

The detection and subsequent volume optimization of biological nanocrystals

Joseph R. Luft,^{1,2,a)} Jennifer R. Wolfley,¹ Eleanor Cook Franks,¹
Angela M. Lauricella,¹ Ellen J. Gualtieri,³ Edward H. Snell,^{1,2} Rong Xiao,⁴
John K. Everett,⁴ and Gaetano T. Montelione^{4,5}

¹Hauptman-Woodward Medical Research Institute, 700 Ellicott Street, Buffalo,
New York 14203, USA

²Department of Structural Biology, State University of New York at Buffalo,
700 Ellicott Street, Buffalo, New York 14203, USA

³Formulatrix, Inc., 10 DeAngelo Drive, Bedford, Massachusetts 01730, USA

⁴Department of Molecular Biology and Biochemistry, Center for Advanced Biotechnology
and Medicine, and Northeast Structural Genomics Consortium, Rutgers, The State
University of New Jersey, 679 Hoes Lane, Piscataway, New Jersey 08854-8021, USA

⁵Department of Biochemistry and Molecular Biology, Robert Wood Johnson Medical
School, Rutgers, The State University of New Jersey, 679 Hoes Lane West, Piscataway,
New Jersey 08854-8021, USA

(Received 10 March 2015; accepted 5 May 2015; published online 15 May 2015)

Identifying and then optimizing initial crystallization conditions is a prerequisite for macromolecular structure determination by crystallography. Improved technologies enable data collection on crystals that are difficult if not impossible to detect using visible imaging. The application of second-order nonlinear imaging of chiral crystals and ultraviolet two-photon excited fluorescence detection is shown to be applicable in a high-throughput manner to rapidly verify the presence of nanocrystals in crystallization screening conditions. It is noted that the nanocrystals are rarely seen without also producing microcrystals from other chemical conditions. A crystal volume optimization method is described and associated with a phase diagram for crystallization. © 2015 Author(s). All article content, except where otherwise noted, is licensed under a Creative Commons Attribution 3.0 Unported License. [<http://dx.doi.org/10.1063/1.4921199>]

I. INTRODUCTION

The advent of X-ray free electron laser (XFEL) sources and the development of serial crystallography techniques revealed that useable X-ray data can be obtained from crystals with dimensions as small as hundreds of nanometers,¹ colloquially termed nanocrystals. Similar results have also been seen from nanocrystals mounted in fixed assemblies^{2,3} and by pushing synchrotron beam line technologies to smaller and smaller beam footprints, having achieved serial crystallography on crystals having an average volume of only $9 \mu\text{m}^3$.⁴ In the general case for these studies, a key question becomes if nanocrystals are a more common result than larger crystals suitable for typical synchrotron-based data collection (which we will term microcrystals measured in tens of micrometers and up), and if there are methods suited to growing nanocrystals that are not so successful for microcrystals. In the case of time-resolved studies, a nanocrystal is an advantageous result, in which a change in molecular conformation associated with a reaction, whether it be triggered by light, chemically, or some other process, can occur over all the individual macromolecules in the diffraction footprint rapidly.

The technical questions surrounding this are (1) how do we identify nanocrystals as they grow, (2) how often do they occur compared to larger samples more suitable for today's

^{a)} Author to whom correspondence should be addressed. Electronic mail: luft@hwi.buffalo.edu



synchrotron, and (3) how do we produce nanocrystals. Visual microscopy can be used to identify crystals, but has a size limit of a few microns, which limits addressing these questions. Other techniques including light-scattering and electron microscopy are available and are discussed in detail elsewhere.⁵ The technical questions we have posed are general; for specific samples they can be answered in detail, but for a generally applicable answer a large number of different macromolecular samples need to be studied. We have adopted Second Order Non-Linear Imaging of Chiral Crystals (SONICCs)⁶ and UV-Two Photon Excited Fluorescence (UV-TPEF)⁷ as high-throughput, integrated methods to identify crystals $<1 \mu\text{m}$ in size (catching but bordering our definition of nanocrystals), and image all samples that come through the high-throughput crystallization screening laboratory operating within the Hauptman-Woodward Institute.⁸ In this manner, we have studied the crystallization behavior of a large number of diverse macromolecules and focused on using the same technology on a small number of samples, in more detail. In this paper, we present these results, we demonstrate an effective method to produce microcrystals from nanocrystals and discuss the production of nanocrystals based on screening and optimization results that we obtain as this production relates to a solubility phase diagram.

II. EXPERIMENTAL

A. Samples and crystallization conditions

To study the frequency of formation of nanocrystals and subsequent optimization to macrocrystals, a set of 60 proteins were selected ranging in molecular weights from 9.8 kDa to 127.2 kDa, and isoelectric points $4.6 \leq \text{pI} \leq 9.8$, from 25 species (including 19 proteins from *Homo sapiens*). These included several proteins of high biological interest. Two samples were truncations of a protein complex of the Heat shock proteins Hsp83 with the Hsp90 co-chaperone. A CLANS sequence similarity matrix⁹ is presented in Figure 1. The supplementary material¹⁰ is a table containing the crystallization results and the following data for each of the 60 NESG proteins, UniProt id, Percent UniProt sequence covered, Gene name, MW, pI, AGF-MALS results, number of residues, percentage of residues predicted to be disordered (using DISOPRED v3.15 and the latest version of the UniRef90 database as a reference database), and the corresponding amino acid sequences. The proteins used were produced using standard protocols of the Northeast Structural Genomics Consortium.¹¹ All sequences and data are calculated for the protein construct used for crystallization screening. Solubility, Maltose-binding protein (MBP) and purification (6xHis) tags required for production were removed for

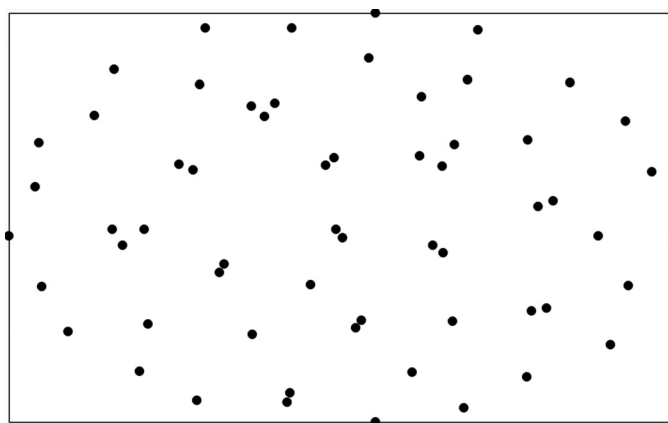


FIG. 1. CLANS visualization of a sequence similarity matrix created from an ALL vs ALL BLAST analysis. Each of the 60 experimental sequences is represented by a dot. Sequence dots were originally distributed randomly and were allowed to evolve through 500 iterations via their pairwise alignment score weights to the state shown in the figure. The distance between sequence dots is a function of their pairwise BLAST alignment scores resulting in similar sequences clustering near one another.

most of these samples prior to crystallization screening. Proteins were flash frozen in 50 μl aliquots and shipped on dry ice. When received at HWI, the samples were rapidly thawed¹² and used for crystallization screening.

Initial crystallization conditions were identified using the Hauptman-Woodward Institute's High-throughput Crystallization Screening Laboratory (HTSlab).¹³ The crystallization method, microbatch-under-oil,¹⁴ as utilized in the HTSlab requires 450 μl of protein to screen 1536 chemical cocktails by combining 200 nl of protein solution with 200 nl of 1536 chemically diverse cocktail solutions under white mineral oil (EMD Millipore, cat #PX0045-3, Billerica, MA). The oil-layer limits evaporative water-loss during the time-course of the experiments, typically, six weeks. The batch approach is advantageous to scaling up results and reproducibility as batch experiments are relatively static with regard to chemical concentrations, when compared to dynamic crystallization approaches such as vapor diffusion.

Two sets of 1536 chemical cocktails were used for this initial crystallization assay. The first set, designed to identify initial crystallization conditions of soluble proteins, is comprised of 688 in-house cocktails based on an incomplete-factorial sampling of salts, buffers, and polymers and 848 chemically complementary cocktails from Hampton Research crystallization screens (Aliso Viejo, CA); cocktails are annually reviewed and refined.⁸ A second set of 1536 cocktails designed to crystallize membrane protein-detergent complexes, grid-samples chemical space near detergent phase boundaries,¹⁵ where the crystallization is more likely to occur.¹⁶ The set of 60 proteins used in this study was not random. There were two sets of proteins selected. The first set of 39 proteins produced visible microcrystals, from the 1536 screen. The second set of 21 proteins did not produce any visible microcrystals from the 1536 screen.

