Nanocrystalline protein domains via salting-out

Daniel G. Greene,[a] Shannon Modla,[b] Stanley I. Sandler,[a] Norman J. Wagner[a] and Abraham M. Lenhoff[a]*

[a]Department of Chemical and Biomolecular Engineering, University of Delaware, Newark, DE 19716, USA, and [b]Delaware Biotechnology Institute, Newark, DE 19713, USA. *Correspondence e-mail: lenhoff@udel.edu

Protein salting-out is a well established phenomenon that in many cases leads to amorphous structures and protein gels, which are usually not considered to be useful for protein structure determination. Here, microstructural measurements of several different salted-out protein dense phases are reported, including of lysozyme, ribonuclease A and an IgG1, showing that salted-out protein gels unexpectedly contain highly ordered protein nanostructures that assemble hierarchically to create the gel. The nanocrystalline domains are approximately 10–100 nm in size, are shown to have structures commensurate with those of bulk crystals and grow on time scales in the order of an hour to a day. Beyond revealing the rich, hierarchical nanoscale to mesoscale structure of protein gels, the nanocrystals that these phases contain are candidates for structural biology on next-generation X-ray free-electron lasers, which may enable the study of biological macromolecules that are difficult or impossible to crystallize in bulk.

1. Introduction

Recent advances in X-ray light sources (Miao *et al.*, 2015) and diffraction analysis (Ayyer *et al.*, 2016) are reducing the requirement for large, well diffracting protein crystals, and several structures of proteins that are difficult to crystallize have been solved using submicrometre- to micrometre-sized protein crystals (Chapman *et al.*, 2011; Kern *et al.*, 2013; Liu *et al.*, 2013; Tenboer *et al.*, 2014). In principle, these new X-ray techniques can enable structure determination via diffraction from single protein molecules (Miao *et al.*, 2001), and beamlines such as those at the European X-ray free-electron laser (XFEL) are being designed to measure diffraction from nanocrystals as small as 15 nm (Ayyer *et al.*, 2015).

Although small crystals are easier to grow than larger crystals (Spence *et al.*, 2012), crystallizing some proteins remains challenging in general. Noncrystalline protein dense phases are commonly observed during the search for protein crystals (Bergfors, 2003; McPherson, 2004), and while these phases are generally considered to be amorphous and hence uninteresting, there are reports of the successful crystallization of a protein from such a seemingly amorphous phase (Bergfors, 2003; Ng *et al.*, 1996; Van Driessche *et al.*, 2018). We recently reported the microstructure of a salted-out ovalbumin dense phase and surprisingly discovered that it contains nanoscale regions of highly ordered protein that are well described by nanocrystalline protein clusters (Greene *et al.*, 2015). Moreover, the nanocrystals are hierarchically assembled into a porous gel network, the structure of which is qualitatively similar to those of the gels formed during crystallization-arrested phase separation of colloidal spheres
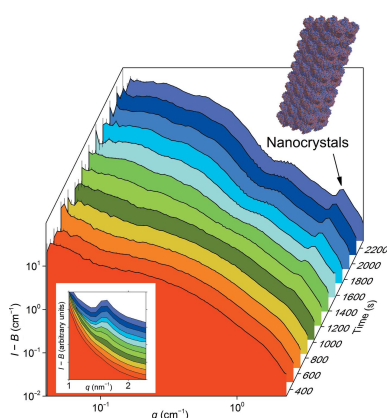
Nanocrystals

(Fortini *et al.*, 2008; Sabin *et al.*, 2012; Tsurusawa *et al.*, 2017; Zhang *et al.*, 2012); such protein gels may have meaningful functions in biological systems (Cai *et al.*, 2017). It is unclear whether the nanocrystals are unique to ovalbumin, but the observation of 'nanorods' and fibers in the crystallization of glucose isomerase (Van Driessche *et al.*, 2018) suggests that they may be representative of a more general phenomenon. We support this postulation here by presenting microstructural measurements of several salted-out protein gels formed from proteins spanning a broad range of folded structures in solution: lysozyme, ribonuclease A (RNAse A) and an IgG1 monoclonal antibody.

2. Materials and methods

2.1. Materials

Lysozyme (CAS 12650-88-3) was purchased from Sigma and RNAse A (CAS 9001-99-4) from Worthington Biochemical Corporation; both were used as received. The IgG1 was provided by Genentech. Ovalbumin was recrystallized three times from fresh hen egg white as described previously (Greene *et al.*, 2015).

Ammonium sulfate (CAS 7783-20-2), sodium chloride, sodium phosphate (CAS 7558-80-7), sodium acetate and glacial acetic acid (CAS 64-19-7) were purchased from Fisher Scientific. Deuterium oxide was purchased from Cambridge Isotope Laboratories. For TEM sample preparation, osmium tetroxide (CAS 20816-12-0), glass-distilled acetone (CAS 67-64-1), ethylene glycol diglycidyl ether (CAS 2224-15-9) resin (Quetol 651), nonenylsuccinic anhydride (NSA; CAS 28928-97-4), *n*-butyl glycidyl ether (CAS 2426-08-6), uranyl acetate (CAS 541-09-3), lead nitrate (CAS 10099-74-8), sodium citrate (CAS 6132-04-3) and 200 mesh Formvar carbon-coated copper grids were purchased from Electron Microscopy Services.

2.2. Sample preparation

Buffers and stock salt solutions were prepared by dissolving the appropriate quantities of buffer components in deionized (18.2 MΩ cm) Milli-Q water or deuterium oxide, adjusted to the appropriate pH/pD and filtered through a 0.22 µm filter. The relation $pD = pH_{read} + 0.4$ for measuring the pD with a glass electrode was used (Glasoe & Long, 1960).

