answer is the same and comes in two parts. Firstly, micro-batch is perhaps a slight misnomer – just because batch is in the name of the method, this, in of itself, does not necessarily mean that a condition that yields crystals does so in a

batch-like process, *i.e.* the nucleation zone has been hit upon the mixing of the protein and reservoir solutions. In the case of the micro-batch, other processes such as evaporation of the liquid from the drop are as least as likely to have been the drivers of crystallisation, *i.e.* upon mixing of the solutions the nucleation zone is not hit and evaporation of the drop is required to transitions to it. The need for evaporation in micro-batch is well exemplified by the early and late microbatch papers see:

Chayen, N.E., Shaw Stewart, P.D., Blow, D.M. *Journal of Crystal Growth*. **122** (1–4), 176–180, (1992).
D'Arcy, A., Mac Sweeney, A., Stihle, M., Haber, A. *Acta D*. **59** (2), 396–399 (2003).
Brumshtein, B., Greenblatt, H.M., Futerman, A.H., Silman, I., Sussman, J.L. *Journal of Applied Crystallography*. (2008).

Micro-batch, like vapor diffusion is often a 'transitionary method'. Therefore, one might face the same need to optimize the crystallization condition to batch in a micro-batch plate as from a vapor diffusion plate. If this is the case, and the same results can be gained by remaining in vapor diffusion plates, it is perhaps easier than introducing a new step.

Secondly, unfortunately, micro-batch is a niche technique and not practiced by the majority of crystallographers. Vapor diffusion is, for better or worse, the most common form of crystallization method. Our feeling, from discussion with crystallographers, if you want to create a method that will be widely used, stay as close to what people are already comfortable with.

- I modestly think that the "standardization" wording should be softened along the manuscript, remarking that this paper includes different strategies that overcome problems found during preparation of samples for XFEL experiments (as they actually conclude in line 692). My suggestion is based on the specific material required for the production of the "phase diagram", which is not available not even in all of the less than 50 worldwide synchrotron light-sources. It would be advisable to comment on alternatives for crystallization laboratories not having liquid handling robots.
We agree that the method essentially requires the use of liquid handling and plate storage robots. To better highlight this drawback, a note has been added to the beginning of the protocol and a larger paragraph discussing the weaknesses of the method has been added to the discussion.

- Whereas I understand that image acquisition during the first 24 hours is of high importance, intervals for labs not allowing for automatic image recording should also been proposed.
We agree with this suggestion and have added suitable alternative timings.

- As the authors state, optimizing crystal morphology is a determinant step in this protocol, and should be included. My suggestion is including the process they followed, leaving their statements on the difficulties to standardize Step 1, and suggesting alternatives where they found some problems.
As stated in the introduction to these responses, this has now been done.

- Most XFEL users will be aware that the amount of protein for XFEL experiments is huge, but this should be stated somewhere in the introduction for beginners not used to this specific methodology.
We readily appreciate this point. To try and give a flavor of the sample requirements of a pump-probe experiment, Table 1 has been inserted and described in the introduction.

- Line 304: "3.4.1. Consider agitating the tubes or plates to prevent crystal settling and to aid nucleation. " I don't understand why shaking the plates or tubes can promote nucleation. I would rather bet that it would favor homogeneous crystallization.
We agree that it favors homogeneous nucleation, but it also has been shown to reduce the onset time of observation crystallization. This information has been added to the results.

- Line 463: "However, when the drop volume was increased to 200 µL, no crystal growth was observed in this condition. Figure 8 shows the results of a large volume (200 and 300 µL, without and with seeds, respectively), batch crystallization of the endothiapepsin. In all of these experiments the endothiapepsin solution was mixed with the precipitant at room temperature (20 ° C), vortexed for 10 s, and then continuously agitated for 24 h on a rocker-shaker. " One possible reason why they did not obtain crystals in the initial scaling may be the fact that the protein was vortexed and agitated so much with the crystallization cocktail: by doing this you are going from a system with convective processes to a system with diffusive processes. I understand that the reason for so much agitation is due to wanting to achieve a system that is as homogeneous and reproducible as possible. Be that as it may, I think the reader would appreciate an explanation of the purpose of such agitation.
The crystallization process is a possibility. This is one of the reasons in this protocol, when working in small volumes, we suggest to mix the protein and precipitant together rather than simple addition. It is very possible that this agitation was not enough so diffusion driven process played a role in crystallization in the plate and not in the Eppendorf. We have added, as per the previous point, a note regarding agitation when scaling.

- Point 2.3 does not seems very realistic to me since people do not have a synchro
We agree with this sentiment, however, it really is vitally important that the crystal diffraction is confirmed at some point, and hopefully more frequently. Non-diffracting samples are a common occurrence during serial crystallography experiments, and, in all honesty, one of the reasons to write this paper.

- Several grammatical mistakes appear thorough the manuscript. I would recommend a careful revision:
We thank the reviewer for their time spotting these.

- Line 36: is increased
Changed.

- Line 39: can adapt
Changed.

- Line 43; Room temperature could be abbreviated as RT, and changed through the manuscript

Changed.

- Line 45: the development
Changed.

- Line 47: but there is also
Included

- Line 50: attraction
We believe the reviewer was referring to: “However, the principal reason why cryo-methods gained traction over RT approaches in the late 1990s was the suppression of radiation damage by sub-zero crystal temperatures[18]”.. “gained traction” was our intention and we think it is correct.

- Line 85: it is important
This line has been rewritten.

- Line 98: techniques
Changed.

Responses to all reviewers

Many thanks for all your comments and the time you took to make them. Based on your comments and those from the editor, the resubmitted manuscript has been extensively revised. For the ease of understanding how the major concerns have collectively changed the manuscript, some specific notes regarding the introduction and protocol are given bellow.

Introduction:

The introduction has been reworked to include a section describing the protein crystallization phase diagram. We initially felt that this was not the paper to include such a description. The aim of this paper was to describe a crystallization methodology. Although, the method makes use of the diagram as a construct, it is not the goal. Since others have already provided detailed descriptions and discussions of the diagram, it was felt that this was enough.