B. Crystallization and imaging

The 1536 well crystallization screening plates were setup with the proteins described at room-temperature, then immediately incubated for one week at 4 °C. They were then stored at 22 °C to obtain information on the impact of temperature-based effects on crystallization, a factor that can have a significant influence.^{17,18} Crystallization screening results were imaged at 22 °C with in-house imaging tables⁸ over the course of six weeks, with the exception of the one week images, which were recorded at 14 °C, immediately following 4 °C incubation. Additionally, a Formulatrix Rock Imager 1000 with SONICC (Bedford, MA) was used to record three types of image data, Visible, Second Harmonic Generation (SHG), and UV-TPEF following in-house imaging at 1 and 4 weeks; this data was collected at 22 °C.

Visual imaging, e.g., using, for example, the laboratory microscope or automated versions thereof, is a familiar aspect of the crystallization laboratory. SHG and UV-TPEF imaging methods are less common. The SHG and UV-TPEF imaging modalities rely on nonlinear optical processes to provide information on the crystallinity of the sample and to discern between protein and salt/small molecule crystals. SHG is a frequency-doubling effect that is coherent so only arises in the presence of non-centrosymmetric crystals. All protein crystals form chiral crystals that are SHG allowed, but proteins in solution or aggregated protein molecules do not generate SHG. In the Formulatrix system, a 200 fs, 1064 nm laser is utilized to image the sample via a raster pattern and the SHG at 532 nm is detected with a photomultiplier tube to generate high-contrast images. The intensity of SHG generated from protein crystals depends on many factors including the space group and hyper-polarizability of the protein molecule, with some samples not generating sufficient signal to be detected. Also, some chiral salt or small molecule crystals will generate SHG leading to false positives.¹⁹ These false negatives and positives can be reduced with the use of UV-TPEF, the third imaging technique. In this mode, the sample is imaged with 532 nm light, with the frequency doubled excitation at 266 nm, which excites aromatic amino acids including tryptophan, tyrosine, and phenylalanine. The multi-photon excited fluorescence from 350 to 400 nm is then detected creating images depicting where there are high concentrations of protein allowing the scientist to determine if their crystal is protein or not.⁷ For SHG and UV-TPEF data acquisition, 5 slices through the drop, spaced 100 μm apart (called z-slices), are integrated together to produce a single, composite

image that can be used for reviewing outcomes. Together, the SHG and UV-TPEF nonlinear imaging modes detect both large and small ($<1\ \mu\text{m}$) crystals while differentiating between protein and salts.

C. Using imaging to guide optimization

From the set of 60 samples, a subset of 14 soluble protein samples, ranging from 9.9 kDa to 48.9 kDa in molecular weight, were selected to study the relationship between SHG and UV-TPEF results and subsequent optimization of the crystallization conditions. These samples are described in Table I. Eleven of these 14 proteins, A-K, produced one or more visible crystals from the initial 1536 screen. Also identified in these cases were multiple precipitates, from chemically distinct cocktails, some were verified to be nanocrystals with a positive SHG signal, and protein with a positive UV-TPEF signal. In these cases, based solely upon the visible images, the precipitates were not outcomes that would have been recommended for optimization; however, the analysis using SHG and UV-TPEF clearly identified the outcomes as crystalline. The other three samples, L-N used as a control, produced no visible crystals from the initial 1536 screen. Only one of these, sample N, produced outcomes that had both a positive, but faint, SHG signal and a positive UV-TPEF signal. Using leftover protein solution from the screening experiments, this set of 14 proteins underwent a second freeze-thaw cycle prior to setting up optimization experiments.

The Visible, SHG, and UV-TPEF images for each of the 14 proteins A-N were reviewed to identify 5–8 cocktails that produced precipitates, where chemical and physical conditions provided signal, these precipitates were verified to be nanocrystals, using SHG, and protein, using UV-TPEF. The chemical diversity of the cocktails that produced nanocrystalline precipitates was used as a selection criterion; however, in many cases, nanocrystalline precipitates were only observed from a limited range of chemical conditions.

Individual optimization experiments were setup following protein and cocktail combinations that are described in Sec. III and were conducted in high-throughput, using automated liquid-handling systems following a modified version of a previously published protocol.²⁰ A simple batch protocol was used for optimization by varying the drop volume ratio (DVR) of the cocktail and protein solutions.^{20,21} The volume of the protein and cocktail solutions is varied incrementally for 16 experiments. Figure 2 graphically displays the percentage of protein and cocktail used for each experiment, along with the 4×4 experiment array used for the

TABLE I. Subset of 14 samples used for optimization studies of initial crystallization “hits.”

Protein code	MW (kDa)	Concentrations (mg/ml)
Visible crystals in screen		
A	10.1	9.5
B	13.7	10.1
C	18.7	9.0
D	29.1	7.7
E	10.2	6.8
F	9.9	8.0
G	24.5	7.2
H	12.5	13.4
I	17.0	10.8
J	15.0	4.4
K	15.9	6.0
No visible crystals in screen		
L	48.9	40.0
M	40.2	11.4
N	30.6	10.4

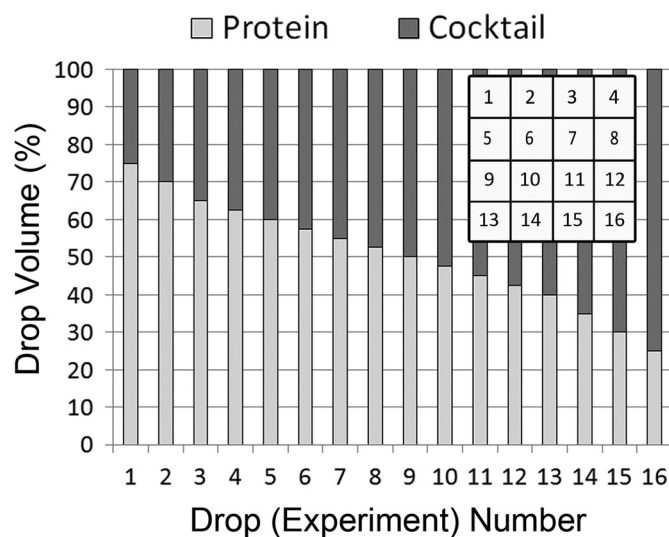


FIG. 2. The drop volume ratio optimization experiments. A stacked bar graph displaying the percentage of protein and cocktail solutions in each of the 16 experiment drops. The inset shows the experiment drop array as it is displayed in figures showing the optimization results.

images presented later in Sec. III. Each set of 16 DVR experiments requires $5 \mu\text{l}$ of protein solution as tabulated in Table II. The cocktail solutions are used at their original concentrations. Two series of DVR experiments were setup, one with the protein at its initial concentration, used during the 1536 screening experiments, and the other at $1/2$ of the initial concentration by diluting the protein stock solution with an equal volume of sample buffer 100 mM sodium chloride, 5 mM Dithiothreitol (DTT), 0.02% (w/v) sodium azide, 10 mM Tris-HCl, pH 7.5 for A-K, M-N. For protein L, the sample buffer was 100 mM sodium chloride, 20 mM Tris, pH 7.4. All of the DVR experiments had a fixed final volume of 600 nl (protein plus cocktail solution) and all were incubated and imaged at 22°C .

TABLE II. The volumes of protein and cocktail solutions used to setup the DVR optimization experiments for conditions that produced nanocrystals from the initial crystallization screen. Note that $\sim 5 \mu\text{l}$ of protein is used for one set of 16 experiments.

Ratio	Protein volume (nl)	Cocktail volume (nl)
1	450	150
2	420	180
3	390	210
4	375	225
5	360	240
6	345	255
7	330	270
8	315	285
9	300	300
10	285	315
11	270	330
12	255	345
13	240	360
14	210	390
15	180	420
16	150	450
TOTAL	4875	4725

The DVR optimization experiments were imaged weekly for six weeks using the Formulatrix Rock Imager 1000 with SONICC at 22 °C, to record Visible, SHG, and UV-TPEF data. Image data were reviewed weekly. All of the images and data reported in this manuscript are based on the four week results.

III. RESULTS

A. SHG and UV-TPEF imaging of large-scale crystallization screening experiments

Figure 3 shows an example of visible, SHG, and UV-TPEF fluorescence images from a result that, from the visible image alone, would likely have been disregarded for subsequent optimization. The theoretical crystal size detection limit for SHG is reported to be 90 nm.²² For the set of 39 proteins that produced visible microcrystals from the crystallization screening 18 of these 39 proteins also produced nanocrystals that were verified by SHG and UV-TPEF from other cocktails in the screen.