Protein solutions were buffer-exchanged to the desired conditions using 10K molecular-weight cutoff Amicon Ultra-4 centrifugal filters (Millipore). Lysozyme was exchanged into 5 mM sodium phosphate buffer pH 7.0, RNAse A into 5 mM sodium acetate/acetic acid buffer pH 4.0, the IgG1 into 5 mM sodium acetate/acetic acid in D_2O pD 5.0 and ovalbumin into 5 mM sodium phosphate in D_2O pD 7.0. Diafiltration was performed until the final pH/pD matched that of the buffer. After buffer exchange, the protein solutions were concentrated to ~100–150 mg ml^{-1} using Amicon Ultra-4 centrifugal filters. The protein concentration was determined from the absorbance at 280 nm.

To reduce the background and improve the contrast in the SANS experiments, ammonium sulfate was partially deuterated by dissolving the salt in D_2O and subsequently freeze-drying the solution to remove water. This process was repeated three times, resulting in 82% deuterated ammonium sulfate.

Salted-out samples were prepared as described previously (Greene *et al.*, 2015). Briefly, the samples were prepared by mixing appropriate amounts of a high-concentration salt stock, the appropriate buffer and the concentrated protein stock, in that order.

2.3. Scattering measurements

Lysozyme measurements in ammonium sulfate and some additional IgG1 measurements (Supplementary Fig. S1) were performed on the G1 BioSAXS beamline at the Cornell High Energy Synchrotron Source with 9.97 keV X-rays over a *q*-range of 0.0635–2.83 nm^{-1}. Samples were oscillated within a flow cell to minimize beam damage (Nielsen *et al.*, 2012). Thirty 1 s exposures were averaged to obtain the final profile. Data were reduced using *BioXTAS RAW* (Nielsen *et al.*, 2009).

Lysozyme measurements in sodium chloride and RNAse A measurements were collected on the 9ID USAXS/SAXS beamline at the Advanced Photon Source. Samples were loaded into 1 mm borosilicate capillaries from Kimble Chase. USAXS data were collected using 17.9 keV X-rays over a *q*-range of 0.0014–2.9 nm^{-1} in fly scan mode with 90 s exposures. The beam was collimated along one dimension with slit length 0.27779 nm^{-1}. More information about the beamline is given in Ilavsky *et al.* (2013). Data were reduced, put onto an absolute scale and desmeared using the *Indra IGOR Pro* macros. Due to the large instrumental background on USAXS at high *q*, complementary SAXS data were also collected. They were measured with a pinhole camera (PILATUS 100K) at the same energy as the USAXS measurements and covered a *q*-range of 0.35–14.5 nm^{-1}. The SAXS data were reduced using the *Nika IGOR Pro* macros (Ilavsky, 2012).

IgG1 and ovalbumin SANS measurements were made on the EQ-SANS instrument at the Spallation Neutron Source at Oak Ridge National Laboratory (ORNL). The sample-to-detector distance was 4 m and the instrument was operated at 30 Hz to produce two neutron bands to cover a *q*-range of 0.04–5.5 nm^{-1}. Samples were prepared at the beamline, loaded into banjo cells with 1 mm path length and sealed with Parafilm. Data were reduced and put onto an absolute scale using the standard methods in *MantidPlot*. Time-dependent data were binned into 1 min slices. Additional time-dependent data for ovalbumin (Supplementary Fig. S3) were collected on the CG-3 beamline at the High Flux Isotope Reactor (HFIR) at ORNL using two sample-to-detector distances, 30 cm and 6 m, with 0.6 nm neutrons to cover a *q*-range of 0.070–7.37 nm^{-1}.

Unless otherwise noted, all data fitting was performed in *SasView* (http://www.sasview.org/). Error bars on fitted parameters are the standard errors reported in *SasView*. All data were fitted with models incorporating instrument resolutions.

research communications

2.4. Microscopy

Transmission electron micrographs were prepared (Greene *et al.*, 2015) by pelleting salted-out samples by centrifugation at 11 400*g* for 2 min. Pellets were transferred to 1.2 × 200 µm specimen holders and frozen using a Leica EM PACT high-pressure freezer. The samples were transferred to a Leica AFS (automated freeze substitution) unit for freeze substitution in 2% osmium tetroxide containing 1% water and 99% acetone. Briefly, the samples were freeze substituted at −85°C for approximately four days, whereupon they were slowly warmed to −20°C over 15 h. The samples were held at −20°C for 3 h, warmed to 4°C over 2 h and then held at 4°C for 2 h. The samples were then removed from the Leica AFS, held at room temperature for 1.5 h and washed with anhydrous acetone. After freeze substitution, the samples were infiltrated with Quetol 651-NSA resin and, after polymerization, were sectioned into 60–70 nm thick slices using a Reichert–Jung Ultracut E ultramicrotome with a DiATOME ultra diamond knife. The sections were transferred onto 200 mesh Formvar carbon-coated copper grids and stained with 2% methanolic uranyl acetate, followed by staining with Reynolds's lead citrate. A Zeiss Libra 120 transmission electron microscope operated at 120 kV was used to image the grids and a Gatan Ultrascan 1000 CCD camera was used to acquire images.