However, this is clearly incorrect since the method asks its readers to create such a diagram. Therefore, a description has been given of the diagram, its regions, and the forces at work during crystallization.

We also further contemplated what is the purpose of the diagram in the context of this protocol. Why were we asking readers to create one? And the answer is to try to empirically determine how to increase the number of crystals generated from a crystallization experiment. Can the nucleation rate be increased by increasing the protein or precipitant concentration, and/or are seeds required?

A discussion has also been added to the introduction regarding the relationship between the crystal number and crystal size. This theme has then been carried through the rest of the protocol, and in particular in reference to the Step 2, where a ‘phase diagram’ is generated. We accept the criticism of the reviewers that the diagram generated in the protocol is not a true phase diagram and we accept the term ‘morphogram’ as a better descriptor.

Protocol

The protocol has also been heavily revised. This was principally due to comments from the editors. From the outset, the idea was to write a protocol that could be used for any protein and then, in the results section, describe how it was specifically applied to endothiapepsin. Although the protocol does advise the following of experimental methods, in some respects it is more of an intellectual framework on how to approach the problem of micro-crystal optimization.

This framework would not be readily applicable to a video protocol, as a specific case would be needed to demonstrate the protocol during filming. Therefore, the protocol has been rewritten for the optimization of endothiapepsin. To try and keep the sense that this protocol could be applied to other soluble proteins, notes have been added to describe the decision-making process.

Despite our initial reluctance, Step 1 – the optimization of the crystal morphology – has been added to the protocol. The reason for not including this step explicitly was that, in terms of a general protocol, it will be very protein specific. The path to obtain crystals can be long and may require optimization of the protein buffer, crystallization temperature or perhaps the construct and purification protocol, to name a few. We did not feel this could be generalized into a common protocol. However, since the protocol has now been rewritten for endothiapepsin, the inclusion of Step 1 now makes a lot more sense.

The flow chart has also been considerably simplified to make it easier to follow.

Reviewer #2:
Manuscript Summary:
The paper describes in practical details how to proceed from existing vapour diffusion conditions to scaled-up batch conditions suitable for growing slurries of micron-sized crystals for serial methods. Overall, I really like the paper - it is well written and gives and honest and real world example on how to use the phase diagram to guide decisions regarding optimization for large volumes of small crystals. Also, room temperature serial crystallography is a booming field and there is not a lot of literature on methods to address the unique sample needs. Usually there is a lot of talk about sample delivery and how to get the crystals in the beam - but really not much about how to get the sample "just right" from the beginning.

Major Concerns:
I have a general comment that should be addressed and that is: if the goal is to find batch conditions in large volumes, the early steps are a waste of time and materials as small drop vapour diffusion set-ups are nowhere near where you need to be (in the phase diagram) for where you aiming to go - as the authors point and out and show as the optimizations progress. I understand that for this paper, this is how the authors arrived there. However, for someone that has managed to fine-tune VD conditions for single, large and well-diffracting crystal conditions (i.e. completed Step 1 ages ago), it is prudent to switch to batch immediately and do all assessment and fine-tuning in batch right away.
We appreciate the tenor of your point. There is definitely an argument for moving immediately to a large volume batch condition once one has been identified. We have found that there is often little difference in results from 500 nL to 10 µL. Changes seem to mainly occur when moving to greater the 50 µL. However, we still feel that gradually scaling the condition can be useful, particularly in cases where the cost of protein production and purification, either in time or finance, are significant. There may be cases where the failure in a smaller volume condition saves the loss of precious material that would have otherwise been lost by rushing to larger volumes.

With regard to the second point about moving straight to batch – we assume this implies, moving straight to micro-batch? There are two reasons why we suggest to do everything in VD plates. Firstly, micro-batch is perhaps a slight misnomer – just because batch is in the name of the method, this, in of itself, does not necessarily mean that a condition that yields crystals does so in a batch-like process, *i.e.* the nucleation zone has been hit upon the mixing of the protein and reservoir solutions. In the case of the micro-batch, other processes such as evaporation of the liquid from the drop are as least as likely to have been

the drivers of crystallisation, *i.e.* upon mixing of the solutions the nucleation zone is not hit and evaporation of the drop is required to transition to it. The need for evaporation in micro-batch is well exemplified by the early and late microbatch papers see:

Chayen, N.E., Shaw Stewart, P.D., Blow, D.M. *Journal of Crystal Growth*. **122** (1–4), 176–180, (1992).
D'Arcy, A., Mac Sweeney, A., Stihle, M., Haber, A. *Acta D*. **59** (2), 396–399 (2003).
Brumshtein, B., Greenblatt, H.M., Futerman, A.H., Silman, I., Sussman, J.L. *Journal of Applied Crystallography*. (2008).

Micro-batch, like vapor diffusion is often a 'transitionary method'. Therefore, one might face the same need to optimize the crystallization condition to batch in a micro-batch plate as from a vapor diffusion plate. If this is the case, and the same results can be gained by remaining in vapor diffusion plates, it is perhaps easier than introducing a new step.

Secondly, unfortunately, micro-batch is a niche technique and not practiced by the majority of crystallographers. Vapor diffusion is, for better or worse, the most common form of crystallization method. Our feeling, from discussion with crystallographers, if you want to create a method that will be widely used, stay as close to what people are already comfortable with.

Sparse-matrix screening can be done in batch and I would also recommend to go immediately for 1 or 2 uL drops with/without seeding. Then the "only" adjustments to be made would be in the scaling and that - as pointed out - typically requires higher concentrations of precipitant. That can then be the focus of a phase diagram. The way the procedure is currently written is unnecessarily complicated with too many decision points of "if yes then proceed to…" "if no, the return to step…". Most scientists will come up with their own decision tree if the parameters to be varied, the selection criteria, and desired outcomes are clearly explained. The current "path" also involves setting up quite a large number of trays for conditions and volumes (i.e. 500 nL VD drops) that will mean very little for the goal (large volume batch).
The first point is partly answered in the previous response, but to expand more. For the majority of target proteins, this method can be followed from Step 2 – moving into batch. Since the majority of protein are grown *via* vapor diffusion, it makes sense to stay in vapor diffusion since the move to batch is likely to require the same optimization that needs to be performed for vapor diffusion.

