Journal of
Applied Crystallography
ISSN 0021-8898

Received 4 February 2011
Accepted 12 May 2011

Novel approach to controlled protein crystallization through ligandation of yttrium cations

Fajun Zhang,[a] Georg Zocher,[b] Andrea Sauter,[a] Thilo Stehle[b,c]* and Frank Schreiber[a]*

[a]Institut für Angewandte Physik, Universität Tübingen, Auf der Morgenstelle 10, 72076 Tübingen, Germany, [b]Interfaculty Institute of Biochemistry, University of Tübingen, Hoppe-Seyler-Strasse 4, Germany, and [c]Department of Pediatrics, Vanderbilt University School of Medicine, Nashville, Tennessee 37232, USA. Correspondence e-mail: thilo.stehle@uni-tuebingen.de, frank.schreiber@uni-tuebingen.de

Crystal structure determination of macromolecules is often hampered by the lack of crystals suitable for diffraction experiments. This article describes a protocol to crystallize the acidic protein bovine β-lactoglobulin in the presence of yttrium to yield high-quality crystals that belong to a new space group. The yttrium ions not only are used to engineer the crystallization, but are an integral part of the crystal lattice and can therefore be used to solve the phase problem using anomalous dispersion methods. Protein crystallization conditions were first optimized using an experimental phase diagram in the protein and salt concentration plane. Crystal growth strongly depends on the position in the phase diagram, and the best crystals grow near the phase transition boundaries. The structure analysis demonstrates the specific binding of yttrium ions to surface-exposed glutamate and aspartate side chains contributed by different molecules in the crystal lattice. By bridging molecules in this manner, contacts between molecules are formed that enable the formation of a stable crystal lattice. The potential application of this strategy to the crystallization of other acidic proteins is discussed on the basis of the universal features of the phase behavior of these proteins and the interactions induced by multivalent ions.

1. Introduction

X-ray crystallography using high-quality single crystals remains the most successful and widely used method for the determination of the three-dimensional structure of macromolecules at the atomic level. With the advent of high-energy synchrotron sources and the development of sophisticated software to solve structures, the crystallization process itself is more and more becoming the rate-limiting step in a successful structure determination (Durbin & Feher, 1996; Piazza, 2004; Anderson & Lekkerkerker, 2002; George & Wilson, 1994). No currently available method allows the prediction of conditions where a specific protein will form single crystals. One of the major difficulties in protein crystallization is the loss of conformational entropy during lattice formation. Therefore, 'surface entropy reduction' approaches based on various binding partners, such as small molecule ligands, peptides and even large crystallization chaperones, remain a promising strategy to promote crystal formation. However, although such binding partners have been proven to be highly effective in a number of cases, no molecule that universally promotes crystallization exists (Koide, 2009).

Despite the importance of protein crystallization, our understanding of the physical mechanisms underlying this process is still limited. In particular, suitable physical control parameters that can be used to predict the occurrence of nucleation and crystallization are not well understood (ten Wolde & Frenkel, 1997). Protein–protein interactions play a crucial role for understanding crystallization, and indeed also protein-aggregation-related physiological diseases (Anderson & Lekkerkerker, 2002; Durbin & Feher, 1996; Piazza, 2004). Over the past two decades, substantial progress has been made in understanding the physical and chemical conditions in protein solutions (Zhang *et al.*, 2007, 2008, 2010; Ianeselli *et al.*, 2010; Bonneté *et al.*, 1999; Curtis *et al.*, 2001; Muschol & Rosenberger, 1995; Petsev & Vekilov, 2000; Shukla *et al.*, 2008; Stradner *et al.*, 2004; Tardieu *et al.*, 1999; Velev *et al.*, 1998; Gunton *et al.*, 2007; ten Wolde & Frenkel, 1997). George & Wilson (1994) first reported that the second viral coefficient, A_2, can be used to guide protein crystallization. They showed that, for a number of proteins, crystallization occurs when A_2 lies in a narrow regime ranging from -1×10^{-4} to -8×10^{-4} mol ml g^{-2}. Slightly negative values of A_2 indicate that protein–protein interactions are dominated by weak attractive forces. By contrast, strong attractive forces will usually lead to the formation of gel or amorphous aggregates instead of crystals.

The attractive forces between protein molecules undergoing crystallization are short ranged and significantly smaller than the size of the protein (Asherie *et al.*, 1996; Galkin & Vekilov, 2000*b*; Hagen & Frenkel, 1994; Lomakin *et al.*, 2003; Muschol & Rosenberger, 1997; Rosenbaum & Zukoski, 1996; ten Wolde & Frenkel, 1997; Vekilov, 2004). Simulations (ten Wolde & Frenkel, 1997) and theoretical calculations (Lutsko & Nicolis, 2006; Nicolis & Nicolis, 2003; Talanquer & Oxtoby, 1998) predict that a short-ranged attraction causes the gas–liquid phase transition point to move below the liquid–solid coexistence curve, resulting in a metastable liquid–liquid phase separation (LLPS) that alters both the equilibrium phase diagram and the crystal nucleation behavior. When approaching the metastable critical point, the critical nucleus first forms highly disordered, liquid-like droplets that eventually turn into crystalline material (ten Wolde & Frenkel, 1997). Theoretical considerations also suggest that the density fluctuation in protein or colloidal solutions could enhance nucleation events near the metastable liquid–liquid coexistence curve (Nicolis & Nicolis, 2003; Talanquer & Oxtoby, 1998; Lomakin *et al.*, 2003; Muschol & Rosenberger, 1997; Rosenbaum & Zukoski, 1996; Vekilov, 2004; Lutsko & Nicolis, 2006). In line with this prediction, the enhancement of the nucleation rate of lysozyme near the LLPS phase boundary has been observed (Galkin & Vekilov, 2000*a*,*b*).