2.5. Analysis

Molecular chains of lysozyme were generated *in silico* from PDB entry 3lzt, which belongs to space group *P*1 (Walsh *et al.*, 1998). Crystalline chains were generated by propagating the lysozyme molecule along the *c* axis (Salemme *et al.*, 1988). Scattering from molecular chains was calculated using *CRYSOL* (Svergun *et al.*, 1995). The data were fitted to all possible linear combinations of monomers, dimers and trimers. The Akaike information criterion (Akaike, 1974) was used to select the best-fitting model. Two sets of calculations were performed, one assuming that there is negligible scattering from the hydration layer and the other that the electron density in the layer is 10% larger than that of the bulk (Svergun *et al.*, 1995). In both cases, we found that the trimer crystalline chain is the best-fitting chain. Linear least-squares fitting of the *CRYSOL* profiles was performed in *MATLAB*.

A nanocrystalline sheet of RNAse A was generated *in silico* from PDB entry 1js0 by propagating the unit cell along the three crystallographic axes. The sheet was 5 × 5 × 3 unit cells in size and was chosen to match the approximate thickness of the nanocrystalline protrusion shown in Fig. 2(*c*).

Power spectra were obtained by taking the Fourier transform of TEM images and integrating the intensity as a function of distance from zero frequency in units of *q*. The fractal dimension, *d*, of an object created by the intersection of a fractal of dimension d_f with a plane is given by $d = d_f - 1$ (Isichenko, 1992), so the intensity is expected to scale as $q^{-d} = q^{1-d_f}$.

Nanocrystalline domain sizes were estimated using Scherrer's equation (Scherrer, 1918), $D = 2\pi/\text{FWHM}$, where *D* is the domain size and FWHM is the full-width at half-maximum of the peak in units of *q*. The relation assumes the crystallite shape parameter is unity, but this parameter can range from ∼0.7 to 2 depending on the crystallite shape (Langford & Wilson, 1978). FWHM was estimated by fitting a Lorentzian peak with power-law background to the data.

3. Results

3.1. Lysozyme

Salted-out lysozyme flocs are fractal-like on large length scales (Georgalis *et al.*, 1997), but the microstructure is unknown on length scales comparable to that of the monomer. Despite this, several studies suggest that lysozyme tends to form chain-like aggregates. A meta study of lysozyme crystals in several space groups revealed that specific intermolecular orientations are conserved across several disparate space groups (Salemme *et al.*, 1988). The resulting molecular chain was hypothesized to be important to lysozyme crystallization (Salemme *et al.*, 1988), and eventually chain-like crystal precursors were found in supersaturated lysozyme solutions (Michinomae *et al.*, 1999). All-atom free-energy calculations have shown that elongated crystal nuclei are energetically preferred over more compact, globular structures (Drenth *et al.*, 2003).

A representative small-angle X-ray scattering (SAXS) spectrum for lysozyme salted-out using 1.25 *M* ammonium sulfate pH 7 is shown in Fig. 1. The SAXS data are well described by small-angle scattering calculated from the molecular chain that is conserved across several disparate space groups (Salemme *et al.*, 1988; Fig. 1, green curve). An empirical power-law dependence of q^{-4} was added to the model of scattered intensity to account for surface scattering from the large flocs. The best-fitting nanocrystalline chain consists of three lysozyme monomers ordered along the chain length and, if approximated as a cylinder (purple curve), is 10.1 ± 0.8 nm long with a radius of 1.5 ± 0.1 nm. The inset in Fig. 1 shows the full, all-atom chain structure overlaid with the best-fit cylindrical representation. Nearly identical chains formed in flocs salted-out with 1.25 *M* sodium chloride (radius and length of 1.3 ± 0.1 and 9.6 ± 0.1 nm, respectively). Chains formed within minutes in the presence of both salts and were observed up to salt concentrations of 1.25 *M* ammonium sulfate and 1.25 *M* sodium chloride, but salt concentrations greater than these yielded no evidence of chain formation.

3.2. Ribonuclease A

RNAse A salted-out in ammonium sulfate solutions at pH 4. On salting-out with 2.25 *M* ammonium sulfate, RNAse A condenses into spherulite particles that contain an amorphous core and are decorated with rectangular protrusions (Figs. 2*a*–2*d*). The protrusions are approximately 30–100 nm thick and consist of parallel, highly ordered striations (Fig. 2*c*). The power spectrum (Fig. 2*e*) of the intensity measured across the protrusion in Fig. 2(*c*) exhibits a broad peak near $q = 0.80$ nm^{-1}, which corresponds to a correlation length scale of 7.85 nm. The theoretical scattering calculated from a

nanocrystalline sheet of RNAse A also exhibits a peak near $q = 0.80$ nm^{-1} (Fig. 2f) with comparable width to that of the power spectrum, which strongly suggests that the protrusions are in fact RNAse A nanocrystals. The low-q region of typical USAXS spectra (Fig. 2g) exhibits a power-law dependence with exponent -2.3, consistent with scattering from a dense, mass-fractal network of fractal dimension 2.3, such as is observed in the interior of the spherulite particles. Indeed, the power spectrum of the particle interior determined from direct imaging exhibits a power-law scaling consistent with a mass fractal of fractal dimension 2.4 ± 0.1 (Supplementary Fig. S1). Unlike lysozyme, there is no evidence for molecular chains in the SAXS profiles (inset in Fig. 2f). Instead, there is a broad peak similar to that observed in scattering from ovalbumin nanocrystals (Greene et $al.$, 2015). The peak grows and sharpens with decreasing salt concentration. The nanocrystal domain sizes, estimated from the peak width, are 139 ± 23 and 63 ± 7 nm for the samples prepared in 1.5 and 2.2 M ammonium sulfate, respectively, which are in good agreement with the thickness of the nanocrystalline protrusions of the spherulite particles.

