Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Crystallization of canavalin as a function of pH

Elizabeth L. Forsythe,[a] Sridhar Gorti[b] and Marc L. Pusey[b]*

[a]BAE Systems, Hunstville, AL, USA, and [b]Biophysics SD46, NASA/MSFC, Huntsville, AL 35812, USA

Correspondence e-mail: marc.pusey@msfc.nasa.gov

Received 22 July 2004
Accepted 3 March 2005

The question was posed of what happens to a protein that is known to grow as an *n*-mer when it is placed in solution conditions where it is monomeric. The trypsin-treated or cut form of the protein canavalin (CCAN) has been shown to nucleate and grow crystals as a trimer from neutral to slightly acidic solutions. Under these conditions the solution is composed almost wholly of trimers. The crystalline protein can be readily dissolved by weakly basic solution, which has been proposed to result in a solution that is monomeric. Results are presented for crystallization experiments of CCAN over the pH 6.4–9.6 range. CCAN was found to crystallize in the canonical rhombohedral form at all pH values. Light scattering and size-exclusion chromatography data gave no indication of monomeric CCAN. The light-scattering data also showed that trimeric CCAN slowly self-associates to form larger species. Results are presented for crystallization experiments of CCAN over the pH 6.4–9.6 range. Fluorescence anisotropy, coupled with light-scattering and gel-filtration experiments, show that the solutions are primarily trimers, with association to form larger species occurring as a function of protein concentration. These results indicate that CCAN crystal nucleation occurs by the self-association of trimers to form progressively larger species in solution, which eventually become critically sized crystal nuclei.

1. Introduction

The mechanism(s) by which protein crystals nucleate and grow are not certain. We have proposed, at least for the tetragonal form of chicken-egg lysozyme, that solution-phase assembly processes are needed to form the growth units for crystal nucleation and growth (Pusey & Nadarajah, 2002). The starting point for the self-association process is the monomeric protein and the final crystallographic symmetry is defined by the interactions of the monomers and subsequent *n*-mers formed. When growth rate and solubility are measured, it has been found that many proteins grow from highly supersaturated solutions, suggesting that this may be a common mechanism for biological macromolecular crystal nucleation and growth. It has been suggested that multimeric proteins generally incorporate the underlying multimer's symmetry into the final crystallographic symmetry (Wukovitz & Yeates, 1995), in which case the associated species in solution would be the growth unit and not the monomeric protein.

An alternative theory for nucleation and growth postulates a system that is essentially always monomeric, with nucleation occurring through a liquid–liquid phase-separation process. In this view, the oversaturated solution spontaneously forms a highly concentrated protein-rich phase, which then, as they are below the critical phase transition concentration, decomposes back to the bulk solution (Berland *et al.*, 1992; Muschol & Rosenberger, 1997; Galkin & Vekilov, 2000). Nucleation is postulated to occur within these phases, which serve to concentrate the solute molecules, thus facilitating the subsequent formation of critical nuclei. These nuclei emerge from the concentrated phases and then grow by addition of monomers from the solution.

As part of our investigations into the protein crystal nucleation process, we posed the question of what happens to a protein that is known to grow as an *n*-mer when it is placed in solution conditions

where it has been shown to be a monomer? The trypsin-treated or cut form of the protein canavalin (CCAN) has been shown to nucleate and grow several crystal forms, all having a trimer as the common basic structural unit, from neutral to slightly acidic solutions (Ko *et al.*, 1993). Under these conditions the solution is composed almost wholly of trimers (Kadima *et al.*, 1990, 1991). The crystalline protein can be readily dissolved by weakly basic solution, which is postulated to result in a monomeric protein solution (Kadima *et al.*, 1990). There are three possible outcomes to an attempt at crystallization of the protein under these postulated monomeric conditions: (i) we will obtain the same crystals as under trimer conditions, but at different protein concentrations, (ii) we will obtain crystals having a different symmetry or (iii) we will not obtain crystals. Obtaining the first result would be indicative that the solution-phase self-association process is critical to the crystal nucleation and growth process. The third result would suggest that it is not or would be indicative of pH effects on the self-association process. The second result would be less clear, as it may reflect a pH-dependent shift in the trimer–trimer interactions or a different growth pathway by monomer assembly, which could be indicated by an absence of trimers in the structure.

Below, we report on the results of crystallization experiments and solution-phase characterization by fluorescence polarization, light scattering and gel-filtration chromatography of CCAN over the pH range 6.4–9.6. The data show that, contrary to our starting hypothesis, CCAN apparently does not dissociate into monomers as a function of pH. However self-association of the trimers, a critical part of the nucleation process, was clearly observed even in undersaturated solutions.