In many cases, the attractive forces between proteins depend critically on surface charge distribution, and modulating these charges can lead to new types of interactions. We have recently studied the effect of ionic strength and the nature of the counter-ion on the effective protein–protein interactions and phase behavior in protein solutions (Ianeselli *et al.*, 2010; Zhang *et al.*, 2007, 2008, 2010). The addition of monovalent salt (NaCl) increases the ionic strength of the protein solutions and screens the surface charges. The effective interaction between proteins is then repulsive and controlled by the surface charge and the excluded volume (Ianeselli *et al.*, 2010; Zhang *et al.*, 2007). When a trivalent cation was used, a reentrant condensation (RC) phase behavior was observed in protein solutions, *i.e.* a condensed two-phase regime is sandwiched between two single-phase regimes (Zhang *et al.*, 2008, 2010). This phase behavior suggests that trivalent metal ions can modulate protein–protein interactions not only in terms of strength, but also by changing the sign of the interaction from repulsion to attraction and *vice versa*. The RC behavior can be explained by the effective charge inversion at the protein surface induced by ion binding (Zhang *et al.*, 2010). As it is likely to be a universal phenomenon, RC could be a useful tool to guide interactions for many proteins, with particular importance for improving protein crystallization, as shown below.

β-Lactoglobulin (BLG) from bovine milk is a small protein that contains 162 amino acids and has a molar mass of about 18.3 kDa (Elofsson *et al.*, 1997). Owing to its importance in the food industry, the phase behavior of BLG in aqueous solutions is particularly well characterized (Sawyer *et al.*, 1999; Piazza *et al.*, 2002; Qin *et al.*, 1998; Townend, Winterbottom & Timasheff, 1960; Sakurai *et al.*, 2001; Townend, Weinberger & Timasheff, 1960; Elofsson *et al.*, 1997; Brownlow *et al.*, 1997). The conformation and association behavior of BLG depend on pH (Elofsson *et al.*, 1997; Qin *et al.*, 1998) and ionic strength (Piazza *et al.*, 2002; Sakurai *et al.*, 2001). The protein starts to dimerize at pH values above 3.5, with the dimeric form dominating at neutral pH. The dimer reversibly dissociates into monomers below pH 3.5 (Townend, Winterbottom & Timasheff, 1960; Townend, Weinberger & Timasheff, 1960). With an isoelectric point of 5.2, the protein is acidic and carries a net charge of $-10e$ at neutral pH in its dimeric form (Elofsson *et al.*, 1997). Inspection of the Protein Data Bank (http://www.rcsb.org) shows that a total of 17 crystal structures of BLG have been determined to date. These can be grouped into three crystal forms that differ in packing interaction and crystal symmetry (Brownlow *et al.*, 1997).

Using BLG as a test case, we have performed a systematic study on modulating the effective interactions in protein solutions by varying protein and yttrium chloride concentrations. Our data show that analysis of the RC behavior of BLG can be used for the generation of a new crystal form that had not been reported previously. The structure determination of these crystals provides direct evidence for the role of multivalent cations such as Y^{3+} in protein crystallization. Crystal packing is largely dependent on Y^{3+} ions that bridge acidic side chains from neighboring protein subunits. Therefore, the Y^{3+} ions exploit the acidity of BLG and help to establish a crystal lattice that would not form in the absence of cations. As an added benefit, the incorporated yttrium ions can also be used to solve the phase problem through anomalous dispersion methods, eliminating the need to produce classical heavy-atom derivatives or selenomethionine-derivatized crystals.

2. Materials and methods

2.1. Materials

Globular BLG from bovine milk (L3908) and yttrium chloride (YCl_3) were purchased from Sigma–Aldrich. Sample solutions were prepared by mixing stock solutions of BLG (67 mg ml^{-1}) and YCl_3 (100 m*M*). The phase diagram (protein concentration c_p *versus* salt concentration c_s) was determined at room temperature ($\sim$295 K) by monitoring the optical transmission of a series of protein solutions containing different salt concentrations (Zhang *et al.*, 2008). The c_p values were determined by UV absorption using an extinction coefficient of 0.966 ml mg^{-1} at a wavelength of 280 nm (Sober, 1970). The ζ-potential values of the sample solutions were determined with light scattering by applying an alternating electrical field to a U-shaped capillary cell (Zetasizer-Nano, Malvern Instruments Ltd). Note that the presence of a high concentration of buffer {such as HEPES [4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid] and Tris buffer} can affect the phase behavior and the solubility of yttrium salts. However, with lower buffer concentration (about 5 m*M*), the effect on the solubility of yttrium salts is negligible. To avoid the effect of other ions, no buffer was used in this work for sample preparation.

2.2. Crystal growth mechanism and preparation of high-quality protein single crystals

Initial crystals were grown at 277 K using the hanging-drop method. Drops containing 5 µl of freshly prepared protein solution were placed above 1 ml reservoirs containing the same salt solution without protein. No other additives were used. Crystals typically grew in a period of 2–4 d. High-quality single crystals suitable for X-ray diffraction measurements were obtained by batch crystallization. For this, the protein solution (1.0 ml) containing the desired yttrium concentration (0.3, 3.0 and 4.0 m*M*) was incubated at 277 K for 7–14 d. Our approach can also be carried out using the hanging-drop method, which requires considerably less protein. Using the hanging-drop method, one can reduce the volume to 2 µl. For standard batch crystallization, our method is thus comparable to other methods, requiring about 1 mg of purified protein. Needle-like crystals of up to 5 mm in length were obtained after two weeks. Crystal bunches were transferred onto a silicon glass plate and cracked to obtain single crystals with dimensions up to 200 × 60 × 60 µm. To reduce osmotic stress the crystals were moved in four steps into a solution containing the identical yttrium concentration but supplemented with 30% glycerol. The crystals were flash frozen in liquid nitrogen and stored until data collection.