3.3. Immunoglobulin G

The salted-out IgG1 exhibits a microstructure that is markedly different from that of lysozyme or RNAse A and is more similar to that of ovalbumin (Greene et $al.$, 2015). The

protein was exposed to 0.75 M ammonium sulfate at pD 5.0 and aged for $\sim$1 h before measurement. A complete description of the phase behavior and second osmotic virial coefficient of the IgG1 in the presence of ammonium sulfate is given in Lewus et $al.$ (2015), where the antibody is denoted mAb B. Transmission electron microscopy (TEM; Figs. 3a, 3b and 3c) reveals that the dense phase consists of uniform tubes packed hexagonally into bundles. The tubes have an outer diameter (OD) of 29 ± 7.5 nm and a wall thickness of 8.5 ± 1.9 nm. Small-angle neutron scattering (SANS) data (Fig. 3d) exhibit broad peaks that are consistent with the structure factor for hexagonally packed cylinders, and the correlation length scale of the primary reflection is $\sim 2\pi/q^* = 29.5$ nm, which is in good agreement with the tube OD. The magnitude of the reflections diminishes at higher q due to the finite nature of the bundle size. Importantly, there is a peak at high q that is sharper than the higher-order reflections of the hexagonal unit cell, indicating that the IgG1 is packing in a highly ordered fashion (Greene et $al.$, 2015). The correlation length scale associated with this peak, $2\pi/q^{**} = 3.59$ nm, is commensurate with the size of an individual molecule, while the domain size of the nanocrystal grain is 8.5 ± 1.5 nm, which agrees well with the wall thickness measured via TEM. The nanocrystalline nature is further supported by SAXS data taken at 1 M ammonium sulfate pH 5 five days after sample preparation (Supplementary Fig. S2), which indicate a sharpening of the high-q peak consistent with nanocrystal growth. The domain size associated with the SAXS peak is 109 ± 5 nm.