Regarding the complexity of the decision tree. We absolutely agree and have adjusted it to contain fewer steps with clearer decision points.

Minor Concerns:
1) Excellent topics raised throughout that should be highlighted in the summary/conclusion are the selection criteria: extent of nucleation, time it takes for crystals to appear, and the benefits of microseeding to achieve/accelerate the previous two parameters.
Thank you highlighting the need for clarification. We believe this has now been achieved by the additions to the introduction regarding the relationship between crystal size and

number. And then, referencing these parameters throughout. A better summary of the results has also been added to the discussion.

2) Please include a brief summary at the end of what your best starting conditions was (from VD, large crystals) and what they ended being for serial crystallography. It seems that MgCl2 and pH was not a huge change but the amount of PEG really needed to be increased and getting seeding "just right" was also important.
A paragraph highlighting this point has been added in the discussion.

3) Finally, quenching is not something I ever considered and clearly this is very important to stop small crystals from continuing to grow while you get ready for beamtime. Maybe a few more sentences to describe this will be very helpful to readers.
A more detailed description of the quenching process has been added to the results.

Figures - there are too many redundant figures, they really need a bit of streamlining:
We appreciate this point and have made some changes based on your points.

Figure 1 should be reduced to 2 panels - one classical phase diagram and then one adapted to show progression for batch. The effect of seeding can be simply described in a sentence or use a different coloured arrow. For colour-blind readers it would be good to avoid red on green and vice versa.
We do appreciate the wish to make figures clear and precise for readers, and we think the suggestion to only have one 'phase diagram' figure is a good. Figure 1 is now the only 'phase diagram' figure. However, we also think that reducing Figure 1 down to only two panels is too reductionist, and then relies too heavily on describing alterations in the text. We feel that these descriptions are more likely to be open to misinterpretation than an extra panel. Therefore, Figure 1 now has 6 panels, describing the key ideas throughout the paper.

Figure 2 - far too busy to be practically used as a decision-making tool. This has to be streamlined to a few simple questions and decision points. It can be made clear that the selection criteria are simply: 1) amount of nucleation, 2) crystal size, 3) time it takes to appear and favourable outcomes are: lots of nucleation, small micron sized crystals and ~24h. All of this is to be done on the background of with/without seeding (once an optimal seed stock concentration has been determined). Based on earlier comment, step 1 should really be, small(ish) volume batch drops using sparse matrix screens - with and without seeds. The next plate - Step 2 - should then be a screen in batch format around your best condition, but varying seed stock concentration (1x, 10x). From there it should be possible after 2 plates to go for scaled up volumes - applying the same selection criteria as before.
We agree with your assessment of the figure, it has now been redone with fewer decision points and with the goals of each step clearly highlighted.

Figure 3 - again, sufficient to use panel A and C only.
Now Figure 4. This has been decreased to two panels as suggested.

Figure 4 - Excellent figure but very "endothiapepsin specific". However, it's good to illustrate the effect of % PEG, pH and MgCl2 concentration. It would ease the reading if you can indicate on each crystal pic the pH, PEG and MgCl values (rounded up/down).

Now Figure 5. The crystallization conditions have been marked on the figure along with our assessments of the crystallization from each.

Figure 5 - redundant if Figure 1 is streamlined. It is sufficient to simply describe in words how you are stepping around phase diagram when changing different parameters and refer to figure 1.
The relevant panel from this figure has now been combined into Figure 1.

Figure 6 & 7 - excellent figures that really illustrate what is happening and what the reader should expect/aim for when setting their own plates.
This figure is essentially the same but, using the same reasoning as the previous point, the essential part of the condition is now given on the figure.

Figure 8 - very difficult to see what is happening and I would leave it out as Figure 9 does a better job to illustrate what is happening. From figure 8 alone it is not obvious what is the best condition as the crystals are very difficult to see (I am assuming based on the text it's the last 2 rows?).
We agree and the figure has now been removed.

Figure 9 - like figure 5, it should be removed and it is sufficient to refer to figure 1 and explain in a sentence or two.
The important panels from this figure have been included into Figure 1.

Figure 10 - excellent figure - but the lines of the curve are difficult to tell apart based on colours alone - can a combination of dashed and continuous lines and colours be used?
We agree. To draw some distinction between lines of similar hue, dashed lines have been added.

Excel worksheet - excellent to provide this for readers to use.
We are glad that you think this will be useful.

Responses to all reviewers

Many thanks for all your comments and the time you took to make them. Based on your comments and those from the editor, the resubmitted manuscript has been extensively revised. For the ease of understanding how the major concerns have collectively changed the manuscript, some specific notes regarding the introduction and protocol are given bellow.

Introduction:

The introduction has been reworked to include a section describing the protein crystallization phase diagram. We initially felt that this was not the paper to include such a description. The aim of this paper was to describe a crystallization methodology. Although, the method makes use of the diagram as a construct, it is not the goal. Since others have already provided detailed descriptions and discussions of the diagram, it was felt that this was enough.

However, this is clearly incorrect since the method asks its readers to create such a diagram. Therefore, a description has been given of the diagram, its regions, and the forces at work during crystallization.

We also further contemplated what is the purpose of the diagram in the context of this protocol. Why were we asking readers to create one? And the answer is to try to empirically determine how to increase the number of crystals generated from a crystallization experiment. Can the nucleation rate be increased by increasing the protein or precipitant concentration, and/or are seeds required?