For the 21 samples that did not produce microcrystals, which could be detected in the visible images, we identified only one case, sample N, where protein nanocrystals were verified by SHG and UV-TPEF. Sample L was a colored solution that, based on the preponderance of SHG signal from nearly all of the experiments, was likely a false-positive SHG signal.

B. Optimization

For the subset of 14 samples that underwent optimization trials all 11 of the proteins that produced visible microcrystals from the initial crystallization screen also produced visible

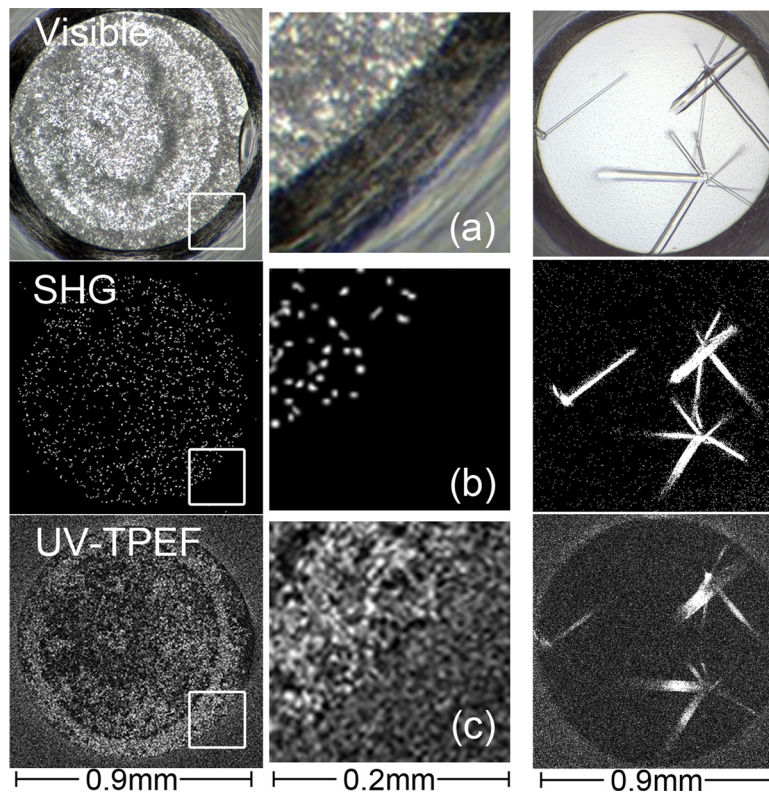


FIG. 3. Examples of precipitates and large crystals verified as protein crystals with Visible, SHG, and UV-TPEF imaging. (a) An enlarged inset showing the visible image of D.2 at four weeks, (b) SHG signal from the experiment shows it is crystalline, and (c) UV-TPEF image shows that it is protein. Far right column shows an example of large crystals of Protein A, produced directly from the initial screen (cocktail 40% (w/v) PEG 1000, 0.1 M ammonium phosphate monobasic, 0.1 M sodium HEPES, pH 7.5) with visible, SHG, and UV-TPEF images.

TABLE III. Cocktails used for optimization experiments for each of the 14 proteins, with an accompanying DVR (optimization) experiment code, a column to indicate the presence or absence of SHG signal from the precipitates selected from the initial crystallization screening experiments, and a column showing which optimization experiments produced crystals identifiable by visual microscopy.

Samples with microcrystals identified in other conditions			
DVR	SHG	Initial crystallization conditions optimized	Microcrystals
Protein A			
A.1	+	40% (w/v) PEG 8000, 0.1 M potassium thiocyanate; 0.1 M sodium acetate, pH 5.0	+
A.2	+	40% (w/v) PEG 8000, 0.1 M sodium nitrate; 0.1 M sodium citrate, pH 4.2	+
A.3	+	HR ^a PEGRx HT (C4); 25% (w/v) PEG 3350, 0.1 M citric acid, pH 3.5	+
A.4	+	HR PEGRx HT (F4); 20% (w/v) PEG 1500, 4% (v/v) MPD, 0.1 M citric acid, pH 3.5	+
A.5	+	HR PEGRx HT (G12); 22% (w/v) PEG 6000, 10% (v/v) 2-propanol, 0.1 M sodium acetate trihydrate, pH 4.0	+
Total success rate			100%
Protein B			
B.1	+	HR Ionic Liquid (11); 27% (w/v) PEG 3350, 5% (w/v) 1-butyl-3-methylimidazolium dicyanamide, 0.09 M MES, pH 5.8	+
B.2	+	24% (w/v) PEG 20 000, 0.1 M potassium thiocyanate, 0.1 M sodium acetate, pH 5.0	+
B.3	+	20% (w/v) PEG 8000, 0.1 M ammonium thiocyanate, 0.1 M sodium acetate, pH 5.0	+
B.4	+	40% (w/v) PEG 8000, 0.1 M ammonium thiocyanate, 0.1 M MES, pH 6.0	+
B.5	+	40% (w/v) PEG 4000, 0.1 M ammonium thiocyanate, 0.1 M MES, pH 6.0	+
Total success rate			100%
Protein C			
C.1	+	HR Grid Screen AS (D6); 3.0 M AS, 0.1 M bicine, pH 9.0	+
C.2	+	40% (w/v) PEG 8000, 0.1 M AS, 0.1 M TAPS, pH 9.0	+
C.3	+	40% (w/v) PEG 8000, 0.1 M lithium sulfate monohydrate, 0.1 M Bis-tris propane, pH 7.0	+
C.4	+	40% (w/v) PEG 1000, 0.1 M sodium chloride, 0.1 M Tris, pH 8.0	+
C.5	+	HR SaltRx HT (E10); 1.8 M sodium phosphate monobasic monohydrate, potassium phosphate dibasic, pH 8.2	+
Total success rate			100%
Protein D			
D.1	+	20% (w/v) PEG 8000, 0.1 M ammonium phosphate-dibasic, 0.1 M sodium acetate, pH 5.0	-
D.2	+	20% (w/v) PEG 4000, 0.1 M lithium sulfate monohydrate, 0.1 M sodium acetate, pH 5.0	+
D.3	+	20% (w/v) PEG 400, 0.1 M magnesium sulfate heptahydrate, 0.1 M sodium acetate, pH 5.0	+
D.4	+	HR PEG/ion HT (E12); 20% (w/v) PEG 3350, 8% (v/v) tacsimate, pH 5.0	+
D.5	+	HR crystal screen HT (E2); 0.5 M sodium chloride, 0.01 M magnesium chloride hexahydrate, 0.01 M hexadecyltrimethylammonium bromide	-
Total success rate			60%
Protein E			
E.1	+	80% (v/v) PEG 400, 0.1 M ammonium bromide, 0.1 M Bis-tris propane, pH 7.0	+
E.2	+	80% (v/v) PEG 400, 0.1 M ammonium chloride, 0.1 M tris, pH 8.0	+
E.3	+	80% (v/v) PEG 400, 0.1 M magnesium acetate tetrahydrate, 0.1 M Bis-tris propane, pH 7.0	+

TABLE III. (Continued.)