4. Discussion and conclusions

We have shown that the salted-out protein gel phases studied all contain crystalline-like nanodomains, suggesting that the phenomenon is a more general feature of such protein gels. Further, the crystalline structure evident in the nanocrystals in each specific protein gel is found to be similar to a bulk crystal structure reported for each protein or, in the case of the IgG, a structurally similar one. These nanocrystalline domains are of the order of 10 nm or larger in size and hence may be suitable candidates for structural biology on next-generation XFEL instruments. This discovery of nanocrystalline ordered domains in protein amorphous phases over a range of proteins reveals a rich, hierarchical structure in salted-out protein gels that were previously thought to be amorphous. These observations have scientific value both in understanding protein crystallization more broadly and in providing a potential new route for preparing samples for structural biology studies. Furthermore, studies of the critical role of nanocrystals in hard-sphere glass formation report that 'we expect structures locally reminiscent of the symmetry of the crystal to be present as common fluctuations of the supercooled liquid' (Leocmach & Tanaka, 2012). This suggests a possible broader, common pathway to dynamic arrest across a very broad range of systems. The present work already includes results for an IgG1, a class of molecules that are typically difficult to crystallize, presumably in part because of the flexibility of the
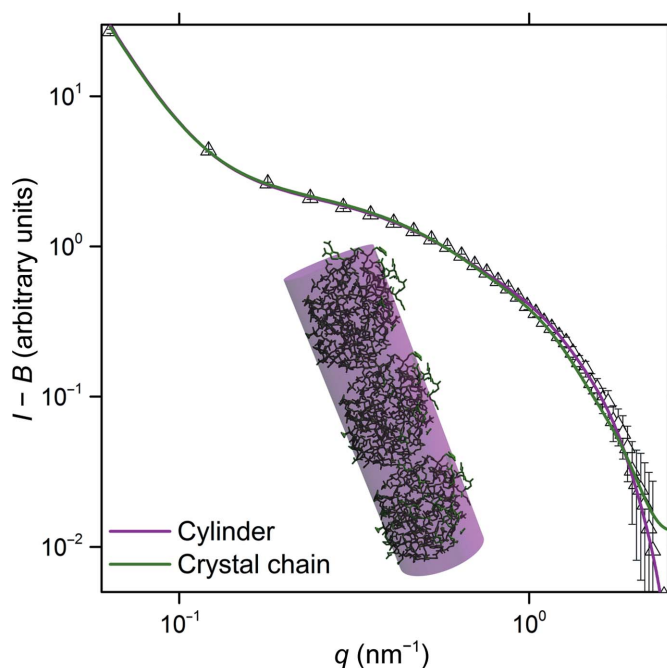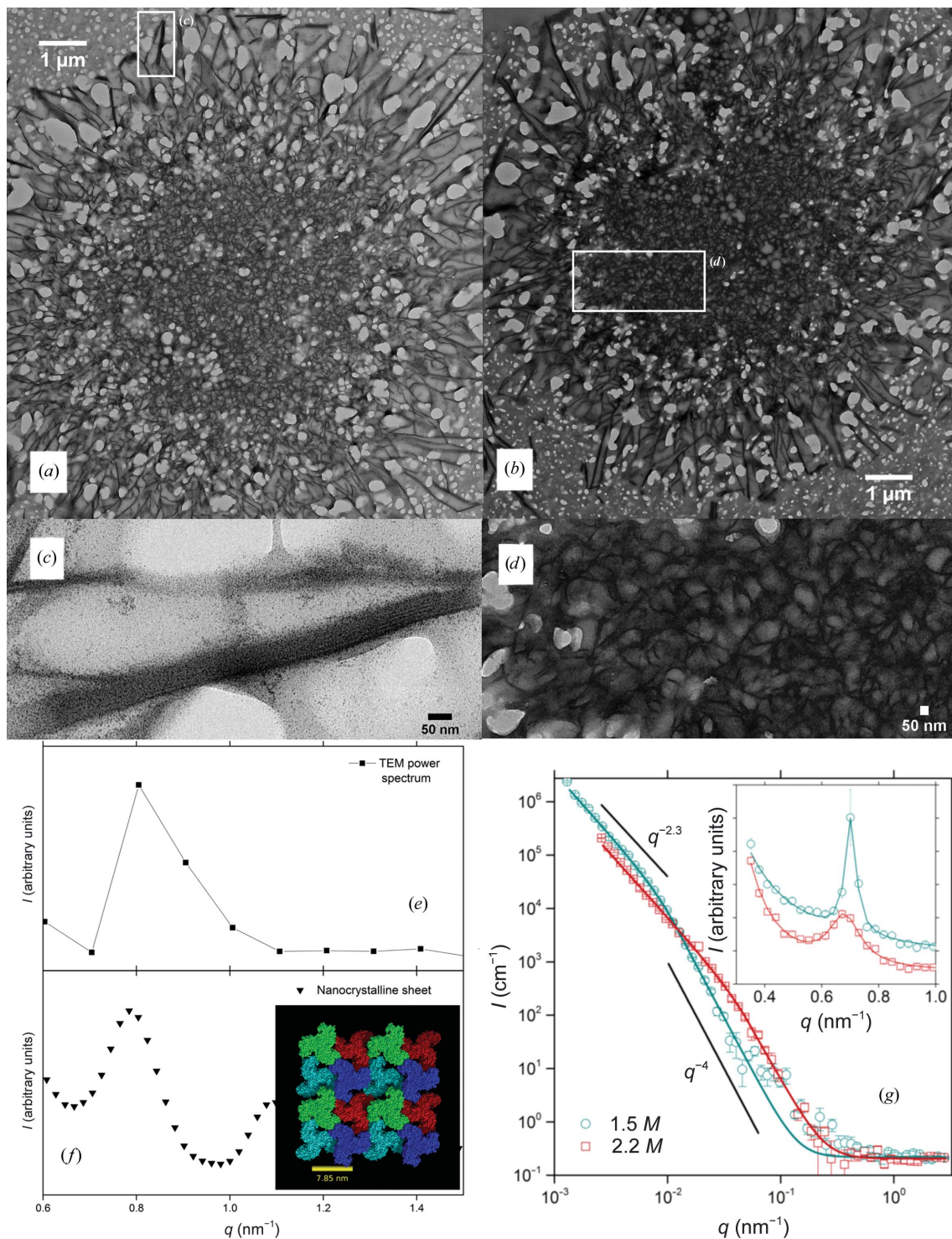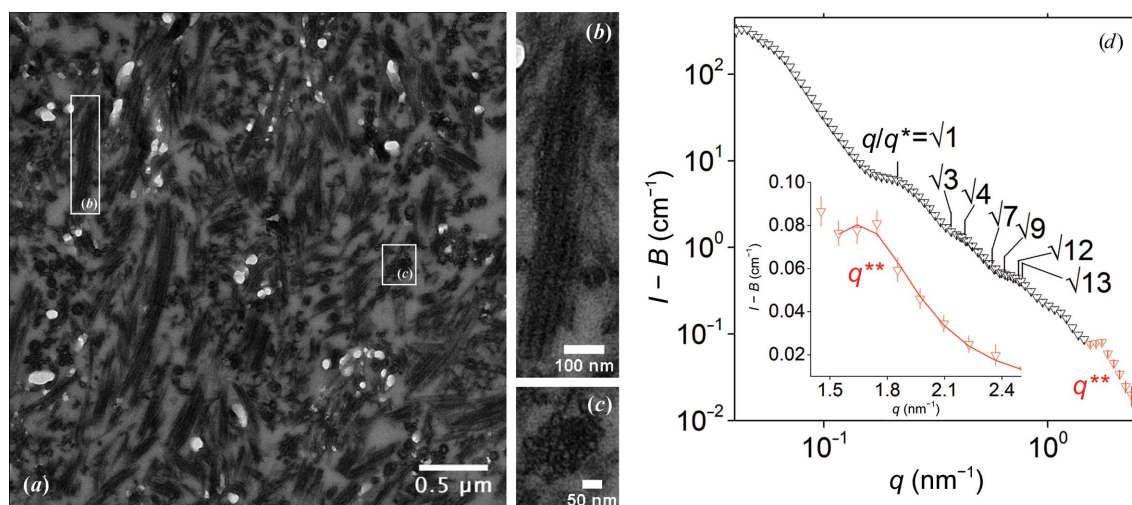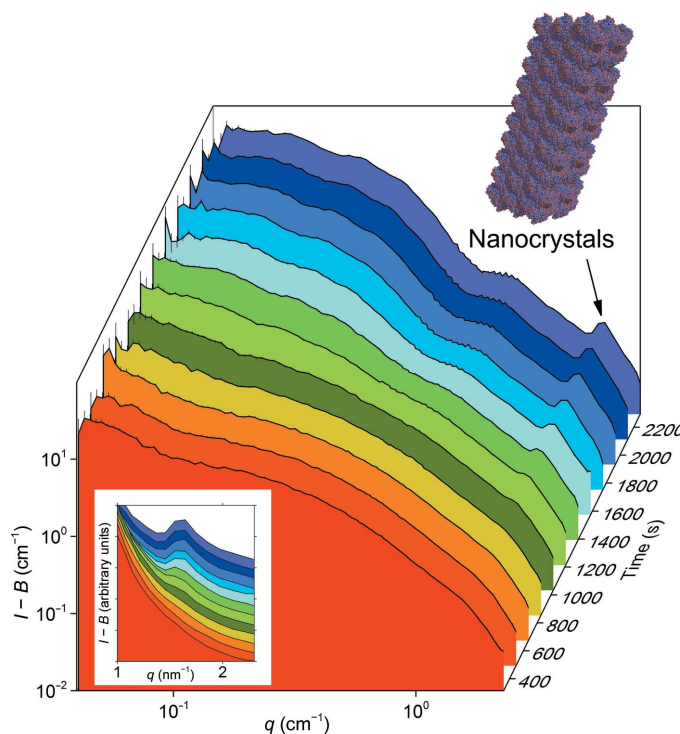