A discussion has also been added to the introduction regarding the relationship between the crystal number and crystal size. This theme has then been carried through the rest of the protocol, and in particular in reference to the Step 2, where a ‘phase diagram’ is generated. We accept the criticism of the reviewers that the diagram generated in the protocol is not a true phase diagram and we accept the term ‘morphogram’ as a better descriptor.

Protocol

The protocol has also been heavily revised. This was principally due to comments from the editors. From the outset, the idea was to write a protocol that could be used for any protein and then, in the results section, describe how it was specifically applied to endothiapepsin. Although the protocol does advise the following of experimental methods, in some respects it is more of an intellectual framework on how to approach the problem of micro-crystal optimization.

This framework would not be readily applicable to a video protocol, as a specific case would be needed to demonstrate the protocol during filming. Therefore, the protocol has been rewritten for the optimization of endothiapepsin. To try and keep the sense that this protocol could be applied to other soluble proteins, notes have been added to describe the decision-making process.

Despite our initial reluctance, Step 1 – the optimization of the crystal morphology – has been added to the protocol. The reason for not including this step explicitly was that, in terms of a general protocol, it will be very protein specific. The path to obtain crystals can be long and may require optimization of the protein buffer, crystallization temperature or perhaps the construct and purification protocol, to name a few. We did not feel this could be generalized into a common protocol. However, since the protocol has now been rewritten for endothiapepsin, the inclusion of Step 1 now makes a lot more sense.

The flow chart has also been considerably simplified to make it easier to follow.

Reviewer #3:
Manuscript Summary:
In their article entitled "optimizing the growth of endothiapepsin crystals for serial crystallography" Beale and Marsh describe a strategy to translate vapor diffusion conditions yielding large crystals into batch conditions leading to large quantities of microcrystals for SMX/SFX. They make us of a phase diagram analysis and provide a worksheet to facilitate its analysis and representation.

Major Concerns:
This text, along with the video, is supposed to introduce the biochemist to a completely new field: massive batch crystallization to prepare suitable samples for serial crystallography. In that respect, it should provide a solid theoretical basis, especially when talking about phase diagrams. However, this is not the case: descriptions are vague, if not erroneous (see for instance, the confusion between seeds and nucleants, precipitants and precipitates). It is crucial that the zones of the phase diagram are clearly identified and their specificities (in terms of molecule behavior) explained in details (see Fig1 and comments below).
One cannot talk about phase diagrams without defining the concepts of solubility and solubility curve, supersaturation vs nucleation probability, kinetics vs thermodynamics of crystal growth. The authors present a lot of experimental data that could be better exploited to illustrate these concepts and bring up rational explanations to phenomena that simply described in the text as facts.
With reference to the introductory note – we completely see the reviewers point regarding the phase diagram description and also humbly agree with our poor use of language in descriptions. The introduction has now been extensively reworked to include such a phase diagram description and the whole document has been corrected so as, for example, not to label seeds as nucleants.

We still feel that going into immense depth describing the diagram is not appropriate from this paper. Others have already done this in a more complete way than we could hope to. This said, on rereading the submitted text, we realized that we had failed to effectively point readers to these articles. These citations have now been included and we hope that a correct picture of the diagram, and its use with regard to this method, has now been given.

For instance, knowing the initial protein concentration in a drop, the resulting number of crystals and their average size, the authors could calculate the amount of crystallized material and estimate the remaining concentration of soluble protein in equilibrium with the crystalline phase (i.e. the solubility). In addition, with such a number of droplets, it

would be easy to actually measure this solubility in the mother liquor and draw a 'real' phase diagram.
This point caused the reflection referred to in the introductory note regarding the point of trying to create the phase diagram – or morphogram – as it is now referred to. Ultimately, we agree with the reviewers point that the 'phase diagram' we suggest to make – is not a true phase diagram. Hence, the decision to take another reviewers suggestion and call it a morphogram.

The reason to essentially perform a minor linguistic adjustment, rather than a major experimental one, was twofold. Firstly, we decided the reason to perform the morphogram experiment was really to determine two things: 1. is crystallization probably happening in batch? 2. Do you need to use seeds to increase the crystal number? Or can enough nucleation be generated through simply mixing protein and precipitant? These questions can be answered by performing the morphogram experiment as we describe. Secondly, we feel the method would lose accessibility if we asked readers to go into more detail, and define the diagram more precisely. Also given the goals and we do not feel that it is strictly necessary.

Also, there is not a single way to perform serial crystallography. The authors should clearly indicate in the introduction what kind of material (what size vs which X-ray source) is required and in which quantity (100 µl does not mean anything). XFEL have been advertised to be able to exploit micro (<10µm), if not nanocrystals (<1µm)...
We very much agree with this assessment. In the first paper [Beale *et al*., (2019)], the importance of deciding what kind of sample is required, was highlighted. Different sources, delivery methods and experiments will dictate the ideal sample. We have tried to give a flavor of these variables in the introduction and in the new table (Table 1), which also tries to give an honest assessment of the mass of sample required for a particular experiment.

While reading the protocol, I was wondering how much protein quantity had been necessary to carry out the endothiapepsin study. It is only at the penultimate paragraph that I found the answer: 100 mg. But who has 100 mg of a none commercial protein to spent? Is it possible to achieve the same goal with a drastically lower sample consumption? These important practical aspects are completely missing in the discussion.
We agree that 100 mg is a lot. Thank you for pointing out that this was poorly highlighted. We hope the new Table 1, discussion thereof, and the addition to the discussion will better clarify the possible expense in serial crystallography. We would stress though that this expense is more to do with the demands of particular experiments rather than this method, per se.

For all these reasons, the article requires a major revision to ensure that any reader with little background in crystallogenesis is able to understand the logic and reproduce the protocol. This would dramatically improve the impact of this work, especially with the support of the upcoming video. The comments below are meant to be constructive and I look forward to reading a revised version and watching the video protocol.
Many many thanks for taking the time to make such in depth comments. It is really appreciated.