Samples with microcrystals identified in other conditions			
DVR	SHG	Initial crystallization conditions optimized	Microcrystals
E.4	+	80% (v/v) PEG 400, 0.1 M magnesium chloride hexahydrate, 0.1 M Bis-tris propane, pH 7.0	+
E.5	+	80% (v/v) PEG 400, 0.1 M potassium nitrate, 0.1 M Bis-tris propane, pH 7.0	+
E.6	–	HR Ionic Liquid 18; 27% (w/v) PEG 3350, 5% (w/v) 1-butyl-2,3-dimethylimidazolium tetrafluoroborate, 0.09 M MES, pH 5.8	+
E.7	–	40% (v/v) PEG 8000, 0.1 M potassium acetate, 0.1 M sodium acetate, pH 5.0	+
E.8	+	40% (v/v) PEG 400, 0.1 M magnesium nitrate hexahydrate, 0.1 M HEPES, pH 7.5	+
Total success rate			100%
Protein F			
F.1	+	HR Slice pH (A12); 15% (w/v) PEG 3350, 0.5 M sodium acetate trihydrate, pH 4.6	+
F.2	+	HR Ionic Liquid (1); 27% (w/v) PEG 3350, 5% (w/v) tetraethylammonium bromide, 0.09 M MES, pH 5.8	–
F.3	+	HR crystal screen cryo (30); 25.5% (w/v) PEG 8000, 15% (v/v) glycerol, 0.17 M AS	–
F.4	+	12% (w/v) PEG 20 000, 0.1 M potassium nitrate, 0.1 M MES, pH 6.0	–
F.5	+	20% (w/v) PEG 1000, 0.1 M lithium sulfate monohydrate, 0.1 M MES, pH 6.0	–
F.6	+	40% (w/v) PEG 1000, 0.1 M ammonium phosphate monobasic, 0.1 M sodium citrate, pH 4.2	–
F.7	+	40% (v/v) PEG 400, 0.1 M sodium phosphate–monobasic, 0.1 M MES, pH 6.0	–
F.8	+	HR PEGRx HT (C11); 20% (w/v) PEG 5000 MME, 0.1 M Bis-tris, pH 6.5	+
Total success rate			25%
Protein G			
G.1	+	HR Grid Screen AS (A1); 0.8 M AS, 0.1 M citric acid, pH 4.0	+
G.2	+	HR Grid Screen AS (B5); 1.6 M AS, 0.1 M Tris, pH 8.0	+
G.3	+	HR crystal screen cryo (6); 24% (w/v) PEG 4000, 20% (v/v) glycerol, 0.16 M magnesium chloride hexahydrate, 0.08 M Tris hydrochloride, pH 8.5	+
G.4	+	24% (w/v) PEG 20 000, 0.1 M potassium chloride, 0.1 M sodium citrate, pH 4.2	+
G.5	+	20% (w/v) PEG 4000, 0.1 M potassium phosphate-monobasic, 0.1 M sodium citrate, pH 4.2	+
G.6	+	20% (w/v) PEG 4000, 0.1 M potassium thiocyanate, 0.1 M TAPS, pH 9.0	–
G.7	+	HR crystal screen cryo (20); 20% (w/v) PEG 4000, 20% (v/v) glycerol, 0.16 M AS, 0.08 M sodium acetate trihydrate, pH 4.6	+
G.8	+	HR PEGRx HT (C7); 10% (w/v) PEG 4000, 0.1 M sodium acetate trihydrate, pH 4.0	+
Total success rate			87%
Protein H			
H.1	+	HR Slice pH (C8); 15% (w/v) PEG 3350, 0.5 M MES, pH 6.2	+
H.2	+	HR Ionic Liquid (11); 27% (w/v) PEG 3350, 5% (w/v) 1-butyl-3-methylimidazolium dicyanamide, 0.09 M HEPES, pH 6.8	+
H.3	+	24% (w/v) PEG 20 000, 0.1 M potassium thiocyanate, 0.1 M MES, pH 6.0	+
H.4	+	20% (w/v) PEG 8000, 0.1 M rubidium chloride, 0.1 M MES, pH 6.0	–
H.5	+	20% (w/v) PEG 4000, 0.1 M sodium nitrate, 0.1 M MES, pH 6.0	+

TABLE III. (Continued.)

Samples with microcrystals identified in other conditions			
DVR	SHG	Initial crystallization conditions optimized	Microcrystals
H.6	+	40% (w/v) PEG 4000, 0.1 M calcium chloride dihydrate, 0.1 M Bis-tris propane, pH 7.0	–
H.7	+	HR crystal screen HT (B6); 20% (w/v) PEG 8000, 0.2 M magnesium acetate tetrahydrate, 0.1 M sodium cacodylate trihydrate, pH 6.5	+
H.8	+	HR crystal screen HT (C4); 30% (w/v) PEG 8000, 0.2 M sodium acetate trihydrate, 0.1 M sodium cocodylate trihydrate, pH 6.5	–
Total success rate			71%
Protein I			
I.1	+	40% (w/v) PEG 8000, 0.1 M sodium chloride, 0.1 M TAPS, pH 9.0	+
I.2	+	40% (w/v) PEG 8000, 0.1 M sodium thiosulfate pentahydrate, 0.1 M CAPS, pH 10.0	–
I.3	+	40% (w/v) PEG 8000, 0.1 M sodium thiosulfate pentahydrate, 0.1 M TAPS, pH 9.0	+
I.4	+	40% (w/v) PEG 4000, 0.1 M magnesium chloride hexahydrate, 0.1 M HEPES, pH 7.5	+
I.5	+	20% (w/v) PEG 1000, 0.1 M magnesium chloride hexahydrate, 0.1 M Tris, pH 8.0	+
I.6	+	40% (w/v) PEG 1000, 0.1 M potassium bromide, 0.1 M Tris, pH 8.0	+
I.7	+	80% (w/v) PEG 400, 0.1 M ammonium bromide, 0.1 M Bis-tris propane, pH 7.0	+
I.8	+	HR PEG/Ion HT (A5); 20% (w/v) PEG 3350, 0.2 M magnesium chloride hexahydrate, pH 5.9	+
Total success rate			87%
Protein J			
J.1	+	HR Ionic Liquid (3); 27% (w/v) PEG 3350, 5% (w/v) 2-hydroxyethylammonium formate, 0.09 M MES, pH 5.8	–
J.2	+	HR Ionic Liquid (12); 27% (w/v) PEG 3350, 5% (w/v) 1,3-dimethylimidazolium dimethyl phosphate, 0.09 M MES, pH 5.8	–
J.3	+	40% (w/v) PEG 8000, 0.1 M ammonium phosphate monobasic, 0.1 M MES, pH 6.0	–
J.4	+	20% (w/v) PEG 4000, 0.1 M potassium acetate, 0.1 M MES, pH 6.0	–
J.5	+	20% (w/v) PEG 1000, 0.1 M rubidium chloride, 0.1 M MES, pH 6.0	+
J.6	+	HR PEGRx HT (B11); 20% (w/v) PEG 2000 MME, 0.1 M MES monohydrate, pH 6.0	–
J.7	+	HR Index HT (F10); 25% (w/v) PEG 3350, 0.20 M sodium chloride, 0.1 M Bis-tris, pH 5.5	+
J.8	+	HR Index HT (G6); 25% (w/v) PEG 3350, 0.20 M ammonium acetate, 0.1 M Bis-tris, pH 5.5	–
Total success rate			25%
Protein K			
K.1	+	HR Ionic Liquid (23); 27% (w/v) PEG 3350, 5% (w/v) triisobutylmethylphosphonium tosylate, 0.09 M Bis-tris Propane, pH 8.8	+
K.2	+	HR Grid Screen AS (D5); 3.0 M AS, 0.1 M Tris, pH 8.0	–
K.3	+	40% (w/v) PEG 8000, 0.1 M calcium chloride dihydrate, 0.1 M MES, pH 6.0	+
K.4	+	40% (w/v) PEG 8000, 0.1 M potassium bromide, 0.1 M Tris, pH 8.0	–
K.5	+	20% (w/v) PEG 4000, 0.1 M manganese chloride tetrahydrate, 0.1 M Bis-tris propane, pH 7.0	–
K.6	+	20% (w/v) PEG 4000, 0.1 M manganese chloride tetrahydrate, 0.1 M MES, pH 6.0	+

TABLE III. (Continued.)