Figure 1
Representative SAXS profile for lysozyme salted-out with 1.25 M ammonium sulfate. The data are fitted to a model of a crystalline chain comprised of three lysozyme monomers (green curve) derived from the conserved molecular chain in crystallographic data (Salemme et $al.$, 1988), with scattering from any hydration layer considered to be negligible. The chain was also approximated as a cylinder (purple curve). The inset shows the best-fit cylinder (10.1 ± 0.8 nm long with a radius of 1.5 ± 0.1 nm) overlaid on the molecular chain. A q^{-4} power-law dependence at low q was added to both models to account for scattering from the large flocs.

Figure 2
The structure of salted-out RNAse A spherulites. (*a*, *b*) TEM micrographs illustrating the amorphous core of the spherulite and the rectangular protrusions. The cavities visible are areas of the sections where the resin did not successfully penetrate. The samples were prepared in 2.25 *M* ammonium sulfate at a protein concentration of 20 mg ml^{-1} and were aged for one day prior to measurement. (*c*) TEM micrograph showing the amorphous interior of the spherulite shown in (*a*). (*d*) TEM micrograph showing a nanocrystalline protrusion on the periphery of the spherulite in (*b*). (*e*) Power spectrum of the intensity profile perpendicular to the long axis of the nanocrystalline protrusion shown in (*c*). (*f*) Theoretical small-angle scattering calculated from an RNAse A nanocrystalline sheet (inset) derived from PDB entry 1js0. The sheet has dimensions of 5 × 5 × 3 unit cells. The correlation length scale of the peak in the power spectrum (*e*) and the theoretical peak (*f*) is 7.85 nm. (*g*) Desmeared USAXS and SAXS (inset) of samples prepared in 1.5 and 2.2 *M* ammonium sulfate at a protein concentration of 20 mg ml^{-1} and aged for one day prior to measurement. The USAXS spectra show a power-law dependence at low *q* with an exponent of −2.3, indicating a dense, mass-fractal like structure consistent with the dense network observed in the interior of the spherulites (*b*), that transitions to an exponent of −4 at higher *q*, consistent with scattering from a smooth surface. Peaks arise in the SAXS spectra near *q* = 0.7 nm^{-1}. These peaks do not appear in the USAXS spectra due to a high background. The domain sizes associated with the peak widths are 139 ± 23 and 63 ± 7 nm for the 1.5 and 2.2 *M* samples, respectively.

Figure 3
The microstructure of an IgG1 salted-out with 0.75 M ammonium sulfate in D_2O. (a) TEM micrograph illustrating the presence of tubular bundles of IgG1. The tubes have an outer diameter of 29 ± 7.5 nm and a wall thickness of 8.5 ± 1.9 nm. (b) Bundle cross-section cut along the long axis. (c) Bundle cross-section cut along the short axis. (d) SANS data. The allowed reflections for the hexagonal unit cell are indicated. The correlation length scale associated with the primary reflection (q^*) is 29.5 nm and there is a sharper peak at higher q (q^{**}). The inset shows the peak at q^{**} fitted to a Lorentzian function to extract the peak width, from which the nanocrystal size is found to be 8.5 ± 1.5 nm.

hinge region. We have not investigated membrane proteins, but they have been studied extensively in the form of nanocrystals and microcrystals (Hunter & Fromme, 2011), including the manipulation of crystallization behavior by varying solution conditions (Fromme & Witt, 1998).

Figure 4
Time-resolved SANS data monitoring the growth of nanocrystalline domains in salted-out ovalbumin. The time indicated along the z axis represents the total time in seconds after sample preparation. The nanocrystals develop to their final size (10 ± 1 nm) in less than 40 min. A thorough analysis of the nanocrystalline domain structure is given in Greene et al. (2015).

The nanocrystals described here are relatively easy to grow and form rapidly in comparison to bulk crystallization. Fig. 4 shows time-resolved small-angle neutron scattering (SANS) measurements, extracted from Supplementary Movie M1, monitoring the growth of nanocrystals in salted-out ovalbumin (a detailed structural analysis of salted-out ovalbumin is given in Greene et al., 2015). Under these conditions (2.1 M ammonium sulfate) the nanocrystals grow in 40 min to their final size of 10 ± 1 nm. Increasing the salt concentration slows the nanocrystal growth (Supplementary Fig. S3). In the other protein systems, the nanocrystal domains form on the order of minutes for lysozyme, one day for RNAse A and about 40 min for the IgG1 (Supplementary Fig. S4, Supplementary Movie M2).