Data were recorded either at beamline PXIII of the Swiss Light Source (Villigen, Switzerland) or with a rotating anode generator (Micromax 007 HF, MSC) equipped with a mar345dtb (Marresearch). Structure determination, refinement, model building and manuscript preparation were carried out using the programs *XDS* and *XSCALE* (Kabsch, 1993), *SHELXD* (Sheldrick, 2008), *SHARP*/*autoSHARP* (Bricogne *et al.*, 2003), *ARP*/*wARP* (Langer *et al.*, 2008), *COOT* (Emsley & Cowtan, 2004; Emsley *et al.*, 2010), *REFMAC5* (Murshudov *et al.*, 1997), *PHENIX* (Adams *et al.*, 2004), *FFT* (Collaborative Computational Project, Number 4, 1994), *RAMPAGE* (Lovell *et al.*, 2003), *POVscript+* (Fenn *et al.*, 2003), and *POVRAY* (http://www.povray.org). For details on structural analysis as well as data collection and refinement statistics (Tables S1 and S2) see the supporting material.[1]

3. Results

3.1. Experimental phase diagram

We first determined the phase diagram of BLG (c_p) as a function of the YCl_3 concentration (c_s) at room temperature (Fig. 1*a*). This diagram provides a guide for optimizing the conditions for protein crystallization. For a given protein concentration c_p, an increase of the salt concentration c_s above a certain threshold (c^*) results in the protein solution becoming turbid and entering a two-phase state. When c_s is increased further (above c^{**}), the protein solution turns clear again. Thus, the two salt concentrations c^* and c^{**} divide the phase diagram into three regimes (Fig. 1*a*). Regimes I and III contain clear protein solution, whereas the protein condenses (or aggregates) in Regime II. The aggregation in Regime II is called 'reentrant condensation' (RC), which originally derives from the studies of DNA condensation in the presence of multivalent counter-ions (Grosberg *et al.*, 2002). In the case of DNA, addition of counter-ions leads to surface charge inversion, *i.e.* the initially net negative charge becomes positive at higher salt concentration. We reason that, given the acidity of BLG, the addition of YCl_3 might produce a similar charge inversion on the protein surface.

To examine the observed RC phase behavior in more detail, we next performed ζ-potential measurements of a BLG solution at 1.3 mg ml^{-1} as a function of c_s (Fig. 1*b*). For $c_s < c^*$, the protein molecules are negatively charged, as indicated by a negative ζ value. The value of ζ increases with increasing c_s, crosses the zero potential line and turns positive. This charge inversion corresponds well with the experimental phase diagram (Fig. 1*a*). We also carried out a second round of experiments with BLG at a concentration of 3.4 mg ml^{-1}.

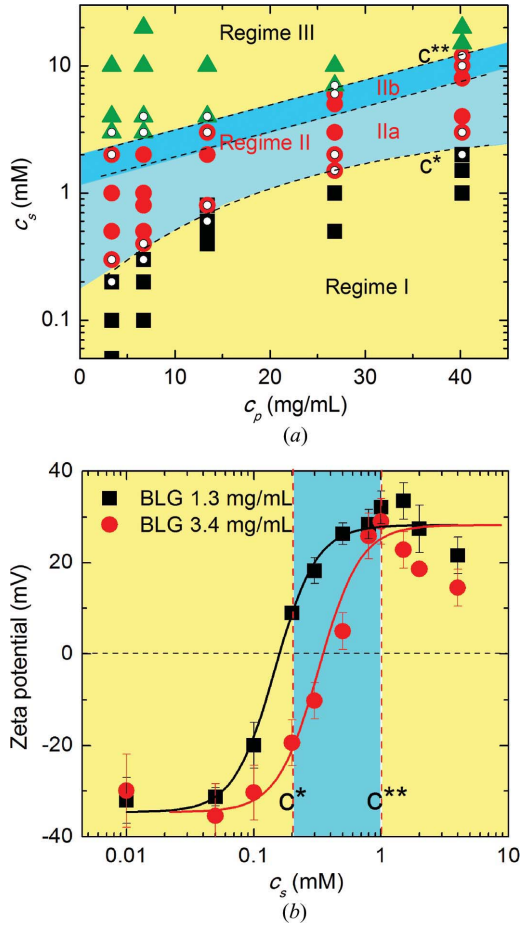

Figure 1
Phase behavior of BLG solutions in the presence of YCl_3. (*a*) Phase diagram at room temperature (295 K). Solid symbols present the sample solutions in different regimes (see text). Small open symbols present the samples solutions where crystallization was observed at 277 K. (*b*) ζ-potential measurement as a function of c_s for c_p of 1.3 and 3.4 mg ml^{-1}. For comparison, c^* and c^{**} for c_p = 3.4 mg ml^{-1} are also shown.

[1] The supplementary figure and tables discussed in this paper, as well as details of the experimental procedure, are available from the IUCr electronic archives (Reference: HE5524). Services for accessing this material are described at the back of the journal.

Here, the ζ potential is close to zero at c^*, indicating that the phase transition at c^* is mainly due to the neutralization of surface charge. However, c^{**} corresponds to a much higher positive ζ potential, probably because sufficient electrostatic repulsion is needed to re-stabilize the protein in solution. We conclude that regimes I and III are dominated by electrostatic repulsion, keeping BLG soluble, and regime II is dominated by an effective attractive potential, resulting in the condensation of protein molecules.

The RC phase behavior and the effective surface charge inversion (Fig. 1) suggest that the effective protein–protein interactions in solution can be modulated from repulsion to attraction and to repulsion again by simply varying c_s (YCl_3). This behavior could thus be very useful for promoting protein crystallization. Good crystallization conditions usually depend on a small attractive potential (George & Wilson, 1994). Therefore, one could expect the highest probability for protein crystallization in areas close to the phase boundaries, where the effective charges are low.