Samples with microcrystals identified in other conditions			
DVR	SHG	Initial crystallization conditions optimized	Microcrystals
K.7	+	20% (w/v) PEG 4000, 0.1 M potassium chloride, 0.1 M CAPS, pH 10.0	–
K.8	+	40% (w/v) PEG 1000, 0.1 M manganese sulfate monohydrate, 0.1 M MES, pH 6.0	+
Total success rate			50%
Samples with no microcrystals identified in other conditions			
DVR	SHG	Initial crystallization conditions optimized	Microcrystals
Protein L			
L.1	+	HR Ionic Liquid (15); 27% (w/v) PEG 3350, 5% (w/v) 1-n-BUTYL-3-methylimidazolium n-octylsulfate, 0.09 M Bis-tris propane, pH 8.8	–
L.2	+	20% (w/v) PEG 8000, 0.1 M sodium nitrate, 0.1 M sodium citrate, pH 4.2	–
L.3	+	40% (w/v) PEG 8000, 0.1 M lithium chloride, 0.1 M sodium acetate, pH 5.0	–
L.4	+	40% (w/v) PEG 8000, 0.1 M potassium phosphate monobasic, 0.1 M CAPS, pH 10.0	–
L.5	+	40% (w/v) PEG 8000, 0.1 M sodium phosphate monobasic, 0.1 M Tris, pH 8.0	–
Total success rate			0%
Protein M			
M.1	–	HR Grid Screen AS (C3); 2.4 M AS, 0.1 M MES monohydrate, pH 6.0	–
M.2	–	40w/v) PEG 4000, 0.1 M sodium nitrate, 0.1 M sodium acetate, pH 5.0	–
M.3	–	40% (v/v) PEG 400, 0.1 M potassium nitrate, 0.1 M HEPES, pH 7.5	–
M.4	–	HR crystal screen cryo (38); 10% (v/v) glycerol, 1.26 M sodium citrate tribasic dihydrate, 0.09 M HEPES-Na, pH 7.5	–
M.5	–	HR Index HT (B8); 1.4 M sodium citrate tribasic dihydrate, 0.1 M HEPES, pH 7.5	–
Total success rate			0%
Protein N			
N.1	+	HR Grid Screen AS (D3); 3.0 M AS, 0.1 M MES monohydrate, pH 6.0	–
N.2	+	12% (w/v) PEG 20 000, 0.1 M lithium chloride, 0.1 M TAPS, pH 9.0	–
N.3	+	20% (w/v) PEG 8000, 0.1 M ammonium phosphate dibasic, 0.1 M TAPS, pH 9.0	–
N.4	+	20% (w/v) PEG 8000, 0.1 M lithium bromide, 0.1 M HEPES, pH 7.5	–
N.5	+	20% (w/v) PEG 8000, 0.1 M potassium phosphate dibasic, 0.1 M HEPES, pH 7.5	–
N.6	+	(1:1) HR Silver Bullet Bio (D12) 0.16% (w/v) L-Homoserine, 0.16% (w/v) 4-hydroxy-L-proline, 0.16% (w/v) argininosuccinic acid disodium salt hydrate, 0.16% (w/v) cytidine, 0.16% (w/v) inosine, 0.16% (w/v) guanine, 0.02 M HEPES sodium pH 6.8 + crystallization reagent (D12) 25% (w/v) PEG3350, 0.1 M HEPES, pH 6.8	–
N.7	+	HR PEG/Ion HT (A2); 20% (w/v) PEG 3350, 0.2 M potassium fluoride, pH 7.3	–
Total success rate			0%

^aHR notates screens from Hampton Research; Hampton Research screens were used as purchased with the exception of the Ionic Liquids and Slice pH screens which were modified for in-house batch crystallization. The Ionic Liquids screen was modified with the addition of 0.09 M buffer and 27% (w/v) PEG 3350 buffer. The Slice pH screen was modified with the dilution of the buffer from its initial 1.0 M to 0.5 M concentration to accommodate the addition of 15% (w/v) PEG 3350 to promote supersaturation in the batch experiments. MPD = (+/-)-2-Methyl-2,4-pentanediol; AS = Ammonium sulfate.

crystals following DVR optimization of nanocrystals from the initial screen. The success of the DVR experiments is scored on a per cocktail basis. If visible crystals were observed following DVR optimization, then the optimization experiments for that cocktail were considered a success. The total number of cocktails used for DVR optimization, each at 16 drop volume ratios and at both full and $1/2$ protein concentrations, is shown in Table II. It must be noted that many of these cocktails were chemically similar, which could account for some, but not all of the high success rates. The screening outcomes at four weeks (Visible and SHG) and the highest visual quality crystal produced following DVR is shown for each of the 11 proteins in Figure 4. For three samples L, M, and N, where visible crystals were not identified during initial screening, there were no visible crystals produced following DVR optimization. This includes sample N, which produced outcomes validated as nanocrystals from initial screening with SHG and UV-TPEF signals.

While chemical diversity was an aim in choosing initial crystallization conditions for optimization, the diversity achieved varied. As an example of where diversity was not achieved,

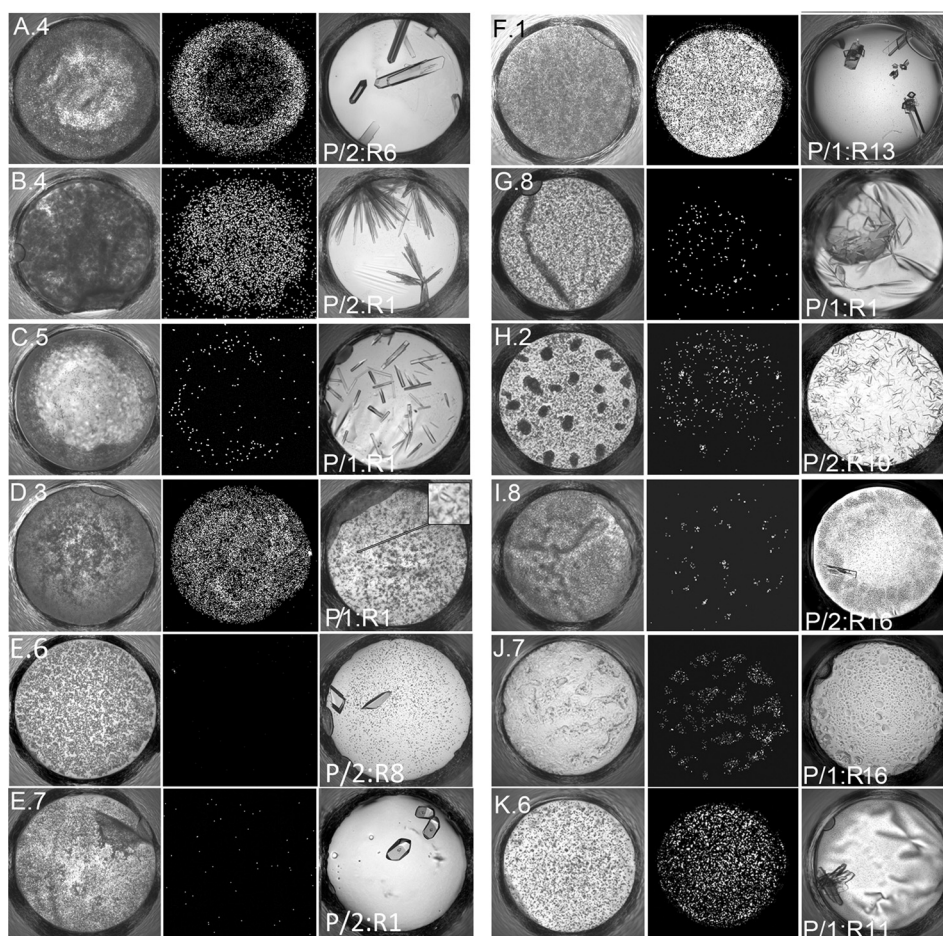


FIG. 4. A display of the 11 proteins (A–K) that produced visible crystals using DVR optimization, showing in each of the two columns the initial nanocrystals, Visible (left), and SHG (center) and the highest visual quality crystals (right) that were produced. The cocktail associated with these outcomes is indicated by a number (1–8) following the protein code; the corresponding cocktail's chemical information is located in Table III. The protein concentration used to produce the optimized crystal is indicated by P/1 for proteins at identical concentrations to the screening experiments, and P/2 for proteins at one half the concentrations used for screening by dilution of the protein solution with an equal volume of sample buffer. The volume ratio that produced the optimized crystal is indicated by R1–R16 and corresponds to the data found in Figure 2 and Table II. All of the initial precipitates had positive SHG signal (indicated by white signal) with the exceptions of experiment E.6 and E.7, where SHG signal was not observed from the precipitate. The well-diameter is 0.9 mm for all of these images.

the cocktails selected for Protein B are all PEG-based, which have small range of pH values 5.0–6.0 and four out of the five cocktails contain a thiocyanate salt.