The nanocrystalline structures appear to be suppressed at higher salt concentrations (for example the absence of molecular chains in lysozyme flocs salted-out at salt concentrations of greater than 1.25 M and a decrease in crystalline domain size in RNAse A nanocrystals with increasing salt concentrations; Fig. 2g). This may be due to the presence of dynamically arrested states such as gels (Dumetz et al., 2008). Successful generation of nanocrystalline domains can be aided by first mapping out the phase boundaries and targeting salt concentrations above, but close to, the instantaneous phase boundaries (Dumetz et al., 2008). The phase boundaries for the proteins studied here have been reported previously (Dumetz et al., 2008; Lewus et al., 2015) and were experimentally confirmed in the present work (Supplementary Figs. S5–S7).

The facile growth of nanocrystalline protein domains by simple salting-out suggests a possible novel approach to structural biology in which conditions yielding nanocrystalline domains are screened for in a high-throughput fashion using a standard SAXS beamline and currently available X-ray-transparent microfluidic devices (Heymann et al., 2014). Once

conditions are found that yield the signature broad peak near $q = 1$ nm^{-1}, follow-up experiments on next-generation XFELs could be used to perform structural biology. Moreover, the systems described here are ideal model systems to study fundamental colloidal phenomena, such as crystallization-arrested phase separation, which has been investigated primarily via computer simulation to this point.

References

Akaike, H. (1974). *IEEE Trans. Autom. Contr.* **19**, 716–723.
Ayyer, K., Geloni, G., Kocharyan, V., Saldin, E., Serkez, S., Yefanov, O. & Zagorodnov, I. (2015). *Struct. Dyn.* **2**, 041702.
Ayyer, K., Yefanov, O. M., Oberthür, D., Roy-Chowdhury, S., Galli, L., Mariani, V., Basu, S., Coe, J., Conrad, C. E., Fromme, R., Schaffer, A., Dörner, K., James, D., Kupitz, C., Metz, M., Nelson, G., Xavier, P. L., Beyerlein, K. R., Schmidt, M., Sarrou, I., Spence, J. C. H., Weierstall, U., White, T. A., Yang, J. H., Zhao, Y., Liang, M. N., Aquila, A., Hunter, M. S., Robinson, J. S., Koglin, J. E., Boutet, S., Fromme, P., Barty, A. & Chapman, H. N. (2016). *Nature*, **530**, 202–206.
Bergfors, T. (2003). *J. Struct. Biol.* **142**, 66–76.
Cai, J., Townsend, J. P., Dodson, T. C., Heiney, P. A. & Sweeney, A. M. (2017). *Science*, **357**, 564–569.
Chapman, H. N., Fromme, P., Barty, A., White, T. A., Kirian, R. A., Aquila, A., Hunter, M. S., Schulz, J., DePonte, D. P., Weierstall, U., Doak, R. B., Maia, F., Martin, A. V., Schlichting, I., Lomb, L., Coppola, N., Shoeman, R. L., Epp, S. W., Hartmann, R., Rolles, D., Rudenko, A., Foucar, L., Kimmel, N., Weidenspointner, G., Holl, P., Liang, M. N., Barthelmess, M., Caleman, C., Boutet, S., Bogan, M. J., Krzywinski, J., Bostedt, C., Bajt, S., Gumprecht, L., Rudek, B., Erk, B., Schmidt, C., Hömke, A., Reich, C., Pietschner, D., Strüder, L., Hauser, G., Gorke, H., Ullrich, J., Herrmann, S., Schaller, G., Schopper, F., Soltau, H., Kühnel, K. U., Messerschmidt, M., Bozek, J. D., Hau-Riege, S. P., Frank, M., Hampton, C. Y., Sierra, R. G., Starodub, D., Williams, G. J., Hajdu, J., Timneanu, N., Seibert, M. M., Andreasson, J., Rocker, A., Jönsson, O., Svenda, M., Stern, S., Nass, K., Andritschke, R., Schröter, C. D., Krasniqi, F., Bott, M., Schmidt, K. E., Wang, X. Y., Grotjohann, I., Holton, J. M., Barends, T. R. M., Neutze, R., Marchesini, S., Fromme, R., Schorb, S., Rupp, D., Adolph, M., Gorkhover, T., Andersson, I., Hirsemann, H., Potdevin, G., Graafsma, H., Nilsson, B. & Spence, J. C. H. (2011). *Nature*, **470**, 73–77.
Drenth, J., Dijkstra, K., Haas, C., Leppert, J. & Ohlenschläger, O. (2003). *J. Phys. Chem. B*, **107**, 4203–4207.
Dumetz, A. C., Chockla, A. M., Kaler, E. W. & Lenhoff, A. M. (2008). *Biophys. J.* **94**, 570–583.
Fortini, A., Sanz, E. & Dijkstra, M. (2008). *Phys. Rev. E*, **78**, 041402.
Fromme, P. & Witt, H. T. (1998). *Biochim. Biophys. Acta*, **1365**, 175–184.
Georgalis, Y., Umbach, P., Raptis, J. & Saenger, W. (1997). *Acta Cryst.* D**53**, 691–702.
Glasoe, P. K. & Long, F. A. (1960). *J. Phys. Chem.* **64**, 188–190.
Greene, D. G., Modla, S., Wagner, N. J., Sandler, S. I. & Lenhoff, A. M. (2015). *Biophys. J.* **109**, 1716–1723.
Heymann, M., Opthalage, A., Wierman, J. L., Akella, S., Szebenyi, D. M. E., Gruner, S. M. & Fraden, S. (2014). *IUCrJ*, **1**, 349–360.
Hunter, M. S. & Fromme, P. (2011). *Methods*, **55**, 387–404.
Ilavsky, J. (2012). *J. Appl. Cryst.* **45**, 324–328.
Ilavsky, J., Zhang, F., Allen, A. J., Levine, L. E., Jemian, P. R. & Long, G. G. (2013). *Metall. Mater. Trans. A*, **44**, 68–76.
Isichenko, M. B. (1992). *Rev. Mod. Phys.* **64**, 961–1043.
Kern, J., Alonso-Mori, R., Tran, R., Hattne, J., Gildea, R. J., Echols, N., Glöckner, C., Hellmich, J., Laksmono, H., Sierra, R. G., Lassalle-Kaiser, B., Koroidov, S., Lampe, A., Han, G. Y., Gul, S., Difiore, D., Milathianaki, D., Fry, A. R., Miahnahri, A., Schafer, D. W., Messerschmidt, M., Seibert, M. M., Koglin, J. E., Sokaras, D., Weng, T. C., Sellberg, J., Latimer, M. J., Grosse-Kunstleve, R. W., Zwart, P. H., White, W. E., Glatzel, P., Adams, P. D., Bogan, M. J., Williams, G. J., Boutet, S., Messinger, J., Zouni, A., Sauter, N. K., Yachandra, V. K., Bergmann, U. & Yano, J. (2013). *Science*, **340**, 491–495.
Langford, J. I. & Wilson, A. J. C. (1978). *J. Appl. Cryst.* **11**, 102–113.
Leocmach, M. & Tanaka, H. (2012). *Nat. Commun.* **3**, 974.
Lewus, R. A., Levy, N. E., Lenhoff, A. M. & Sandler, S. I. (2015). *Biotechnol. Prog.* **31**, 268–276.
Liu, W., Wacker, D., Gati, C., Han, G. W., James, D., Wang, D. J., Nelson, G., Weierstall, U., Katritch, V., Barty, A., Zatsepin, N. A., Li, D. F., Messerschmidt, M., Boutet, S., Williams, G. J., Koglin, J. E., Seibert, M. M., Wang, C., Shah, S. T. A., Basu, S., Fromme, R., Kupitz, C., Rendek, K. N., Grotjohann, I., Fromme, P., Kirian, R. A., Beyerlein, K. R., White, T. A., Chapman, H. N., Caffrey, M., Spence, J. C. H., Stevens, R. C. & Cherezov, V. (2013). *Science*, **342**, 1521–1524.
McPherson, A. (2004). *Methods*, **34**, 254–265.
Miao, J., Hodgson, K. O. & Sayre, D. (2001). *Proc. Natl Acad. Sci. USA*, **98**, 6641–6645.
Miao, J., Ishikawa, T., Robinson, I. K. & Murnane, M. M. (2015). *Science*, **348**, 530–535.
Michinomae, M., Mochizuki, M. & Ataka, M. (1999). *J. Cryst. Growth*, **197**, 257–262.
Ng, J. D., Lorber, B., Witz, J., Théobald-Dietrich, A., Kern, D. & Giegé, R. (1996). *J. Cryst. Growth*, **168**, 50–62.
Nielsen, S. S., Møller, M. & Gillilan, R. E. (2012). *J. Appl. Cryst.* **45**, 213–223.
Nielsen, S. S., Toft, K. N., Snakenborg, D., Jeppesen, M. G., Jacobsen, J. K., Vestergaard, B., Kutter, J. P. & Arleth, L. (2009). *J. Appl. Cryst.* **42**, 959–964.
Sabin, J., Bailey, A. E., Espinosa, G. & Frisken, B. J. (2012). *Phys. Rev. Lett.* **109**, 195701.