2. Materials and methods

Sodium chloride, ADA, HEPES, sodium phosphate and Tris buffers were from Sigma Chemical Co. (St Louis, MO, USA). Unless otherwise mentioned, buffers were prepared by dissolving the appropriate amounts of the buffering species and the salt to ~95% of the final solution volume. The pH was then adjusted to the desired value with the buffer counter-ion and the volume adjusted to the final amount with distilled water (dH_2O).

2.1. Purification and preparation of CCAN

Trypsin-treated or cut canavalin (CCAN) was prepared as previously described from Jack beans (Sumner & Howell, 1936; Smith *et al.*, 1982). The protein was further purified by recrystallization, the procedure involving dissolution of the crystals into a minimal amount of dilute ammonium hydroxide solution, keeping the pH $\leq$ 10. This solution was clarified by centrifugation, then crystallized by dialysis against 0.05 *M* sodium phosphate, 0.15 *M* NaCl pH 6.8. The recrystallization procedure was repeated at least three times, with the final crystals collected by centrifugation and stored at 253 K. CCAN concentrations were determined by UV absorbance, using a mass extinction coefficient of A_{280} = 0.438 ml mg^{-1}.

2.2. Crystallization of CCAN

Crystallizations were set up in Cryschem sitting-drop plates, using 500 µl of reservoir solution, and sealed with clear tape. Each plate was set up with four identical rows, with the NaCl concentration varying from 0 to 0.5 *M* in 0.1 *M* increments across the row. The reservoir buffer concentrations were 0.1 *M* in either ADA (pH 6.4, 6.8, 7.2) or Tris–HCl (pH 7.8, 8.2, 8.8, 9.2 and 9.6). The crystalline protein was dissolved as above, then concentrated using a Centriprep centrifugal ultrafiltration cell at 298 K. The concentrated CCAN solution was diluted with dH_2O to prepare stock solutions having concentrations of 15, 25, 50, 100, 150, 200, 300 and 400 mg ml^{-1}. The volume ratio of protein to reservoir solution in the crystallization drops was varied down the rows for each crystallization plate, being 1:1, 2:1, 3:1 and 4:1 protein:reservoir solution for rows 1 to 4, respectively. The same protein solution was used throughout for any one plate. The plates were incubated at 293 K and examined on a daily basis.

2.3. Fluorescence anisotropy analysis of the CCAN solutions

CCAN was labeled on the N-terminal amines with the fluorescent probe pyrene acetic acid succinimidyl ester (pyr; Molecular Probes, catalog No. P6115). The fluorescent probe was dissolved in dimethyl formamide. Basically, the crystalline protein was solubilized by dissolution in weak base solution, washed several times by diafiltration with dH_2O, and the calculated volume of fluorescent probe solution added with stirring to give an estimated 25% derivatization, based upon the measured protein concentration. After ~30 min incubation at room temperature in the dark, the protein solution was concentrated then passed down a G-50 column equilibrated in 0.01 *M* HEPES, 0.02 *M* NaCl pH 7.5. The protein fractions were pooled, concentrated, and stored at 277 K. Pyrene absorbs at 280 nm and the ratio of the free probe absorption at 280 and 346 nm was used to determine a corrected protein absorption value at 280 nm. After correction the protein was found to be 18% derivatized.

The steady-state fluorescence anisotropy value, *R*, was calculated using the equation:

$$R = (I_{VV} - G^* I_{VH})/(I_{VV} + 2^* G^* I_{VH}), \quad (1)$$

where I_{VV} and I_{VH} are the blank corrected measured intensities with subscripts H and V referring to the horizontal and vertical orientation of the excitation and emission polarizers, respectively. *G*, an optical correction factor, was calculated from:

$$G = I_{HV}/I_{HH}. \quad (2)$$

As *G* is an optical correction factor, dependent upon the instrumentation characteristics, it only has to be determined one time for a given set of optical conditions.

Fluorescence anisotropy measurements were made at room temperature, using an SLM spectrofluorometer, with the excitation and emission slits set at 2 nm and the excitation and emission monochromators set at 340 and 376 nm, respectively. Three 0.1 *M* buffer solutions with 0.2 *M* NaCl, HEPES pH 7.0 and Tris–HCl pH 8.0 and 9.0 were prepared. Solutions for the assay were prepared by mixing equal volumes of buffer plus dH_2O or buffer plus protein solution. For a typical measurement series, the first measurements were made on the buffer blank, followed by one or more additions and measurements at low concentrations of pyr-CCAN, followed by a series of additions and measurements of increasing concentrations of CCAN. The anisotropy data obtained were plotted against the log of the protein concentration, calculated using a monomeric MW of 42 kDa (McPherson, 1980).