3.2. Temperature-dependent phase behavior and crystal growth

All samples used to determine the phase diagram were next used to perform crystallization experiments at 277 K. Samples that crystallized after one week are labeled in the phase diagram (Fig. 1*a*) with smaller open circles. As expected, crystallization preferentially occurred in regions close to the phase boundaries, in most cases below c^* and above c^{**}. These observations indicate that one should be able to prepare protein solutions at concentrations just below c^* or above c^{**} at room temperature, and then follow crystallization in real time by moving the solutions to lower temperatures, assuming the solubility of BLG is reduced at lower temperatures.

We therefore first evaluated the temperature-dependent phase behavior of the supernatants in Regime II. Regime II was further divided into two sub-regimes (Regimes IIa and IIb) because of the somewhat different phase behavior of the supernatants (Fig. 1*a*) as a function of temperature. In regime IIa, the supernatants of protein solutions (aggregates are removed by centrifugation) are clear, with a single phase present in the entire temperature window (277–313 K). At lower temperature, crystals grow directly from the homogeneous supernatant after several hours to days. In Regime IIb, protein solutions show a reversible temperature-dependent phase behavior. An example is shown in Fig. 2(*a*), which assays phase behavior by depicting light absorption at wavelengths ranging from 350 to 800 nm as a function of temperature. The insets show the same sample solution below and above a transition temperature at which the solution changes from turbid to clear. This procedure is reversible, with no significant differences after several heat–cool cycles. The absorption data were fitted using a sigmoidal function and the transition temperature, T_{tr}, was determined. T_{tr} decreases with initial salt concentration (Fig. 2*b*). For sample solutions with 3.4 and 6.7 mg ml^{-1} protein concentrations, this reversible aggregation occurs in the range of salt concentration of 1.2–2.4 m*M* and 1.8–3.2 m*M*, respectively. Crystallization of the supernatants occurs at temperatures below T_{tr}.

We next followed crystal growth from Regimes IIa and IIb by optical microscopy at 277 K using the hanging-drop method. As starting points we chose a c_p value of 6.7 mg ml^{-1} and c_s values of 3.0 and 0.3 m*M*. The c_s values are close to c^{**} (∼2.5 m*M*) and c^* (∼0.3 m*M*). Both solutions were clear at room temperature. Figs. 3(*a*)–3(*e*) show a series of photographs of drops containing 3.0 m*M* YCl_3 at 277 K. The hanging drops become turbid after several hours, and tiny droplets or aggregates are observed after 2 d (Fig. 3*b*). After 3 d, crystals start to form and grow homogeneously throughout the sample (Fig. 3*c*). The crystals are needle-like single crystals with homogeneous sizes of 0.1–0.2 mm in length (Figs. 3*c*–3*e*). No crystal growth was observed for the same solution at room temperature (above T_{tr}). This crystallization behavior cannot be explained by classical nucleation theory because in that case one would not observe such a metastable phase before crystallization (see Figs. 3*g* and 3*h*). Rather, a two-step growth mechanism is a more suitable explanation

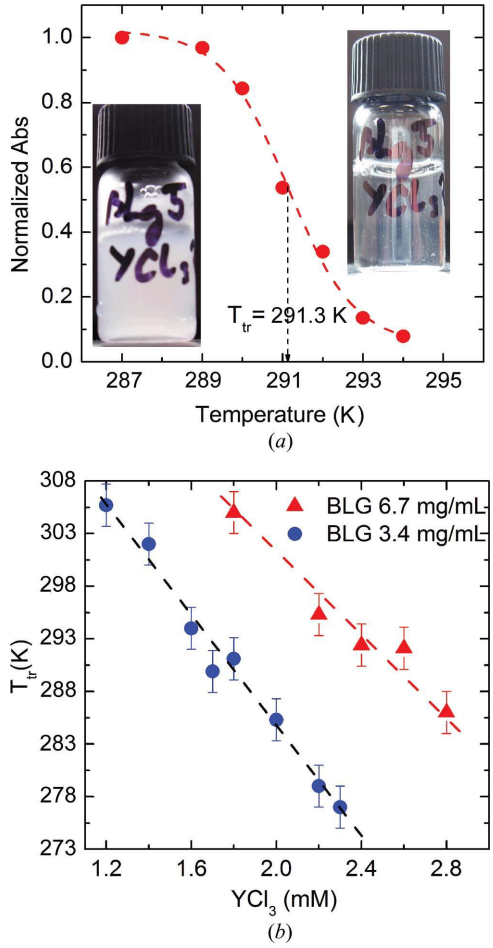

Figure 2
Reversible phase separation in Regime IIb. (*a*) Plot of UV–vis absorption as a function of temperature for the supernatant of BLG (3.4 mg ml^{-1}) with 1.8 m*M* YCl_3. The transition temperature, T_{tr} = 291.3 K, was determined by fitting a sigmoidal function; the insets show the corresponding sample solution below and above T_{tr}. (*b*) Plot of T_{tr} as a function of c_s for two series of sample solutions.

(Vekilov, 2004). For samples with low c_p values at the beginning of the experiment, the metastable phase forms clusters (Fig. 3*b*). For samples with high c_p values, a liquid–liquid coexistence phase can be observed (Fig. 3*f*). For instance, the sample solution comprising 33.5 mg ml^{-1} of BLG and 8.0 m*M* YCl_3 shows a mixture of liquid droplets and crystals after 3 d at 277 K. This two-step crystal growth process is in contrast to that found in the second sample solution with c_s = 0.3 m*M*, where crystals grow directly from heterogeneous nucleation and grow into a bundle of needle-like crystals directly from the homogeneous solution as predicted by classical nucleation theory (Figs. 3*g* and 3*h*).