A full-series of DVR experiments, at two protein concentrations, is shown for protein sample A in Figure 5. At 9.5 mg/ml, the DVR outcomes, although many appear to be precipitate visually, are verified as protein (positive UV-TPEF signal) and crystals (positive SHG signal). A visible crystal is observed at ratio 14 (210 nl protein to 390 nl cocktail), note that unlike the nanocrystals, while this larger crystal has a UV-TPEF signal, verifying that it is protein, it does not have an SHG signal. At 4.8 mg/ml, the DVR experiments produced large crystals easily discerned from visible images. The UV-TPEF verifies these outcomes to be protein. None of these larger crystals provides an SHG signal. The crystal outcomes vary considerably in their size and frequency. These trends generally correlate or incrementally change with the drop volume ratio. Oftentimes, there is a noticeable difference in crystal frequency and size when comparing the full and $1/2$ concentrated series of DVR experiments. Higher protein concentrations would

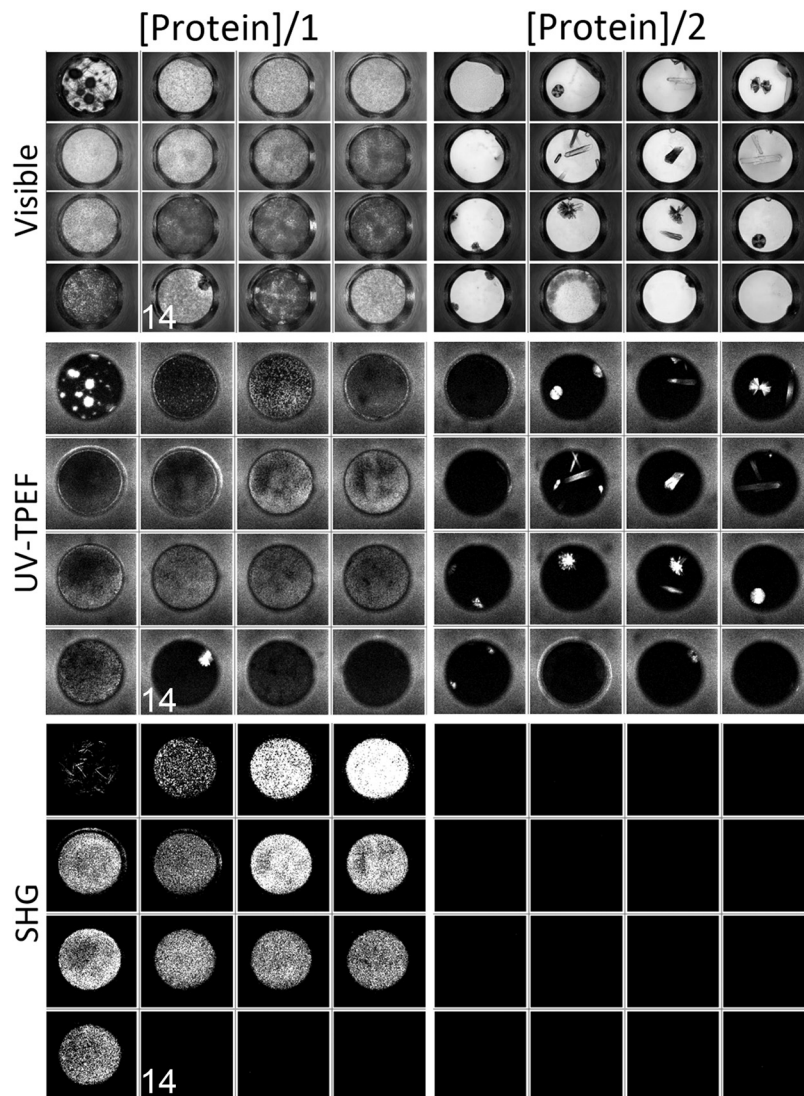


FIG. 5. The full set of DVR experiments for Protein A, Visible, UV-TPEF, and SHG at four weeks. Refer to Figure 2 to review the drop volumes used to make up the 4×4 experimental arrays. Note there are two sets of DVR experiments, column [Protein]/1 = 9.5 mg/ml protein concentration and column [Protein]/2 = 4.8 mg/ml. For each of the two sets of DVR experiments, an array of visible, UV-TPEF, and SHG imaging data are displayed. Note that ratio 14 of [Protein]/1 lacks an SHG signal, similar to the outcomes of the DVR experiments setup with [Protein]/2.

generally, but not always, correlate with higher levels of supersaturation and the subsequent production of a greater number of smaller crystals.

Figure 5 illustrates that SHG characterization can fail to detect crystals. While the majority of crystals provide SHG signal the SHG signal strength is, among other factors, dependent upon the symmetry and orientation of the crystal.²² If we look more closely at Figure 5, it is interesting to note a change from SHG positive (left bottom of Figure 5), possibly lower symmetry crystals produced at higher protein concentrations, to SHG negative (right bottom of Figure 5), potentially higher symmetry crystals produced at lower protein concentrations. The crystal morphology that lacks SHG signal is present for the entire set of 16 DVR experiments setup using an initial protein concentration of 4.8 mg/ml. This crystal morphology and subsequent lack of SHG signal was observed for one of the DVR experiments setup using the 9.5 mg/ml solution of protein sample A (Ratio 14, 35% protein, 65% cocktail), which corresponds to a protein concentration of 3.3 mg/ml. This value falls within the range of the DVR experiments setup using the 4.8 mg/ml protein solution that produced crystals, possibly of higher symmetry, that lacked SHG signal.

IV. DISCUSSIONS

Returning to our original questions, i.e., how do we identify nanocrystals as they grow?, how often do they occur compared to larger samples more suitable for today's synchrotron?, and how do we produce nanocrystals?, this study has provided some answers for soluble proteins. We have not studied nanocrystals as they grow; details of the various techniques to follow this are presented elsewhere.⁵ However, we have determined how often they occur compared to microcrystals. For the crystallization and imaging study, we observe that for the 60 samples (representing structurally uncharacterized soluble macromolecules of biological interest) out of the 39 that produced microcrystals, 18 also produced nanocrystals from the 1536 condition screen. Only one sample, protein N, did not produce visible microcrystals from the screen, but did provide SHG and UV-TPEF signals to support the presence of nanocrystals. For this sample set, only a small percentage of cases (<2%) failed to produce microcrystals if nanocrystals were detected.

As noted above, UV-TPEF and SHG characterization are not perfect. The loss of SHG signal for the protein in Figure 5 when grown at a lower concentration is consistent with published data for chicken egg-white lysozyme, where the crystal morphology was demonstrated to vary with protein concentration.²³ In the lysozyme study, holding all other variables constant, X-ray diffraction data verified that higher symmetry tetragonal crystals were grown from lower protein concentrations (100 mg/ml), while lower symmetry orthorhombic crystals were grown from higher protein concentrations (150 mg/ml). In our case, because we are analyzing visual data from samples sent by another laboratory, we do not have X-ray data available to confirm the relationship between crystal symmetry and the presence of SHG signal.

Chemical diversity was used as a goal in the selection of cocktails for optimization through DVR. In some cases this was not possible; precipitates with a positive SHG and UV-TPEF signature were sometimes only produced from cocktails that were chemically similar. This could be an indication that the required chemical conditions for the production of these nanocrystals fell within a chemically restricted region. Chemical similarity among sets of cocktails used for DVR optimization of each protein means that the success rate of the optimization from nanocrystals to microcrystals should not be used as a quantitative metric but rather seen as a qualitative result that it is possible to scale up from an initial nanocrystal. The vast majority of the cocktails in the 1536 screen contain PEG, salt and a buffer. It is therefore not surprising that these chemical components are found in the subsets of cocktails selected for DVR optimization for the individual proteins.

The chemical effects of changing the sample to cocktail drop volume ratios are chemically complex, multi-dimensional, and although related, do not map directly to a trajectory that describes the supersolubility of the protein on a simple two-dimensional phase diagram, Figure 6. For example, the co-dilution of all of the solutes from the protein and cocktail solutions will