2.4. Gel-filtration analysis of CCAN

A Separose S-200 column (1.5 × 60 cm) was equilibrated with 0.01 *M* HEPES, 0.1 *M* NaCl pH 7.0. CCAN solutions of 20, 5 and 1 mg ml^{-1} were prepared in the same buffer by dilution of concentrated stock CCAN solutions. Aliquots of 1 ml of the protein solution were carefully washed onto the S-200 column, which was then eluted at a flow rate of ~3 ml min^{-1} using a peristaltic pump. Fractions of 40 drops were collected. The column was calibrated using the proteins

catalase (250 kDa), hemoglobin (66 kDa) and ovotransferrin (80 kDa) (Sigma), all prepared by dissolution of the protein directly into the column buffer.

2.5. Quasi-elastic light-scattering (QLS) analysis of CCAN

QLS measurements were made to provide a direct comparison with the data of Kadima *et al.* (1990), as well as to verify the size of the protein species eluted from the gel-filtration chromatography experiments. QLS measurements were performed using a He–Ne laser (Spectra Physics model 147), a photon-counting module (EG&G model SPCM-AQ) and a correlator (Brookhaven Instruments Corporation model 9600). Light exiting the laser (wavelength λ = 633 nm) was focused into a single-mode polarization preserving optical fiber, using a laser-to-fiber source coupler (Wave Optics), which serves as a light guide. The output end of the optical fiber was coupled to a lens to collimate the exiting polarized light into a sample placed within a temperature-controlled holder. The output end of the optical fiber was affixed to the sample holder that rests on a goniometer (Newport model URM100CC). The photon-counting module placed at a fixed location on an optical table collects light scattered by the sample that had passed through two precisely aligned pinholes placed in the path of the scattered light. The correlator computes the intensity correlation function from the TTL pulse train generated by light impinging upon the photon-counting module. A computer then stores and analyzes the intensity correlation functions of the scattered light.

Protein concentrations of 0.5 and 2.0 mg ml^{-1} from peak fractions Nos. 36 and 32, the peak trailing edge and center fractions, respectively, collected from the S-200 column on two different occasions were directly used in QLS measurements after filtration with 0.22 µm pore-size Nalgene filters. The experiments were conducted using cylindrical cells that contained approximately 200–300 µl of the filtered sample solutions. The cylindrical cells were placed in the scattering chamber filled with refractive-index matching fluid and the temperature of the chamber was controlled with Peltier devices and set to 293.1 $\pm$ 0.1 K. All light-scattering measurements were performed at a scattering angle of 40°. The intensity correlation functions, $c(\tau)$, obtained from QLS measurements are given as (Berne & Pecora, 1976),

$$c(\tau) = \langle I(t)I(t+\tau)\rangle_t, \tag{3}$$

where the brackets denote time, t, average and $I(t)$ represents the time-dependent fluctuations of the intensity of scattered light. At the detector, the observed fluctuations in intensity are a measure of the fluctuations in the average magnitude of the Poynting vector expressed in terms of the scattered electric field $E(t)$ as $I(t) \simeq E(t)E^*(t)$, according to Poynting's theorem (Jackson, 1962). The autocorrelation function, $c(\tau)$, is then rewritten in terms of the normalized electric field correlation function, $g(\tau)$, as (Berne & Pecora, 1976),

$$C(\tau) = I^2 + \beta I^2 g^2(\tau), \tag{4}$$

where I^2 is the square of the average scattered light intensity observed at a scattering angle φ and an instrumental constant β, which accounts for both a finite sample time and a finite scattering angle (Jakeman & Pike, 1968; Jakeman *et al.*, 1970). For the instrument used in this laboratory, β was determined as ~0.8.

2.6. Solubility of CCAN

Solubility of CCAN in 0.05 *M* HEPES, 0.1 *M* NaCl pH 7.0 was determined by dialysis. Briefly, a concentrated CCAN solution was crystallized by dialysis against the indicated buffer, producing a mass of small crystals in the dialysis bag. Dialysis was carried out in a 300 ml water jacketed beaker, with the solution temperature controlled by a circulating bath/chiller. The dialysis buffer was replaced with fresh solution and stirring was continued for the duration of the experiment. At periodic intervals the dialysis bag was opened and an aliquot of protein solution removed, after which the bag was resealed and dialysis continued. The aliquot removed was briefly centrifuged, diluted with dH_2O and the protein concentration determined by absorbance at 280 nm. The liquid volume was maintained in the bag by periodic additions of dialysis buffer. The bags were inverted at least one time between each measurement, to allow the microcrystals to fall through the solution and promote equilibration with the solution. After an aliquot was removed for quantification, the temperature was adjusted by increments of 1–2 K. Measurements were taken daily, over the temperature range of 289–297 K.