Minor Concerns:
The structure of the protocol (headings) is not homogeneous. Explicit titles should be inserted for each major step throughout or should be removed every where (see section 2 and JoVE instructions).
This has now been done.

introduction
l45 : 'developed' development?
Changed.
L51: 'suppression of radam' not really suppressed, but slowed down
Changed.
l57: 'large quantities (...µl) of micro-crystals (...µm)' as 'µl' refers to solutions and not crystal volumes, it would be more correct to say 'large quantities of solutions (or suspensions) of micro-crystals'
Changed.

Figure 1: this figure was adapted from Beale et al (2019) but a part of the legend was omitted, which makes it confusing.
The purple line (called the 'limit of protein supersaturation') has an official name: it is the solubility curve. The authors should use it to avoid any misunderstanding.
The color code of the diagram's zones is not defined: undersaturated, metastable, nucleation.
What about the orange zone (top right in B and C)? What does it correspond to?
The whole strategy described in this paper is based on the interpretation of the phase diagram (and I like this idea). This is why it is important that the legend of this first figure is clear and complete.
Figure 1 has now been reworked to include all phase diagram figures. An extra panel has been added, as panel A, which simply describes each part of the diagram. The solubility line has been added and corrected, not only here, but in the rest of the document as well.

Fig1A: 'transition' is very misleading too, it looks like a zone name (just like nucleation), but it seems to be an event (drop transition/ equilibration). If this is what the authors mean, they should explain it and add 'crystal growth' between 'nucleation' and 'end'.
Thank you for highlighting this. The figure labels and legend have now been changed to clarify the meaning of transition. A larger explanation has also been given in the text.

Fig1A: 'An example of the nucleation zone penetration limits of a transitionary phase crystallization method. The blue dotted line marks a theoretical limit of the transition into the nucleation zone. ' What does this mean??? Impossible to understand without reading Beale et al (2019).
Note that usually the limit between the metastable and the nucleation zones (black line) is shown as a dotted line because it is transient and moves towards the solubility curve over time (a question of probability of nucleation).
This has primarily been covered by the previous point. The division between the metastable and nucleation zones has also been greyed to imply the fluidity between the two.

Fig1C: it is not correct that 'protein seeds act as nucleants', they help bypass the nucleation step and crystal growth can then start at lower supersaturation without the need of any nucleation event.
Thank you for highlighting this point – this error has been corrected here and in the rest of the document.

Fig1D: the authors must better explain this panel in the text and in the legend. The first step seems to indicate that crystals are obtained in 20 days... in the undersaturated zone... which is simply not possible! How can crystals grow in 10, 4, 2 days in the metastable zone (bleu zone)? So, please give a clear description of what each dot (starting/end position?) and arrow correspond to in this process of optimization.
This panel has ultimately been removed as it was deemed more confusing than helpful.

L82/l90/title of Fig1: the term 'transitionary phase (crystallization) method' is very misleading when talking about 'phase diagram'. Could 'phase' be replaced by 'step'?
We agree. The sentence has been reworked essentially as you have proposed.

L110: have therefore been? Excluded
This sentence has been removed as Step 1 is no longer excluded.

Protocol:

L133: 'phase diagram... can only be reliably created... via batch' Why? Do the authors mean 'in this context' or more generally? In the latter case, it is certainly not true, see early works on phase diagrams https://doi.org/10.1107/S0907444998010890
L135: 'wether protein crystallization is batch-like' what does batch-like mean? Please explain
We agree that L133 is incorrect and has been removed. With regard to the line in L135, we agree that there is confusion in this phasing, and so here, and elsewhere in the document, we have changed such descriptions to simply say batch.

L140: this is usual called a grid screen => Creat a grid screen based on the crystallization condition in a 96-well plate...
Changed.

L142: does the 'precipitant concentration' refer to the concentration before or after mixing with the protein? In the latter case, how would you achieve 40% (w/v) PEG 3350?
Thank you for highlighting this potential confusion. Throughout the document, all concentrations refer to premixed reagent concentrations. A note has been added to the beginning of the protocol and this has also been made clear in the 'morphogram generator'.

L144: '5 - 40%' precise % (w/v) throughout the text
Due to the changes in the format of the protocol, this sentence has been removed.

L147: how do you create your seed stock? How do you assess its 'concentration/effectiveness'? How do you make the seed stock preparation reproducible to ensure that your resulting phase diagram is reproducible? Give some hints, or at least a few good references. The reader may not be familiar with this important aspect of the protocol.

There is now a full protocol for Steps 2.1.1. and 3.2.4.1., when a seed-stock is made for optimizations and the scaled protocol, respectively. Notes have also been added to aid the inexperienced reader.

L161: 'until crystals appear'... in one drop, in all drops, in most of them??? How long shall one wait?
This has now been clarified to "…every 0, 3, 6, 12, 18, 24 h, then every day for the first week, and every week for the next four."

Phase diagram...
L164: A note could be added below the title to indicate that a worksheet is available as supplementary material to plot the phase diagram. This is an important added value of this work.
A note has been added.

L178: again, seeds do not act as nucleants. With a nucleant, you go through a nucleation step but you lower the activation energy. With seeds, you directly start the crystal growth and bypass the nucleation step. Please rephrase.
As per the response above, this has now been changed.

L174: 'and therefore not batch' weird sentence...
This sentence has been removed.

L182: is it not possible to use an image analysis software like imagemagic to estimate the number more automatically and accurately?
Thank you for this suggestion. We would like to explore easy ways to do this analysis automatically. The Formulatrix software already make a similar calculation with regard to the presence of crystals in the drop. It may be very possible to use part of the data for this calculation as a rapid means to count crystals. However, we have presented the manual way as one that all can do.

L197: 'some precipitant' precipitate!
Mea culpa..

L201: Add a title to 2.3 like 'Crystal analysis'
Added.

L219-227: problem of text organization. 2.4 should be 2.3.3, 2.5 should be followed by the title 'Optimize crystal growth time' ('Optimizing crystal growth time' to be homogeneous with 2.2...)
This has been amended as per the reworking of the protocol.