3.3. Structural basis for crystal formation with YCl_3

Starting from the initial crystals, high-quality protein single crystals could be produced by optimizing the growth temperature as well as protein and salt concentrations (Fig. S1). These crystals were then used to collect diffraction data sets and determine structures of BLG. An initial structure was established using single anomalous dispersion data collected from crystals grown with 4.0 m*M* YCl_3. Data collection statistics are given in Table S1. The initial structure was then used to solve data sets from crystals grown at YCl_3 concentrations of 3.0 and 0.3 m*M*, respectively (Table S1). The final structures have excellent quality and allowed us to unambiguously locate bound yttrium ions as a result of their anomalous scattering behavior.

All crystals belong to space group $P2_12_12_1$ and contain one dimer in their asymmetric unit. In all cases, three fully occupied yttrium ions mediate crystal contacts that help to form the crystal lattice (Fig. 4*a*). A fourth yttrium site is occupied depending on the yttrium concentration. None of the 17 known crystal structures of BLG deposited in the Protein Data Bank has the crystal packing reported here. This provides evidence for a critical role of the yttrium ions in the generation of the lattice observed here. A packing analysis using *PISA* (Krissinel & Henrick, 2007) revealed that the biological dimer forms the largest interface by shielding a surface area of 470 Å^2 from the solvent. Six additional, smaller, intermolecular contacts, with buried surface areas ranging from 67 to 164 Å^2, account for the crystal packing. Each of the four yttrium ions is coordinated by acidic residues contributed by two protein molecules, and these contacts serve to form the crystal lattice (Fig. 4). The ionic interface built up by the metal ions is 552 Å^2 and is larger than the dimer contact of the BLG protomers.

To test whether the presence of yttrium affects the overall fold of BLG, we compared the structures reported here with previously determined BLG structures using *DALI* (Holm & Sander, 1993). As expected, we did not see a significant effect of the yttrium ions on the fold of a protomer. A more detailed superimposition yields Cα r.m.s. deviation values of 0.4–1.3 Å when superposing 151–161 Cα atoms of the main chain. These small values demonstrate that the BLG structure is identical to the previously determined ones. Most of the small differences are located at terminal residues or at the loop regions. As these regions are exposed and possess some flexibility, the differences are most likely due to the crystal packing of the different crystal forms.

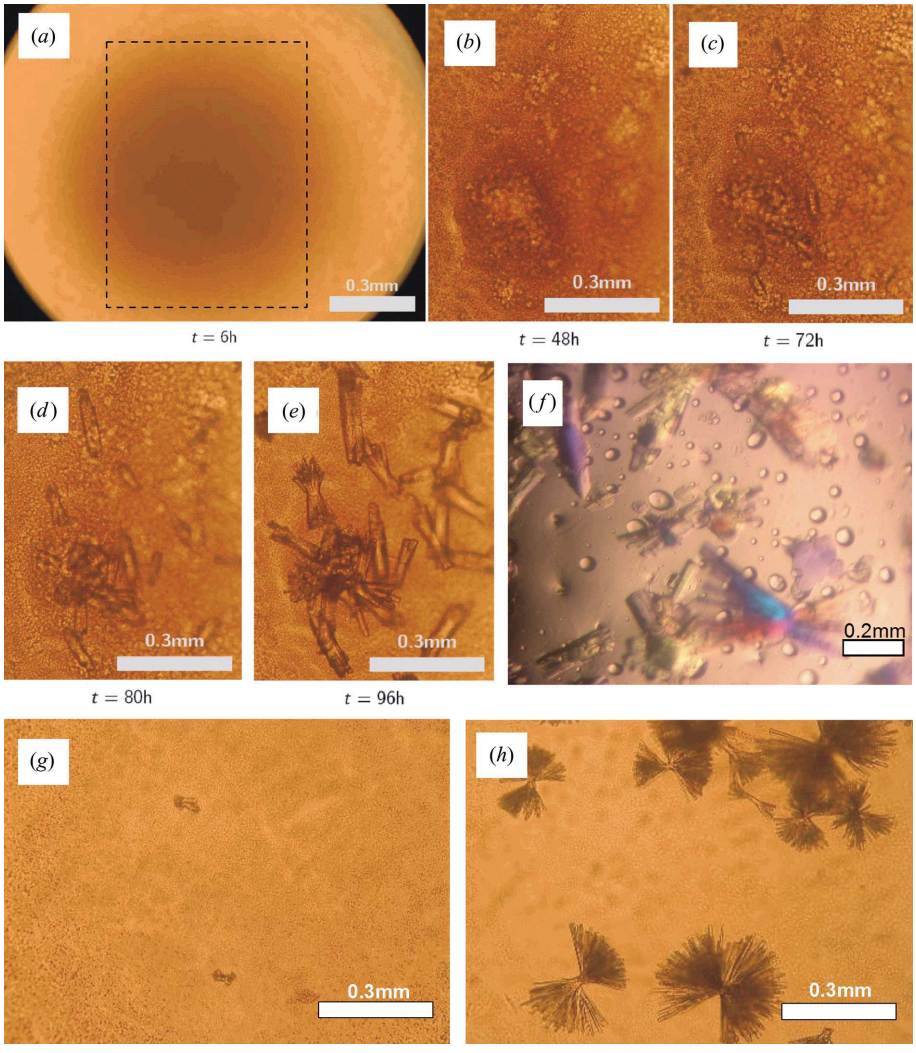

Figure 3
(*a*)–(*e*) Two-step crystal growth from reversible phase separation close to *c***: c_p = 6.7 mg ml^{-1} with c_s = 3.0 m*M* at 277 K at different times. (*f*) A snapshot of c_p = 33.5 mg ml^{-1} with c_s = 8.0 m*M* at 277 K, 3 d after sample preparation, where both dense liquid droplets and crystals were clearly seen. (*g*)–(*h*) One-step crystal growth from protein solution close to *c**: c_p = 6.7 mg ml^{-1} and c_s = 0.3 m*M* YCl_3 at 277 K after (*g*) 8 h and (*h*) 48 h.