Figure 1
CCAN crystals grown at pH 6.4 (left) and pH 9.6 (right). The pH 6.4 crystals were grown from 0.1 *M* ADA, 0.4 *M* NaCl pH 6.4, using 25 mg ml^{-1} protein with a crystallization drop at a protein:reservoir solution ratio of 3:1. The pH 9.6 crystals were grown from 0.1 *M* Tris–HCl pH 9.6 using 200 mg ml^{-1} protein with a crystallization drop having a protein:reservoir solution volume ratio of 4:1.

Figure 2
Rod-shaped crystals of CCAN, obtained at 0.1 *M* Tris–HCl, 0.3 *M* NaCl pH 8.2, using 150 mg ml^{-1} protein solution with a crystallization drop having a protein:reservoir solution volume ratio of 2:1.

3. Results

3.1. Crystallization of CCAN

Crystals were obtained at all pH values tested and they generally appeared within 1–3 d. As the pH was increased, the canonical rhombohedral crystals were anticipated to be obtained at progressively higher starting protein concentrations. In reality, most wells had too much protein and resulted in a large mass of crystals in the drop. In some cases, however, limited nucleation occurred and single, or just a few, crystals were obtained. In instances where high starting concentrations were employed these crystals often grew rather large. Examples of crystals grown at both ends of the pH spectrum are shown in Fig. 1. Higher salt concentrations resulted in CCAN crystals having a different morphology as shown in Fig. 2. X-ray diffraction data (not shown) indicated the same space group and unit-cell dimensions as the hexagonal form of canavalin (Ko *et al.*, 1993). These did not appear at all pH values, but were present at the highest pH values, generally as dense mats of rods or needle-shaped crystals.

3.2. Fluorescence anisotropy results

Pyrene (pyr) has a long fluorescence lifetime, which depending upon the solution conditions may be over 100 ns (Andreoni *et al.*, 1994; Pan & Berglund, 1997). These lifetimes are well matched to the anticipated molecular masses, the rotational rates, of the CCAN monomer, trimer, and higher order aggregates. Fluorescence anisotropy data were collected over the concentration range 4×10^{-6} to 6×10^{-4} *M* protein, calculated on the basis of the monomeric molecular weight, at pH 7.0, 8.0 and 9.0. Solution containing pyr-CCAN was only added to 10^{-5}*M*, then only unlabeled protein added thereafter. The data obtained is plotted in Fig. 3 using (1) and shows that the anisotropy is essentially the same for all protein at low concentrations and that the self-association leading to crystals occurs over the same concentrations at all pH values. No attempts were made to correct the data for motion of the probe independent of that of the protein. Anisotropy data collected using CCAN that had been labeled with carboxyrhodamine qualitatively agreed with the pyr-CCAN data, although as would be expected for a shorter lifetime probe species the total range in anisotropy values was considerably less (data not shown).

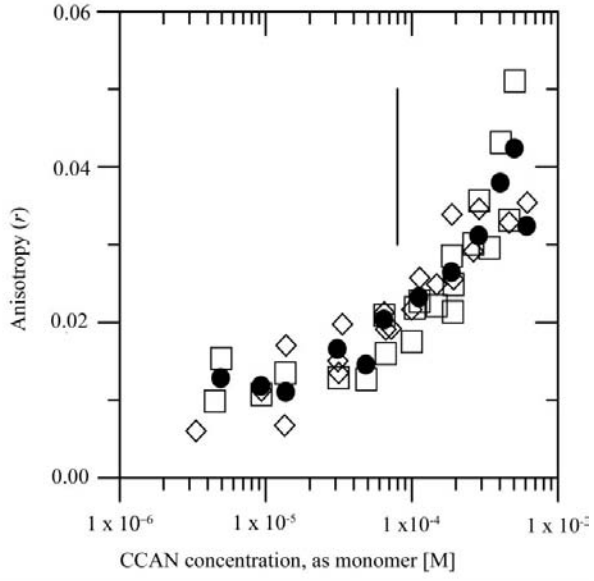

Figure 3
Fluorescence anisotropy of CCAN. A starting solution of pyr-CCAN was progressively titrated with unlabeled CCAN and the anisotropy determined as a function of the total protein concentration. Diamonds, pH 7.0; squares, pH 8.0; circles, pH 9.0. The vertical line shows the solubility of CCAN at 293 K, 0.05 *M* HEPES, 0.1 *M* NaC pH 7.0.

3.3. Solubility

Solubility data was obtained by the dialysis method. CCAN crystals dissolve very slowly at pH 7.0, and the measurements were repeated until a constant value of 3.9 mg ml^{-1} was obtained at 293 K. The solubility is indicated on Fig. 3 by a vertical line. Variations of the temperature by ± 4 K showed little change in the solubility.