L222-236: what do the authors exactly mean by 'starting point'? Is it the x,y coordinates in the phase diagram just after mixing the solutions to prepare a VP ou batch drop?
A more complete description has been added to the results to clarify this point.

L224: 'mixture is already in the metastable region' actually 'in the metastable or nucleation region' since you cannot distinguish these regions using seeing...
We would humbly disagree. We think the comparison between the seeded and unseeded drops does allow the nucleation zone and metastable regions to be distinguished. If users of this protocol have failed to see any crystals grow in their unseeded drops in the morphogram experiment, they will have proceeded to this step. The seeded drops now might yield some information. If crystals have appeared in the seeded drops and not the unseeded, this would imply that the drops with crystals may be in the metastable region. They may be multiple drops, perhaps indicating in which direction the optimization must go in order to move the initial drop mixture towards the nucleation zone, and therefore batch.

L238: 'return to step 2.1' Why?????? It seems you already found what you were looking for, why would you step back? Something is missing in the explanation...
A note has been added to further explain. We feel there is value in having a plotted morphogram prior to attempting to scale the condition. Yes, crystallization now happens in less the 24 h – an indication of a batch crystallization process. But the question of how to increase X_n has still to be answered. Are seeds required, or can it be done with nucleation alone? The morphogram does not uniquely provide this information, but is a simple tool to try and answer this question.

L243: 'batch crystallography' => batch crystallization
This sentence has now been changed.

L244/247: not 'nucleants' but 'crystals or growth events' A nucleant is a compound that promotes nucleation, nothing to do with seeds, which are tiny pieces of crystals.
We very much agree and this sentence has been removed.

L245: 'predetermined by the volume of seeds' no, by the volume of seed stock solution (added to the experiment). If you disagree, please explain how you determine the seed volume.
We very much agree and this sentence has been removed.

L254: 'Anecdotally' this does not seem anecdotal at all, but a very practical point. Please elaborate in view of effects of drop volumes on nucleation probabilities described in the literature.
We agree that this is an interesting point. From the proteins that we have tried micro-crystallize using this method, say 6 since it was formalized into method and 6 further proteins whose scaling it was based on. 3 of the 6 needed some tweaking from the initial morphogram, and in this respect, *i.e.* higher protein and/or precipitant concentration when in a larger volume. We are sure that there are extensive examples of this in the literature; however, we would say that they are not obvious in the XFEL literature.

L284: 'increase the final drop volume to the maximum allowed by the plate and/or robot.' Give typical volume values to illustrate this point.
This has now been given.

L304: 'chimney§well plates' give the reference in the table of materials

The reference to chimney well plates has been removed as this was not done in the case of endothiapepsin.

L311: 'hemocytometer' give the reference in the table of materials

This has been added.

L327: 'move onto another possible scaling route' which one? Not clear, please explain

This point has been clarified.

Representative results

This part is much too long and should be more precise and reduced by a third, or even by half. In particular, the paragraph on morphology optimization does not bring anything, since the concept of morphology is not clearly defined and the influence of supersaturation (central point in a phase diagram) on the growth of the faces of a crystal and thus of its morphology are not discussed. If 'no protocol has been provided' (l332), then just skip this from the manuscript.

As described in the introductory note – this section has now been added to the protocol and so has been left in the results.

Regarding the length of the Results as a whole, efforts have been made to remove unnecessary sentences. However, ideas for 'other proteins' *i.e.* not endothiapepsin have had to be left in the results as moving them to the protocol does not fit with the guidelines for JoVE paprs.

We appreciate it is long, however, we both felt that a detailed description of these steps is of value since they are rarely described.

L333: '...and seemingly bound to the whims of the 'crystallization gods' that to attempt to establish a single protocol might cost the writer their sanity [see McPherson and Gavira (2014)}' Should the reader see some kind of humor here? How is it possible the write such a sentence and cite Alex McPherson who spent his entire scientific life trying to bring rationality in this field??? Thanks to him (and others), growing useful (not beautiful!) crystals is no longer alchemy or magic, but a true scientific discipline called crystallogenesis. A more grounded discussion on the theoretical basis of crystallogenesis would be very beneficial to this paper.

This was a poor attempt at some humor, and poorly written – the fault lying here with John – rather than May. The idea was to convey a sense of uncertainty. Often when you read crystallization methods papers, it can feel like that the method implies that crystals will always grow if you follow the steps. We really didn't want to give that impression, particularly with the first step, which in essence, is finding a crystallization hit. A far from certain enterprise. We feel Steps 2 and 3 are on firmer grown as they describe ways to optimize a hit once found.

In either case, the whole paragraph has now been removed as Step 1 has now been included in the results.

L347: streak-seeding?

Yes – streak-seeding was meant – the sentence has now been removed to reduce the results size.

L353 Diaylzed x2000? Performed 2000 times? Or against a 2000x larger volume?

The sentence has now been clarified.

L360-361: SPG? Please explain

This sentence has been removed as per the cutting down of the results. Crystals grown in SPG buffer are still highlighted in the Figure 4 (previous Figure 3), and the buffer components are given in the legend.

Figure3: indicate the compositions of PACT solutions A4, A5, C10, D10

The figure has now been reduced to two panels – with A4 and C10. The composition of each is given on the figure.

L368: 'although a credible micro-crystallization starting point, are, perhaps, not the ideal.' What does it mean?

This sentence has now been removed, as per the cutting down of the results. Essentially the original sentence was trying to say, perhaps poorly, that the long thin crystals produced from drops like A4 in the PACT screen could be optimized into micro-crystals. But they are perhaps not ideal in terms of their morphology.

L373: produced single crystals?

We have left the box-like description in, as we thinking this is a valid point when thinking about ideal crystals for serial experiments.

L376: morphology and diffraction quality were the principle criteria?

This sentence has been removed as per the cutting down of the results.

L398/ Fig11: Mg => Mg2+

Both changed – Fig. 11 is now

L410/Fig4,6: TRIS => TRIS-HCl

All Tris references have been changed to TRIS-HCl.