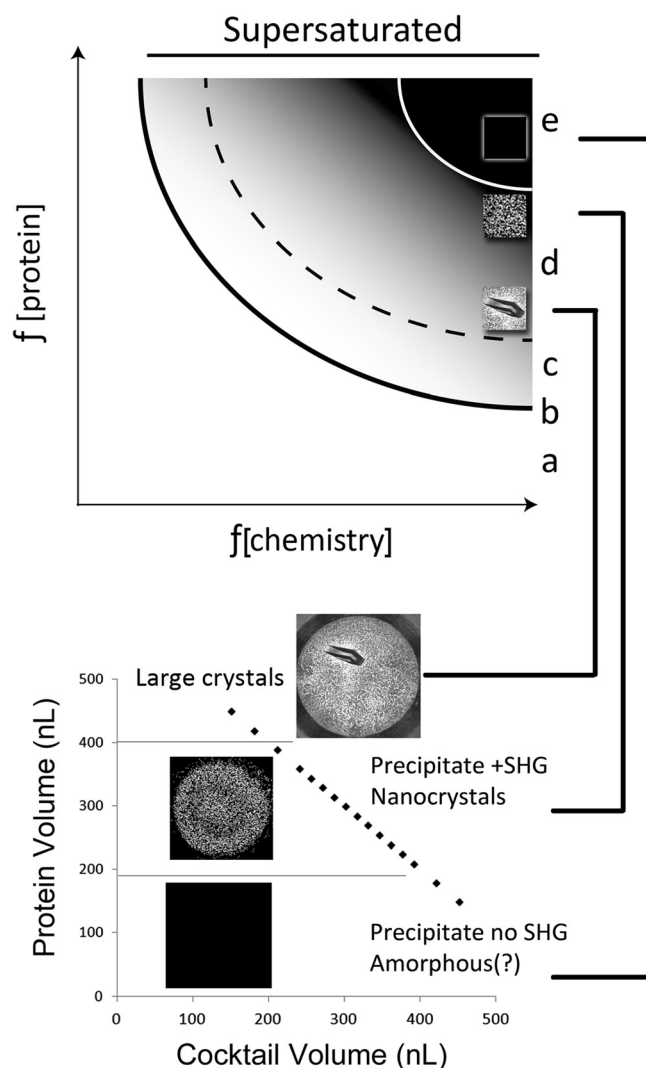


FIG. 6. Correlating DVR outcomes with a phase diagram. We can approximate the location of DVR outcomes to their location on a phase diagram to help guide interpretation of results. The top figure is a simple phase diagram showing the protein solubility as a function of protein concentration and solution chemistry. Nucleation is a stochastic process with a probability of occurrence that relates to the level of supersaturation. The regions in the diagram are (a) undersaturated, (b) saturated, (c) metastable (crystal growth occurs, nucleation is thermodynamically, but not kinetically probable), (d) labile (spontaneous homogeneous nucleation is thermodynamically and kinetically probable), and (e) precipitation (disordered aggregation). Nucleation and crystal growth decrease the concentration of protein in solution. When the concentration of protein is sufficiently decreased, the solution will reach a state where additional crystal nuclei will not form. Based upon the number and size of the crystals produced, we can infer the location of an experiment on a phase diagram. A few large crystals are more likely to occur in the labile region closer to the metastable; many small crystals are more likely to occur in the labile regions closer to the precipitation region. By applying this knowledge, we can relate the outcomes of DVR experiments, shown as the series of 16 points in the bottom figure, to a phase diagram to provide a rational framework to guide follow-up experiments.

lead to difficult to calculate changes in the solution pH. Very small changes in solution pH, on the order of hundredths of a pH unit, can effect crystallization.²⁴ Similarly collective, concentration-based variations in the solutes are readily calculated, but the effects of individual chemical components on the solubility of the protein can be difficult to delineate. That stated the phase diagram can still be used as a means to understand and correlate the experimental outcomes to a first approximation. Trends in protein solubility can be mapped to the phase diagram based on particular outcomes.²⁵ Clear drops are undersaturated, saturated, or metastable. Larger crystals will nucleate in the labile zone in a location closer to the metastable zone,

where additional spontaneous, homogeneous nucleation events cannot occur. Nanocrystals will form in the labile zone nearer to the precipitation zone, at high levels of supersaturation, where many additional nucleation events can take place prior to the solution becoming sufficiently depleted of protein that the level of supersaturation reaches a state, where no additional nucleation takes place. The use of detection methods that can identify nanocrystals in drops appearing visually to be either clear or precipitated enables a more detailed empirical determination of the phase diagram for a protein, and can be used to more precisely guide further optimization, including production of microcrystals or nanocrystals. If we know where we are in the phase diagram, and can reproduce those conditions using simple drop volume ratio methods, then we can target regions of higher or lower supersaturation to obtain the desired outcomes.

Phase diagram approaches to understanding crystallization and their use in transitioning from nanocrystals to microcrystals and vice versa require an accurate assessment of experimental outcome.²⁵ The use of SHG and UV-TPEF signalling combined with visible imaging provides an objective resolution to the problem of positive identification of macromolecular microcrystalline material. We can verify a precipitate is crystalline using SHG and verify that it is a macromolecule if UV-TPEF signal is present. This is an important development taking one aspect of crystallization that has long been considered an *art*, interpretation of precipitates, to a *science*. An example of this are experiments E.6 and E.7, where although the precipitates lacked SHG signal, they were considered similar in appearance to precipitates that had provided SHG signal and because of this visual similarity were selected for DVR optimization. The crystals produced from the optimization experiments lend credence to the concept of using SHG as a training tool to identify precipitates that are promising for optimization.

Our results suggest that at least for soluble proteins, within the set studied, there are few if any cases where an experiment is limited by crystal volume. Utilizing an XFEL, it is now possible to determine a crystal structure, such as that reported for Photosystem I, using crystals that range in size from 200 nm to 2 μm .¹ Even if crystal can be grown larger, there are additional advantages of data collection using nanocrystals. For example, time-resolved studies need the majority of the macromolecules in the crystal to have undergone the same reaction. Whether triggered by light, chemical reaction, or other means, the propagation time for any activation will be less in a smaller crystal. Similarly, as size increases, the chance for static disorder also increases. There may be cases where good data are only available from perfect domains of small crystals. Nanocrystals are small and therefore less material makes up each crystal. In the future this may become another advantage of XFEL sources, though with current technology the quantity of crystals required is large. Resources for nanocrystallography remain scarce and are highly utilized. Using the combination of SHG and UV-TPEF to detect nanocrystals, and the optimization protocols described to identify chemical conditions that produce suitably diffracting crystals for standard single crystal X-ray diffraction methods, we may be able to alleviate some of the burden on the XFEL and make the most effective use of this resource for those experiments that require nanocrystals.

Other approaches can be used to achieve nanocrystal formation when microcrystals are identified by targeting nucleation and growth rate. Classically, nucleation rates can be increased by adjusting the temperature to increase the level of supersaturation (effectively part of the DVR strategy), growth rate can be slowed by adjusting the temperature to decrease the level of supersaturation. The DVR approach is a simple first-stage optimization that makes use of the existing sample and crystallization cocktail. Building on this through refinement and the introduction of additional variables known to impact crystallization would undoubtedly lead to improved results.

SHG and UV-TPEF have been used to identify conditions where nanocrystals are produced. Both SHG and UV-TPEF imaging are now routine in the HTSlab crystallization pipeline and over 800×1536 well experiment plates (one protein per plate), of both soluble and membrane protein crystallization assays, have been imaged using the Formulatrix Rock Imager 1000. Technical and software modifications were developed with Formulatrix for the system to reach a sustained throughput of 5×1536 well plates/24 h, and we anticipate that this data will help to reveal the true prevalence of nanocrystals in the absence of microcrystals and the required

need (or not) for XFEL capability in regular structural studies of nanocrystals (assuming diffraction quality is maintained irrespective of volume considerations). The majority of this experimental data is distributed to external investigators using the HTSlab and not reviewed in-house. However, over time this data will become available through current investigators and future investigators making use of our academic crystallization screening service.

V. CONCLUSION

Nanocrystallography, enabled by XFELs, has the potential to open up crystallographic studies to samples that were previously recalcitrant. However, we note that for this study the frequency that a nanocrystal is obtained when a microcrystal cannot be obtained is very low (<2%) for the soluble proteins used in this study. This suggests a more nuanced role for XFEL applications to those studies in the small number of cases where larger crystals are unavailable or those where there is an experimental requirement for a nanocrystals. The optimization process described a simple screening of sample and crystallization cocktail ratio, is effective, but was not successful in all cases. Based upon this study it is possible, in the majority of cases, to go from nanocrystal to microcrystal and vice versa using these techniques.

ACKNOWLEDGMENTS

Funding support was partially provided by NIH GM100494, U54-GM094597 from the Protein Structure Initiative, NSF 1231306 and The New York State Economic Development Western NY Council. We thank Formulatrix, in particular, Roberto Dorich, Safdar Haider, and Zak Umanoff for working with us to develop the software and hardware capabilities required to adapt the Rock Imager 1000 with SONICC to our 1536 well plate requirements. We also thank Ms. H. Janjua and L. Mao for expert technical assistance. J.R.L. would also like to take this opportunity to thank the many scientists of the NESG, who provided the samples for this study, and many other opportunities for research.