3.4. S-200 gel-filtration experiments

The anisotropy data suggested that CCAN remains as a trimer at all pH values employed. This was further tested using size-exclusion chromatography. CCAN solutions were made up at 20, 5 and 1 mg ml^{-1} (monomer concentrations of 4.8×10^{-4}, 1.2×10^{-4} and 2.4×10^{-5} *M*, respectively) in 0.05 *M* HEPES, 0.1 *M* NaCl pH 7.0 buffer. 1 ml aliquots of the CCAN solutions were passed through the S-200 column at a pump speed of ~3.0 ml min^{-1}, fractions of 40 drops were collected and assayed for protein by absorption at 280 nm. The results are shown in Fig. 4. The column was subsequently re-equilibrated in 0.05 *M* Tris, 0.1 *M* NaCl pH 9.0 and a second 1.0 mg ml^{-1} aliquot of CCAN, prepared by dilution of a stock solution into this buffer, was chromatographed at the higher pH. This data, superimposed on that collected at pH 7.0 on Fig. 4, shows that the dilute protein solution at pH 9.0 coelutes with that at pH 7.0, indicating that the size is equivalent. From the molecular-weight markers, the elution volume is consistent with what would be expected for the CCAN trimer in all instances.

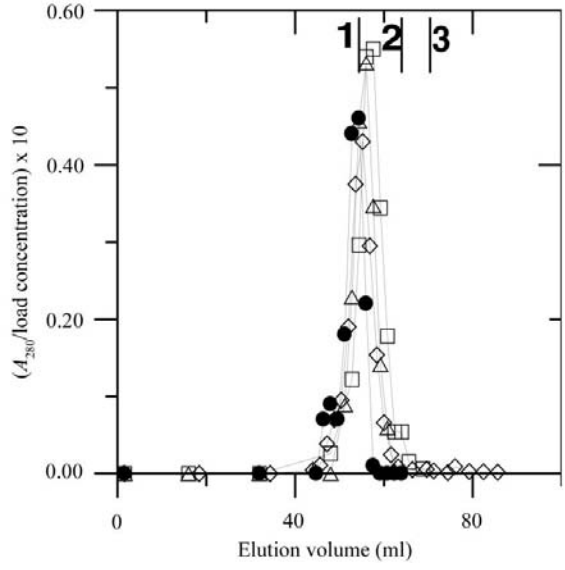

Figure 4
S-200 gel-filtration experiments for CCAN as a function of concentration and pH. 1 ml aliquots of protein at starting concentrations of 20 mg ml^{-1} (diamonds), 5 mg ml^{-1} (squares) and 1 mg ml^{-1} (triangles), were chromatographed on an S-200 column (1.5 × 60 cm) equilibrated in 0.05 *M* HEPES, 0.1 *M* NaCl pH 7.0. Data points denoted by circles are for protein at 1 mg ml^{-1} chromatographed on the same column, which was re-equilibrated in 0.05 *M* Tris–HCl, 0.1 *M* NaCl pH 9.0. The three vertical lines refer to the elution peak locations for (1) catalase (MW = 250 kDa), (2) ovotransferrin (MW = 80 kDa) and (3) hemoglobin (MW = 64 kDa).

3.5. Light-scattering analysis

CCAN was used as eluted from the S-200 gel filtration column. Peak fraction 36, a peak trailing edge fraction containing 0.5 mg ml^{-1} protein, and fraction 32, the peak center fraction containing 2.0 mg ml^{-1} protein, were used. Quasi-elastic light-scattering measurements were performed on these two solutions, as well as dilutions of fraction No. 36, and the resulting intensity correlation functions obtained at a scattering angle of 40° for fractions Nos. 32 and 36 are shown in Fig. 5. In all cases, the intensity correlations exhibited unimodal behavior that can be characterized by a distribution of exponentially relaxing processes. For the ideal case of solutions containing mono-disperse particles, the normalized electric field correlation function $g(\tau)$ decays as a simple exponential: $g(\tau) = \exp(-Dq^2\tau)$. Here, D is the translational diffusion coefficient of the particles, and q is the magnitude of the wavevector defined by the scattering angle φ and the wavelength λ of the laser as: $q = 4\pi n/\lambda\sin(\theta/2)$. In the present experiments, q was approximately 9.03×10^4 cm^{-1}.