Salemme, F. R., Genieser, L., Finzel, B. C., Hilmer, R. M. & Wendoloski, J. J. (1988). *J. Cryst. Growth*, **90**, 273–282.

Scherrer, P. (1918). *Nachr. Ges. Wiss. Göttingen*, **1918**, 98–100.

Spence, J. C. H., Weierstall, U. & Chapman, H. N. (2012). *Rep. Prog. Phys.* **75**, 102601.

Svergun, D., Barberato, C. & Koch, M. H. J. (1995). *J. Appl. Cryst.* **28**, 768–773.

Tenboer, J., Basu, S., Zatsepin, N., Pande, K., Milathianaki, D., Frank, M., Hunter, M., Boutet, S., Williams, G. J., Koglin, J. E., Oberthuer, D., Heymann, M., Kupitz, C., Conrad, C., Coe, J., Roy-Chowdhury, S., Weierstall, U., James, D., Wang, D. J., Grant, T., Barty, A., Yefanov, O., Scales, J., Gati, C., Seuring, C., Srajer, V., Henning, R., Schwander, P., Fromme, R., Ourmazd, A., Moffat, K., Van Thor, J. J., Spence, J. C. H., Fromme, P., Chapman, H. N. & Schmidt, M. (2014). *Science*, **346**, 1242–1246.

Tsurusawa, H., Russo, J., Leocmach, M. & Tanaka, H. (2017). *Nat. Mater.* **16**, 1022–1028.

Van Driessche, A. E. S., Van Gerven, N., Bomans, P. H. H., Joosten, R. R. M., Friedrich, H., Gil-Carton, D., Sommerdijk, N. & Sleutel, M. (2018). *Nature*, **556**, 89–94.

Walsh, M. A., Schneider, T. R., Sieker, L. C., Dauter, Z., Lamzin, V. S. & Wilson, K. S. (1998). *Acta Cryst.* D**54**, 522–546.

Zhang, T. H., Klok, J., Tromp, R. H., Groenewold, J. & Kegel, W. K. (2012). *Soft Matter*, **8**, 667–672.