A superimposition of structures obtained from Regimes IIa and IIb shows that the structures are identical (Cα r.m.s. deviation of 0.7 Å, 159 aligned residues). We then determined the positions and the relative occupancies of the yttrium ions in crystals grown using both regimes by calculation of electron density maps with the Bijvoet differences as coefficients (see Table S3). Three yttrium ions are coordinated by at least three acidic residues (Figs. 4*b*–4*d*). The ions have the same position and occupancy (close to 100%) in all crystals. Only the fourth yttrium, which is coordinated by residues D130 and E158 (Fig. 4*e*), exhibited a difference. This site is poorly occupied (∼30%) in crystals grown from Regime IIa, which used 0.3 m*M* YCl_3. By contrast, the occupancy of this ion is significantly higher (55%) in crystals grown with Regime IIb, which used 3.0 m*M* YCl_3.

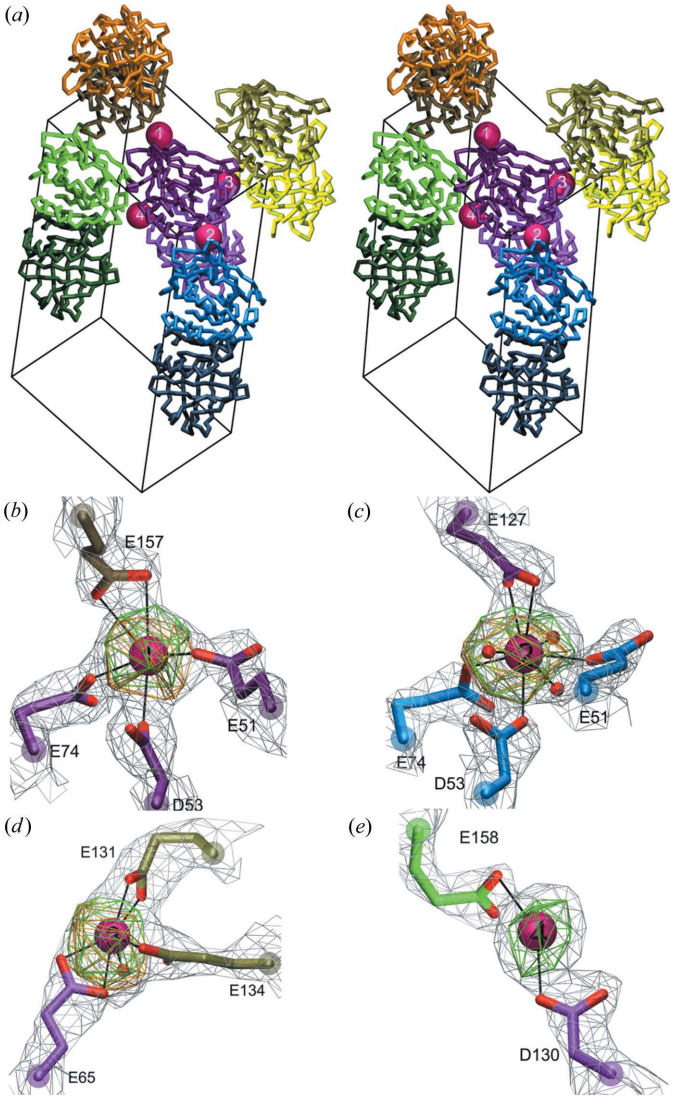

Figure 4
(*a*) Stereo view of the crystal packing, with the linking yttrium sites represented by spheres (magenta). Different asymmetric units are shown in different colors, where two chains of the dimer are represented by a dark and a light color. The very small proteinogenic crystal contacts that do not include yttrium ions are not shown. (*b*)–(*e*) Enlargements of (*a*), representing the coordination of the yttrium sites. The ($2F_{obs}$–F_{calc}) electron density map at a σ level of 1.3 and the anomalous difference maps for crystals grown from 6.7 mg ml^{-1} of BLG with 0.3 and 3.0 m*M* YCl_3 at a σ level of 5 are depicted in grey, green and orange, respectively.

The sites that ligate yttrium clearly have some specificity. Yttrium is always contacted by at least two and in most cases three acidic residues (E, D). While each BLG dimer contains 54 surface-exposed acidic residues, only a small fraction of these cluster in such a way that two or three acidic residues can together engage the same cation. The binding of multivalent metal ions at such 'acidic surface patches' clearly alters the charge distribution at the protein surface as each yttrium neutralizes three acidic side chains. This change in charge pattern then leads to the formation of new crystal contacts.

The bound cations bridge two molecules since, for each cation, there is one residue contributed from the neighboring protein (Fig. 4). Thus, both the specific binding of yttrium to an acidic surface patch as well as the bridge formation to a neighboring side chain are crucial for protein crystallization. One of the major reasons for the difficulties of protein crystallization is the low number of contacts in protein crystals in proportion to their molecular weight, and hence the relatively weak interaction produced by each crystal contact (Durbin & Feher, 1996). In the absence of yttrium or other cations, patches of acidic residues at the surface of a protein would probably be poor contact points for crystallization owing to their flexibility and similar charge. However, by engaging yttrium ions they can create a new, well ordered contact point at the protein surface. The orientation-dependent bridge interaction to another side chain can then significantly enhance the probability of contact formation between protein molecules.

4. Discussion

Crystallization of macromolecules remains a challenge, preventing an understanding of the three-dimensional structures of a large number of highly relevant proteins. New strategies to overcome this bottleneck are urgently needed. We describe here a protocol to obtain high-quality crystals for the acidic protein BLG by using trivalent metal ions and present an analysis of the phase behavior of the BLG/cation system. In addition to Y^{3+}, we have tested two other trivalent metal ions, La^{3+} and Al^{3+}. La^{3+} displays a behavior very similar to Y^{3+} and could thus also be used to initialize protein crystallization with our protocol. On the other hand, Al^{3+} does not behave similarly, which is mainly because of the strong hydrolysis property of Al^{3+} (data not shown). Our protocol could be of general use to promote protein crystallization in difficult cases. Advantages of our strategy are that it (i) requires no additives except the cation, (ii) works with relatively low protein concentrations and (iii) has predictable phase behavior. These factors significantly reduce the time and effort that is required to optimize protein crystallization.