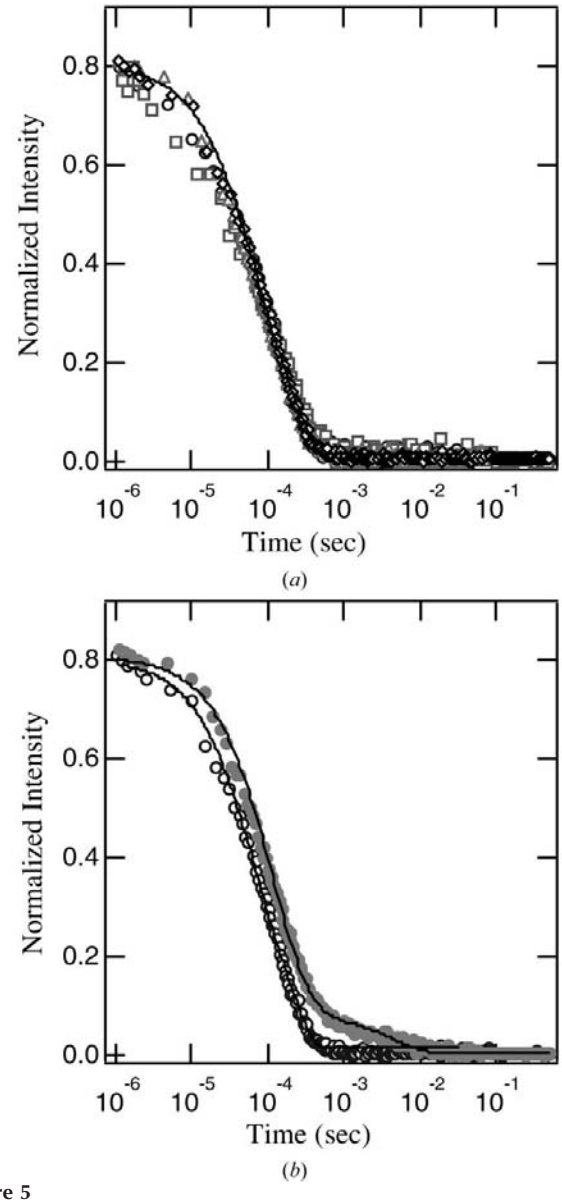

Figure 5
All intensity correlation functions were obtained at a scattering angle of 40° and temperature of 293 K. In (*a*), the intensity correlation functions were obtained from four different protein concentrations: 0.5, 1.0, 1.5 and 2.0 mg ml^{-1}, represented by squares, circles, triangles and diamonds, respectively. In (*b*), the intensity correlation functions were obtained at a single protein concentration, 2.0 mg ml^{-1}, , immediately as well as ∼10 h after sample preparation. The solid lines are fit to data using either the method of cumulants or a double-exponential method (see text). The average diffusion coefficient of the data is $\langle D\rangle = (6.3 \pm 0.6) \times 10^{-7}$cm^2 s^{-1}, which yields an average hydrodynamic radius of 3.4 ± 0.3 nm. A second population of scatterers observed by the intensity correlation function is indicated by the shoulder at ∼10^{-3}–10^{-2} s in (*b*) and have an average hydrodynamic radius of ∼100 nm.

A common complication in QLS measurements is the determination of sample polydispersity from the observed intensity correlation functions. At present, accurate determinations of the distribution of scatterers are not possible, due to low signal-to-noise ratios of the observed intensity correlation functions (see Fig. 5*a*). For practical considerations, the method of cumulants for the electric field correlation function $g(\tau)$ was employed, which fitted well all the observed data, yielding meaningful and consistent results. The method of cumulants by Koppel (1972) is a widely used numerical technique that determines the average diffusion coefficient $\langle D\rangle$ of polydisperse solutions, valid for short times, which defines $g(\tau)$ as

$$g(\tau) = \exp\left[-\langle D\rangle q^2\tau + \sum_{i=2}^{n}\frac{(-1)K_i\tau^i}{i!}\right], \quad (5)$$

where the K_is are higher order coefficients of the expansion. In the analyses of the intensity correlation functions, only the coefficients K_2 and K_3 that represent the variance and skewness of the distribution, respectively, were additionally used.

Using the method of cumulants, the averaged diffusion coefficients, $\langle D\rangle$, obtained from the individual correlations functions were $(6.5 \pm 0.7) \times 10^{-7}$, $(6.2 \pm 0.3) \times 10^{-7}$, $(5.7 \pm 0.4) \times 10^{-7}$ and $(6.4 \pm 0.1) \times 10^{-7}$ cm^2 s^{-1} for sample concentrations of 0.5, 1.0, 1.5 and 2.0 mg ml^{-1}, respectively. The average hydrodynamic radius obtained from the above diffusion coefficients, determined by using the Stokes–Einstein relation, is 3.4 ± 0.3 nm. The Stokes–Einstein relation for the diffusion coefficient, D, is defined as (Berne & Pecora, 1976) $D = k_BT/6\pi\eta a$, where k_BT is the thermal energy, η is the solution viscosity and a is the particle radius. Any perceived difference in the determined averaged diffusion coefficients or hydrodynamic radii with respect to protein concentration are statistically insignificant. Also shown in Fig. 5(*b*) is the behavior of the measured intensity correlation functions obtained, for the sample concentration of 2.0 mg ml^{-1}, both immediately as well as approximately 10 h after sample preparation. Most notably, the intensity correlation function obtained ∼10 h after sample preparation exhibits characteristics of a bimodal distribution of scatterers, simply labeled as fast (f) and slow (s). As a result, the electric field correlation function $g(\tau)$ is defined by