L440/Fig5 'can influence drop starting point'. Explain what dots and arrows correspond to. If a dot represents a starting point, then one would expect a second dot at the end of the arrow (representing a modification). By the way, one would hardly see any difference in batch among all dots shown in Fig.5A, because they are all in undersaturated conditions. The logic of this part is totally unclear...

We can see the reviewer point. The phase diagram on the batch location was not intended to mean anything specifically, just set the scene for the vectors or optimizations. The goal of the figure was to show how changes in protein or precipitant concentrations, or drop ratios, would change the starting point on the diagram. And how combinations of the changes could be used to move toward batch conditions. We do see that it is confusing. We have amended the figure and use the example in B, to show how such an optimization could be performed.

L449: 'precipitant' => precipitate

Again.. mea cupla..

L457: 'as can be seen...' What are we supposed to see in Figure 7? It does not tell anything about crystal sizes...
We see what you are trying to say, but the figure does not try to say something specific about crystal size. The idea is to show the reader what such an increased size, hanging drop might look like. Scale bars are given to show the approximate size of the visible crystals. The part in the text the reviewer is referring to, is actually saying that crystals can be observed. Considering that no crystals are observed when moving from 12 µL to 200 µL. We felt it was important to show what this step looked like.

L497/L500: again seeds do not influence 'the amount of nucleation' but the number of crystals
Again mea culpa -

L508: 'the batch crystallization was quenched' please explain in details using crystallogenesis concepts (solubility, supersaturation, kinetics of growth...) what you mean by quenching and give a clear protocol: L509 'using the precipitant solution', L634 'by the addition of crystallization buffer'. What is actually added and what is the effect on the system?
A complete description of the final crystallization condition of endothiapepsin has now been added to the protocol, a larger description of quenching and how this was approached is given in the results. This is a case where a precise phase diagram would be an advantage as the precise location of the solubility line would be known. This point has been added regarding the down sides of this method.

Discussion:

the discussion about morphology optimizing should be removed.
Due to the changes in the protocol this part has been left.

L672: 'using the nucleation time' How do you assess this time? Without DLS measurements, you can hardly observe a nucleation event, only its consequence, which is the growth of a crystal. Please explain
This sentence has been removed as the discussion has been rewritten, but the reviewer is correct that – crystal growth – rather than optimization time was meant.

L677:'The protocol attempts to discover this possibility by continuously advising to check the diffraction quality of the crystals produced at each stage of the optimization. ' Excellent advice for somebody living next to a synchrotron facility. What would you suggest to a crystal grower having only beamtime access every 3-6 months?
This point has now been added as a limitation to the protocol. We have thought about this point. It may be very possible to create a perfect sample for an XFEL without testing its diffraction. However, we would strongly advise against this.

L681: 'The creation of a phase diagram in Step 2, although not strictly necessary,' Really? But I thought it was the central point of this manuscript... I'm afraid you've lost the reader...

This sentence has been removed. The feeling behind it, was to honestly look at the protocol and say that, some of it may very well be unnecessary – many XFEL sample, past, present and future will be created not following it. But we appreciate the reviewers view that it statements like this dampen the impact.

L688/690/692: 'hope, hopefully' hope has no place in a scientific paper. This paragraph gives the impression that you have no confidence in the protocol you have just described. Please make this conclusion more enthusiastic and convincing!
We agree. These have now been removed.

Supplementary worksheet

Click here to access/download;Supplemental File (Figures, Permissions, etc.);phase-diagram-generator.xlsx

↓ insert protein name and concentration units

[protein] (mg/mL)	Row	Drop	5.00	8.18	11.36
			1	2	3
100.0	A	no-seed	20	20	40
		seed	20	15	30
87.5	B	no-seed	5	10	20
		seed	26	25	25
75.0	C	no-seed		5	10
		seed	15	25	20
62.5	D	no-seed			10
		seed	10	15	20
50.0	E	no-seed			1
		seed	5	10	20
37.5	F	no-seed			
		seed	3	10	3
25.0	G	no-seed			
		seed		3	10
12.5	H	no-seed			

12.5	seed

↑ insert protein drop concentrations here

Automatic input generated for plot

1				
x	y	no-seed	seed	x
5	100	20	20	8.18
5	87.5	5	26	8.18
5	75	0	15	8.18
5	62.5	0	10	8.18
5	50	0	5	8.18
5	37.5	0	3	8.18
5	25	0	0	8.18
5	12.5	0	0	8.18

7				
x	y	no-seed	seed	x
24.09	100	90	70	27.27
24.09	87.5	75	50	27.27
24.09	75	35	40	27.27
24.09	62.5	40	30	27.27
24.09	50	15	25	27.27

24.09	37.5	15	25	27.27
24.09	25	5	20	27.27
24.09	12.5	1	15	27.27

[precipitant] [% (w/v)]									← insert p
14.55	17.73	20.91	24.09	27.27	30.46	33.64	36.82	40.00	← insert p
4	5	6	7	8	9	10	11	12	
30	70	80	90	100					
20	40	50	70	100					
25	25	50	75						
40	60	50	50	100					
10	25	20	35	30					
25	25	35	40	60	100				
10	20	30	40	20					
25	25	20	30	30	10				
5	10	10	15	10					
15	20	20	25	20	25				
2	10	10	15	3					
20	25	50	25	30	20	20			
	2	3	5	10	2				
10	20	15	20	25	30	40	100		
			1			1	1	2	

5	10	10	15	25	3	5	2	3

2			3				4	
y	no-seed	seed	x	y	no-seed	seed	x	y	no-seed
100	20	15	11.36	100	40	30	14.55	100	30
87.5	10	25	11.36	87.5	20	25	14.55	87.5	25
75	5	25	11.36	75	10	20	14.55	75	10
62.5	0	15	11.36	62.5	10	20	14.55	62.5	10
50	0	10	11.36	50	1	20	14.55	50	5
37.5	0	10	11.36	37.5	0	3	14.55	37.5	2
25	0	3	11.36	25	0	10	14.55	25	0
12.5	0	0	11.36	12.5	0	0	14.55	12.5	0