At a given protein concentration, depending on the salt concentration, crystal growth can follow two different mechanisms: the one-step classical nucleation mechanism at $c_s \simeq c^*$, or a two-step mechanism at $c_s \simeq c^{**}$. The crystal

growth observed in our protein–salt (YCl_3) mixtures can follow both of these mechanisms. The results are similar, as high-quality single crystals have been obtained from both regions. Crystal structure analysis demonstrates that the specific binding of yttrium to the acidic residues (E and D) causes an effective charge inversion at the protein surface, which is the driving force of the RC phase behavior. The binding sites and number of cations on the protein surface suggest that the multivalent metal ions act as bridges in protein crystallization. Our observations of crystal growth mechanisms near c^* and c^{**} are in good agreement with theoretical predictions that, far from the metastable critical point, the critical nucleation events follow the classical pathway, *i.e.* the nucleus has the same order and crystallinity as a stable bulk crystal. When approaching the metastable critical point, the nucleus forms highly disordered aggregates or liquid droplets (ten Wolde & Frenkel, 1997), leading to metastable LLPS. The two coexisting liquid phases are metastable with respect to the crystalline state. Over time, they decay to more stable solid phases such as crystals or (amorphous) aggregates (Vekilov, 2004).

It is worth noting that several proteins have already been crystallized using yttrium as an additive (Bouyain *et al.*, 2005; Guo *et al.*, 2002; Selmer *et al.*, 2002; Xu *et al.*, 2001). With the exception of SelB (Selmer *et al.*, 2002), which has a theoretical pI of 8.06, all proteins crystallized in the presence of yttrium are acidic. Investigation of the crystal packing of the basic SelB protein shows that its crystal lattice is mainly formed by two large protein–protein contacts, neither of which involves yttrium. Although the yttrium ions contribute some additional crystal contacts that probably serve to further stabilize the lattice, they are not the principal mediators of crystal packing in this case.

5. Conclusions

In summary, we show here that a globular BLG solution in the presence of multivalent counter-ions (Y^{3+}) exhibits a rich phase behavior, which can be successfully exploited for protein crystallization. The RC phase behavior allows for the modulation of effective protein interactions in solution as well as for the optimization of protein crystallization conditions. Considering the universality of this RC behavior in protein solutions induced by multivalent metal ions, and the comprehensive understanding on the role of the ions in protein crystal growth mechanisms, we expect that the method described here can be of general use. About 50% of proteins in mammalian genome sequences are predicted to be acidic, and thus the strategy presented here offers new approaches for the successful crystallization of at least some of these proteins (Zhang *et al.*, 2010).

6. Protein Data Bank accession code

The atomic coordinates of the models together with the structure factors have been deposited in the Protein Data Bank under codes 3ph5 and 3ph6.

We acknowledge contributions by A. Gallice in the early stage of this project and financial support from DFG. We thank the beamline scientists at the Swiss Light Source for their support.

References

Adams, P. D., Gopal, K., Grosse-Kunstleve, R. W., Hung, L.-W., Ioerger, T. R., McCoy, A. J., Moriarty, N. W., Pai, R. K., Read, R. J., Romo, T. D., Sacchettini, J. C., Sauter, N. K., Storoni, L. C. & Terwilliger, T. C. (2004). *J. Synchrotron Rad.* **11**, 53–55.
Anderson, V. J. & Lekkerkerker, H. N. (2002). *Nature* (*London*), **416**, 811–815.
Asherie, N., Lomakin, A. & Benedek, G. B. (1996). *Phys. Rev. Lett.* **77**, 4832–4835.
Bonneté, F., Finet, S. & Tardieu, A. (1999). *J. Cryst. Growth*, **196**, 403–414.
Bouyain, S., Longo, P. A., Li, S., Ferguson, K. M. & Leahy, D. J. (2005). *Proc. Natl Acad. Sci. USA*, **102**, 15024–15029.
Bricogne, G., Vonrhein, C., Flensburg, C., Schiltz, M. & Paciorek, W. (2003). *Acta Cryst.* D**59**, 2023–2030.
Brownlow, S., Morais Cabral, J. H., Cooper, R., Flower, D. R., Yewdall, S. J., Polikarpov, I., North, A. C. & Sawyer, L. (1997). *Structure*, **5**, 481–495.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Curtis, R. A., Blanch, H. W. & Prausnitz, J. M. (2001). *J. Phys. Chem. B*, **105**, 2445–2452.
Durbin, S. D. & Feher, G. (1996). *Annu. Rev. Phys. Chem.* **47**, 171–204.
Elofsson, U. M., Paulsson, M. A. & Arnebrant, T. (1997). *Langmuir*, **13**, 1695–1700.
Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.
Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* D**66**, 486–501.
Fenn, T. D., Ringe, D. & Petsko, G. A. (2003). *J. Appl. Cryst.* **36**, 944–947.
Galkin, O. & Vekilov, P. G. (2000*a*). *J. Am. Chem. Soc.* **122**, 156–163.
Galkin, O. & Vekilov, P. G. (2000*b*). *Proc. Natl Acad. Sci. USA*, **97**, 6277–6281.
George, A. & Wilson, W. W. (1994). *Acta Cryst.* D**50**, 361–365.
Grosberg, A. Y., Nguyen, T. T. & Shklovskii, B. I. (2002). *Rev. Mod. Phys.* **74**, 329–345.
Gunton, J. D., Shiryayev, A. & Pagan, D. L. (2007). *Protein Condensation – Kinetic Pathways to Crystallization and Disease.* New York: Cambridge University Press.
Guo, F., Esser, L., Singh, S. K., Maurizi, M. R. & Xia, D. (2002). *J. Biol. Chem.* **277**, 46753–46762.
Hagen, M. H. J. & Frenkel, D. (1994). *J. Chem. Phys.* **101**, 4093–4097.
Holm, L. & Sander, C. (1993). *J. Mol. Biol.* **233**, 123–138.
Ianeselli, L., Zhang, F., Skoda, M. W., Jacobs, R. M., Martin, R. A., Callow, S., Prévost, S. & Schreiber, F. (2010). *J. Phys. Chem. B*, **114**, 3776–3783.
Kabsch, W. (1993). *J. Appl. Cryst.* **26**, 795–800.
Koide, S. (2009). *Curr. Opin. Struct. Biol.* **19**, 449–457.
Krissinel, E. & Henrick, K. (2007). *J. Mol. Biol.* **372**, 774–797.
Langer, G., Cohen, S. X., Lamzin, V. S. & Perrakis, A. (2008). *Nat. Protoc.* **3**, 1171–1179.
Lomakin, A., Asherie, N. & Benedek, G. B. (2003). *Proc. Natl Acad. Sci. USA*, **100**, 10254–10257.
Lovell, S. C., Davis, I. W., Arendall, W. B. III, de Bakker, P. I., Word, J. M., Prisant, M. G., Richardson, J. S. & Richardson, D. C. (2003). *Proteins*, **50**, 437–450.
Lutsko, J. F. & Nicolis, G. (2006). *Phys. Rev. Lett.* **96**, 046102.
Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.