$$g(\tau) = A_f\exp(-D_fq^2\tau) + A_s\exp(-D_sq^2\tau), \quad (6)$$

where D_f and D_s and the amplitudes A_f and A_s represent the diffusion coefficients and relative contributions of each population of scatterers, respectively. Using this expression to characterize the intensity correlation function obtained ∼10 h after sample preparation, the apparent hydrodynamic radii of the fast and slow populations of scatterers were determined to be 4.0 and 100 nm, respectively.

4. Discussion

This project was initiated to test the role of solution-phase assembly processes in macromolecular crystallization. Based upon previous work (Kadima *et al.*, 1990), it was thought that more basic pH resulted in dissociation of the trimer to monomers. Accordingly, protein crystallizations carried out at higher pH values would probe the

relevance of the monomer-–trimer assembly process to formation of the crystallization growth units. The initial crystallization results were in accord to what was expected, that crystallization could be carried out at higher pH values, but only by elevating the protein concentration to drive formation of the postulated trimer growth unit.

The conclusion of Kadima *et al.* (1990) that trimeric CCAN dissociates to monomers at basic pH was called into question with the results from the fluorescence anisotropy experiments. If the monomer–trimer association process was pH-dependent, then the expected results would be three distinct curves, with the increase in anisotropy occurring at progressively higher concentrations with increased pH. Instead, the anisotropy data at higher pH values were co-linear with that collected at pH 7.0, which indicated that the self-assembly process was not pH-dependent. The next step, then, was to determine what the baseline species at lower concentrations was prior to the self-assembly step (the concentration range beginning at 10^{-6} to 10^{-5} *M* in Fig. 3). Gel-filtration experiments unambiguously showed that the baseline species was not the monomer, but instead most likely the trimer.

The light scattering data of Kadima *et al.* (1990) gives a measured trimer radius of 3.6–4.05 nm, compared with the 3.4 nm radius determined in this work. Upon standing, the solution was found to progressively become aggregated, resulting in one containing scatterers of ∼4.0 and ∼100 nm radius. The second value is large compared with the 23.8 nm diameter particles found by Kadima *et al.* (1990), which is close to the 23.4 nm dimension estimated for eight canavalin trimers in a rhombohedral unit cell arrangement (McPherson, 1980; Kadima *et al.*, 1990). While the particle sizes for the trimer are comparable in this and the work of Kadima *et al.* (1990), the solution behavior is not. We found the solution to be relatively stable with respect to the trimer as a function of decreasing protein concentration. Kadima *et al.* (1990) found that their solution became monomeric (radius = 1.8 nm), with sizes smaller than this as the solution became more dilute. Further, they found that the monomer aggregates over a relatively narrow protein concentration range, between 0.15 and 0.4% (1.5 and 4 mg ml^{-1}, their Fig. 1), rapidly progressing through the trimer to the larger unit cell sized species, where it plateaus. We did not find any indication of this, although our fluorescence anisotropy data does indicate that the protein is self-associating over the much broader 10^{-5} to 6×10^{-4} *M* concentration range (0.47 to ∼28 mg ml^{-1}). Cuvettes with the terminal anisotropy experiment solutions left in them developed rhombohedral crystals overnight at all three pH values.

Kadima *et al.* (1990) concluded that CCAN undergoes a pH-dependent dissociation process, from the trimer to the monomer, as the pH is raised above 7.0, with the protein being essentially all monomer by pH 8.0. They also suggest that the barrier to crystal nucleation is trimer formation from the monomer. These conclusions are not supported by our findings. No indication of trimer dissociation to the monomer, either as a function of decreased protein concentration or of increased pH, has been found in the work reported herein. However, trimer association to larger species has been observed, as reported by Kadima *et al.* (1990). Additionally, and more interesting, is that the self-association of the trimers to larger species occurs at virtually the same concentrations for all three pH values tested in this work and that it is apparent at concentrations below the measured crystalline solubility.