8			9				10	
y	no-seed	seed	x	y	no-seed	seed	x	y	no-seed
100	100	100	30.46	100	0	0	33.64	100	0
87.5	0	100	30.46	87.5	0	0	33.64	87.5	0
75	30	60	30.46	75	0	100	33.64	75	0
62.5	20	30	30.46	62.5	0	10	33.64	62.5	0
50	10	20	30.46	50	0	25	33.64	50	0

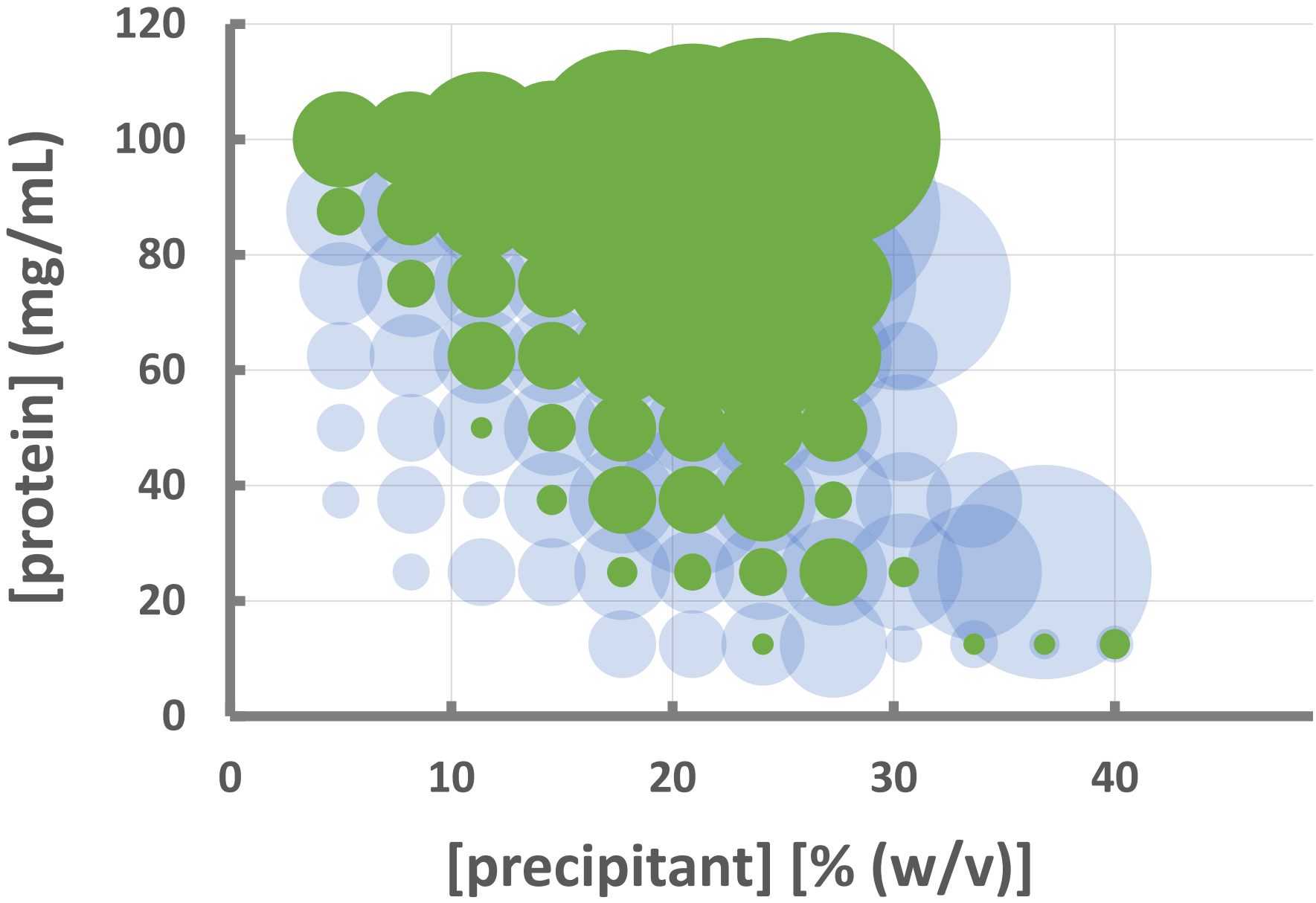
37.5	3	30	30.46	37.5	0	20	33.64	37.5	0
25	10	25	30.46	25	2	30	33.64	25	0
12.5	0	25	30.46	12.5	0	3	33.64	12.5	1

›recipitant name and concentration units
›recipitant drop concentrations here

	5				6			
seed	x	y	no-seed	seed	x	y	no-seed	seed
30	17.73	100	70	40	20.91	100	80	50
25	17.73	87.5	25	60	20.91	87.5	50	50
20	17.73	75	25	25	20.91	75	20	35
20	17.73	62.5	20	25	20.91	62.5	30	20
20	17.73	50	10	20	20.91	50	10	20
20	17.73	37.5	10	25	20.91	37.5	10	50
10	17.73	25	2	20	20.91	25	3	15
0	17.73	12.5	0	10	20.91	12.5	0	10

	11				12			
seed	x	y	no-seed	seed	x	y	no-seed	seed
0	36.82	100	0	0	40.00	100	0	0
0	36.82	87.5	0	0	40.00	87.5	0	0
0	36.82	75	0	0	40.00	75	0	0
0	36.82	62.5	0	0	40.00	62.5	0	0
0	36.82	50	0	0	40.00	50	0	0

20	36.82	37.5	0	0	40.00	37.5	0	0
40	36.82	25	0	100	40.00	25	0	0
5	36.82	12.5	1	2	40.00	12.5	2	3


120
100
80
60
40
20
0
[protein] (mg/mL)
0
10
20
30
40
[precipitant] [% (w/v)]