Muschol, M. & Rosenberger, F. (1995). *J. Chem. Phys.* **103**, 10424–10432.
Muschol, M. & Rosenberger, F. (1997). *J. Chem. Phys.* **107**, 1953–1962.
Nicolis, G. & Nicolis, C. (2003). *Physica A*, **323**, 139–154.
Petsev, D. N. & Vekilov, P. G. (2000). *Phys. Rev. Lett.* **84**, 1339–1342.
Piazza, R. (2004). *Curr. Opin. Colloid Interf. Sci.* **8**, 515–522.
Piazza, R., Iacopini, S. & Galliano, M. (2002). *Europhys. Lett.* **59**, 149–154.
Qin, B. Y., Bewley, M. C., Creamer, L. K., Baker, H. M., Baker, E. N. & Jameson, G. B. (1998). *Biochemistry*, **37**, 14014–14023.
Rosenbaum, D. F. & Zukoski, C. F. (1996). *J. Cryst. Growth*, **169**, 752–758.
Sakurai, K., Oobatake, M. & Goto, Y. (2001). *Protein Sci.* **10**, 2325–2335.
Sawyer, L., Kontopidis, G. & Wu, S. Y. (1999). *Int. J. Food Sci. Technol.* **34**, 409–418.
Selmer, M., Wilting, R., Holmlund, D. & Su, X.-D. (2002). *Acta Cryst.* D**58**, 1871–1873.
Sheldrick, G. M. (2008). *Acta Cryst.* A**64**, 112–122.
Shukla, A., Mylonas, E., Di Cola, E., Finet, S., Timmins, P., Narayanan, T. & Svergun, D. I. (2008). *Proc. Natl Acad. Sci. USA*, **105**, 5075–5080.
Sober, H. A. (1970). *Handbook of Biochemistry: Selected Data for Molecular Biology*. Cleveland: The Chemical Rubber Co.
Stradner, A., Sedgwick, H., Cardinaux, F., Poon, W. C., Egelhaaf, S. U. & Schurtenberger, P. (2004). *Nature* (*London*), **432**, 492–495.
Talanquer, V. & Oxtoby, D. W. (1998). *J. Chem. Phys.* **109**, 223–227.
Tardieu, A., Le Verge, A., Malfois, M., Bonneté, F., Finet, S., Riès-Kautt, M. & Belloni, L. (1999). *J. Cryst. Growth*, **196**, 193–203.
Townend, R., Weinberger, L. & Timasheff, S. N. (1960). *J. Am. Chem. Soc.* **82**, 3175–3179.
Townend, R., Winterbottom, R. J. & Timasheff, S. N. (1960). *J. Am. Chem. Soc.* **82**, 3161–3168.
Vekilov, P. G. (2004). *Cryst. Growth Des.* **4**, 671–685.
Velev, O. D., Kaler, E. W. & Lenhoff, A. M. (1998). *Biophys. J.* **75**, 2682–2697.
Wolde, P. R. ten & Frenkel, D. (1997). *Science*, **277**, 1975–1978.
Xu, H. E., Lambert, M. H., Montana, V. G., Plunket, K. D., Moore, L. B., Collins, J. L., Oplinger, J. A., Kliewer, S. A., Gampe, R. T., McKee, D. D., Moore, J. T. & Willson, T. M. (2001). *Proc. Natl Acad. Sci. USA*, **98**, 13919–13924.
Zhang, F., Skoda, M. W., Jacobs, R. M., Martin, R. A., Martin, C. M. & Schreiber, F. (2007). *J. Phys. Chem. B*, **111**, 251–259.
Zhang, F., Skoda, M. W., Jacobs, R. M., Zorn, S., Martin, R. A., Martin, C. M., Clark, G. F., Weggler, S., Hildebrandt, A., Kohlbacher, O. & Schreiber, F. (2008). *Phys. Rev. Lett.* **101**, 148101.
Zhang, F., Weggler, S., Ziller, M., Ianeselli, L., Heck, B. S., Hildebrandt, A., Kohlbacher, O., Skoda, M. W. A., Jacobs, R. M. J. & Schreiber, F. (2010). *Proteins Struct. Funct. Bioinformatics*, **78**, 3450–3457.