In retrospect, it is perhaps not surprising that there is no pH-dependence, over the range tested, for the monomer–trimer equilibrium. X-ray crystallographic analysis of CCAN shows that the monomer–monomer contacts which form the trimer are predominately hydrophobic, in other words, should not be pH-sensitive (Ko *et al.*, 1993). Similar results are found for the kidney bean vicilin phaseolin (Lawrence *et al.*, 1994). Structural homology studies for the peanut vicilin, Ara h 1, indicate that again the monomer–monomer contacts are primarily hydrophobic interactions (Maleki *et al.*, 2000). Analysis of the monomer–monomer contacts for the vicilin soybean β conglycinin indicated that nonpolar groups accounted for 65% of the buried surface area in the contact region, indicating that hydrophobic interactions are important to trimer formation (Maruyama *et al.*, 2001). Fluorescence anisotropy experiments with intact Ara h 1 showed that monomer–trimer association occurred in the 10^{-7} *M* concentration range and that the process was not sensitive to high salt concentration (Maleki *et al.*, 2000). No indication was obtained in our results for CCAN undergoing a monomer-trimer association process at similar concentrations, which we speculate may be a result of the proteolytic cleavage process. A more valid comparison may be obtained from similar studies on intact, or non-protease treated, canavalin. CCAN may dissociate to monomers at lower concentrations than $\sim 4 \times 10^{-6}$ *M*, although data collected to as low as 9.2×10^{-8} *M* at pH 7.0 (not shown) did not show any evidence of this.

However, one result is made very clear by the anisotropy experiment. Crystal nucleation of CCAN proceeds by the concentration dependent self-association of growth units in the solution. This self-association is clearly apparent well before the measured solubility. Another mechanism for increased anisotropy would be increased solution viscosity, due to added protein. However, the protein concentrations at which increased anisotropy becomes apparent are below 1×10^{-4} M, or less than 5 mg ml^{-1}. While there may be some increase in viscosity, it is not sufficient to account for the anisotropy shift observed.

We thank Dr Joseph Ng, University of Alabama, Huntsville for the collection of X-ray diffraction data.

References

Andreoni, A., Bottiroli, G., Colasanti, A., Giangare, M. C., Riccio, P., Roberti, G. & Vaghi, P. (1994). *J. Biochem. Biophys. Methods*, **29**, 157–172.
Berland, C. R., Thurston, G. M., Kondo, M., Broide, M. L., Pande, J., Ogun, O. & Benedek, G. B. (1992). *Proc. Natl Acad. Sci. USA*, **89**, 1214–1218.
Berne, J. B. & Pecora, R. (1976). *Dynamic Light Scattering.* New York: Wiley.
Galkin, O. & Vekilov, P. G. (2000). *Proc. Natl Acad. Sci. USA*, **97**, 6277–6281.
Jackson, J. D. (1962). *Classical Electrodynamics.* New York: Wiley.
Jakeman, E., Oliver, C. J. & Pike, E. R. (1970). *J. Phys. A*, **3**, L45–L48.
Jakeman, E. & Pike, E. R. (1968). *J. Phys. A*, **1**, 128–138.
Kadima, W., McPherson, A., Dunn, M. F. & Jurnak, F. A. (1990). *Biophys. J.* **57**, 125–132.
Kadima, W., McPherson, A., Dunn, M. F. & Jurnak, F. (1991). *J. Cryst. Growth*, **110**, 188–194.
Ko, T. P., Ng, J. D. & McPherson, A. (1993). *Plant Physiol.* **101**, 729–744.
Koppel, D. E. (1972). *J. Chem. Phys.* **57**, 4814–4820.
Lawrence, M. C., Izard, T., Beuchat, M., Blagrove, R. J. & Colman, P. M. (1994). *J. Mol. Biol.* **238**, 748–776.
McPherson, A. (1980). *J. Biol. Chem.* **255**, 10472–10480.
Maleki, S. J., Kopper, R. A., Shin, D. S., Park, C. W., Compadre, C. M., Sampson, H., Burks, A. W. & Bannon, G. A. (2000). *J. Immunol.* **164**, 5844–5849.
Maruyama, N., Adachi, M., Takahashi, K., Yagasaki, K., Kohno, M., Takenaka, Y., Okuda, E., Nakagawa, S., Mikami, B. & Utsumi, S. (2001). *Eur. J. Biochem.* **268**, 3595–3604.
Muschol, M. & Rosenberger, F. (1997). *J. Chem. Phys.* **107**, 1953–1962.
Pan, B. & Berglund, K. A. (1997). *J. Cryst. Growth*, **171**, 226–235.
Pusey, M. & Nadarajah, A. (2002). *Cryst. Growth Des.* **2**, 475–483.
Smith, S. C., Johnson, S., Andrews, J. & McPherson, A. (1982). *Plant Physiol.* **70**, 1199–1209.
Sumner, J. B. & Howell, S. F. (1936). *J. Biol. Chem.* **113**, 607–610.
Wukovitz, S. W. & Yeates, T. O. (1995). *Nature Struct. Biol.* **2**, 1062–1067.