- ¹H. N. Chapman, P. Fromme, A. Barty, T. A. White, R. A. Kirian, A. Aquila, M. S. Hunter, J. Schulz, D. P. DePonte, U. Weierstall, R. B. Doak, F. R. N. C. Maia, A. V. Martin, I. Schlichting, L. Lomb, N. Coppola, R. L. Shoeman, S. W. Epp, R. Hartmann, D. Rolles, A. Rudenko, L. Foucar, N. Kimmel, G. Weidenspointner, P. Holl, M. Liang, M. Barthelmeß, C. Caleman, S. Boutet, M. J. Bogan, J. Krzywinski, C. Bostedt, S. Bajt, L. Gumprecht, B. Rudek, B. Erk, C. Schmidt, A. Homke, C. Reich, D. Pietschner, L. Struder, G. Hauser, H. Gorke, J. Ullrich, S. Herrmann, G. Schaller, F. Schopper, H. Soltau, K.-U. Kuhnel, M. Messerschmidt, J. D. Bozek, S. P. Hau-Riege, M. Frank, C. Y. Hampton, R. G. Sierra, D. Starodub, G. J. Williams, J. Hajdu, N. Timneanu, M. M. Seibert, J. Andreasson, A. Rocker, O. Jonsson, M. Svenda, S. Stern, K. Nass, R. Andritschke, C.-D. Schroter, F. Krasniqi, M. Bott, K. E. Schmidt, X. Wang, I. Grotjohann, J. M. Holton, T. R. M. Barends, R. Neutze, S. Marchesini, R. Fromme, S. Schorb, D. Rupp, M. Adolph, T. Gorkhover, I. Andersson, H. Hirsemann, G. Potdevin, H. Graafsma, B. Nilsson, and J. C. H. Spence, "Femtosecond X-ray protein nanocrystallography," *Nature* **470**(7332), 73–77 (2011).
- ²A. E. Cohen, S. M. Soltis, A. Gonzalez, L. Aguila, R. Alonso-Mori, C. O. Barnes, E. L. Baxter, W. Brehmer, A. S. Brewster, A. T. Brunger, G. Calero, J. F. Chang, M. Chollet, P. Ehrensberger, T. L. Eriksson, Y. P. Feng, J. Hattne, B. Hedman, M. Hollenbeck, J. M. Holton, S. Keable, B. K. Kobilka, E. G. Kovaleva, A. C. Kruse, H. T. Lemke, G. W. Lin, A. Y. Lyubimov, A. Manglik, I. I. Mathews, S. E. McPhillips, S. Nelson, J. W. Peters, N. K. Sauter, C. A. Smith, J. H. Song, H. P. Stevenson, Y. S. Tsai, M. Uervirojnangkoom, V. Vinetsky, S. Wakatsuki, W. I. Weis, O. A. Zadovnyy, O. B. Zeldin, D. L. Zhu, and K. O. Hodgson, "Goniometer-based femtosecond crystallography with X-ray free electron lasers," *Proc. Natl. Acad. Sci. U.S.A.* **111**(48), 17122–17127 (2014).
- ³M. S. Hunter, B. Segelke, M. Messerschmidt, G. J. Williams, N. A. Zatsepin, A. Barty, W. H. Benner, D. B. Carlson, M. Coleman, A. Graf, S. P. Hau-Riege, T. Pardini, M. M. Seibert, J. Evans, S. Boutet, and M. Frank, "Fixed-target protein serial microcrystallography with an x-ray free electron laser," *Sci. Rep.* **4**:6026 (2014).
- ⁴C. Gati, G. Bourenkov, M. Klinge, D. Rehders, F. Stellato, D. Oberthür, O. Yefanov, B. P. Sommer, S. Mogk, M. Duszynski, C. Betzel, T. R. Schneider, H. N. Chapman, and L. Redecke, "Serial crystallography on in vivo grown microcrystals using synchrotron radiation," *IUCrJ.* **1**(2), 87–94 (2014).
- ⁵G. Calero, A. E. Cohen, J. R. Luft, J. Newman, and E. H. Snell, "Identifying, studying and making good use of macromolecular crystals," *Acta Crystallogr., Sect. F: Struct. Biol. Commun.* **70**, 993–1008 (2014).
- ⁶L. M. Hauptert and G. J. Simpson, "Screening of protein crystallization trials by second order nonlinear optical imaging of chiral crystals (SONICC)," *Methods* **55**(4), 379–386 (2011).
- ⁷J. T. Madden, E. L. DeWalt, and G. J. Simpson, "Two-photon excited UV fluorescence for protein crystal detection," *Acta Crystallogr., D: Biol. Crystallogr.* **67**(10), 839–846 (2011).
- ⁸J. R. Luft, E. H. Snell, and G. T. DeTitta, "Lessons from high-throughput protein crystallization screening: 10 years of practical experience," *Expert Opin. Drug Discovery* **6**(5), 465–480 (2011).
- ⁹T. Frickey and A. Lupas, "CLANS: A Java application for visualizing protein families based on pairwise similarity," *Bioinformatics* **20**(18), 3702–3704 (2004).

- ¹⁰See supplementary material at <http://dx.doi.org/10.1063/1.4921199> for crystallization results, biochemical, and biophysical characterization of the 60 NESG proteins used for this study.
- ¹¹R. Xiao, S. Anderson, J. Aramini, R. Belote, W. A. Buchwald, C. Ciccocanti, K. Conover, J. K. Everett, K. Hamilton, Y. J. Huang, H. Janjua, M. Jiang, G. J. Kornhaber, D. Y. Lee, J. Y. Locke, L. C. Ma, M. Maglaqui, L. Mao, S. Mitra, D. Patel, P. Rossi, S. Sahdev, S. Sharma, R. Shastry, G. V. T. Swapna, S. N. Tong, D. Y. Wang, H. A. Wang, L. Zhao, G. T. Montelione, and T. B. Acton, "The high-throughput protein sample production platform of the northeast structural genomics consortium," *J. Struct. Biol.* **172**(1), 21–33 (2010).
- ¹²J. P. Deng, D. R. Davies, G. Wisedchaisri, M. T. Wu, W. G. J. Hol, and C. Mehlh, "An improved protocol for rapid freezing of protein samples for long-term storage," *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **60**, 203–204 (2004).
- ¹³J. R. Luft, R. J. Collins, N. A. Fehrman, A. M. Lauricella, C. K. Veatch, and G. T. DeTitta, "A deliberate approach to screening for initial crystallization conditions of biological macromolecules," *J. Struct. Biol.* **142**(1), 170–179 (2003).
- ¹⁴N. E. Chayen, P. D. S. Stewart, and D. M. Blow, "Microbatch crystallization under oil—A new technique allowing many small-volume crystallization trials," *J. Cryst. Growth* **122**(1–4), 176–180 (1992).
- ¹⁵M. Koszelak-Rosenblum, A. Krol, N. Mozumdar, K. Wunsch, A. Ferin, E. Cook, C. K. Veatch, R. Nagel, J. R. Luft, G. T. DeTitta, and M. G. Malkowski, "Determination and application of empirically derived detergent phase boundaries to effectively crystallize membrane proteins," *Protein Sci.* **18**(9), 1828–1839 (2009).
- ¹⁶R. M. Garavito, D. Picot, and P. J. Loll, "Strategies for crystallizing membrane proteins," *J. Bioenerg. Biomembr.* **28**(1), 13–27 (1996).
- ¹⁷G. K. Christopher, A. G. Phipps, and R. J. Gray, "Temperature-dependent solubility of selected proteins," *J. Cryst. Growth* **191**(4), 820–826 (1998).
- ¹⁸J. P. Astier and S. Veessler, "Using temperature to crystallize proteins: A mini-review," *Cryst. Growth Des.* **8**(12), 4215–4219 (2008).
- ¹⁹R. G. Closser, E. J. Gualtieri, J. A. Newman, and G. J. Simpson, "Characterization of salt interferences in second-harmonic generation detection of protein crystals," *J. Appl. Crystallogr.* **46**(6), 1903–1906 (2013).
- ²⁰J. R. Luft, J. R. Wolfley, M. I. Said, R. M. Nagel, A. M. Lauricella, J. L. Smith, M. H. Thayer, C. K. Veatch, E. H. Snell, M. G. Malkowski, and G. T. DeTitta, "Efficient optimization of crystallization conditions by manipulation of drop volume ratio and temperature," *Protein Sci.* **16**(4), 715–722 (2007).
- ²¹I. Rayment, "Small-scale batch crystallization of proteins revisited: an underutilized way to grow large protein crystals," *Structure* **10**(2), 147–151 (2002).
- ²²D. J. Kissick, D. Wanapun, and G. J. Simpson, "Second-order nonlinear optical imaging of chiral crystals," *Annu. Rev. Anal. Chem.* **4**, 419–437 (2011).
- ²³E. L. Forsythe, E. H. Snell, and M. L. Pusey, "Crystallization of chicken egg-white lysozyme from ammonium sulfate," *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **53**, 795–797 (1997).
- ²⁴A. McPherson, "Increasing the size of microcrystals by fine sampling of pH limits," *J. Appl. Crystallogr.* **28**, 362–365 (1995).
- ²⁵J. R. Luft, J. R. Wolfley, and E. H. Snell, "What's in a drop? correlating observations and outcomes to guide macromolecular crystallization experiments," *Cryst. Growth Des.* **11**(3), 651–663 (2011).