Effects of Protein Purity and Precipitant Stereochemistry on the Crystallization of Thaumatin†

Neer Asherie,* Charles Ginsberg, Arieh Greenbaum, Samuel Blass, and Sarah Knafo

Department of Physics and Department of Biology, Yeshiva University, New York, New York 10033-3312

Received June 12, 2008; Revised Manuscript Received August 21, 2008

ABSTRACT: Thaumatin is frequently used as a model protein in crystallization studies because it rapidly forms crystals in the presence of tartrate ions. The thermodynamic and kinetic properties of thaumatin crystals have been studied for almost 10 years, and the results are contradictory. Here we show that by using a homogeneous preparation of thaumatin and controlling the stereochemistry of the tartrate precipitant, it is possible to achieve consistent results for the protein solubility. To understand the role of protein impurities in the crystallization of thaumatin, we examined two commercial sources of the protein and characterized the heterogeneities therein. To examine the effect of precipitant stereochemistry, we crystallized thaumatin with L, D, and DL (racemic) tartrate ions. We suggest that the inconsistencies among previous results stem in part from the different behavior of thaumatin with the L and D enantiomers: the solubility of thaumatin crystals increases with temperature in L-tartrate, whereas it decreases with temperature in D-tartrate. Our results demonstrate the importance of using pure protein and stereochemically pure precipitants in the crystallization of proteins.

1. Introduction

It has been more than 160 years since the first protein crystals were reported,1 but crystallizing proteins is still a challenge.2,3 The inability to predict the conditions under which a given protein crystallizes inhibits progress on many biophysical problems.4 The best-known problem is the production of crystals for the X-ray structure determination. The trial and error approach to crystallizing proteins has led to the “crystallization bottleneck”: less than 20% of purified proteins produce diffraction-quality crystals.5

Protein crystallization also dominates other phenomena.6 Several diseases, including genetic cataracts7 and hemoglobin SC disease,8 are caused by the formation of protein crystals in vivo. The effective release of drugs such as insulin depends on the properties of the crystals used.9 Finally, crystallization is used in industry to purify proteins,10,11 and the processing constraints often require crystals of a specific size and habit.12 A fundamental understanding of protein crystallization is essential for better analyzing and controlling these processes.

One approach to gaining insight into protein crystallization is to use model proteins.13 The most widely studied model protein is lysozyme, because it is commercially available in large quantities from several vendors at an affordable price and can be crystallized under many different conditions and in many different habits.1,14 Other model proteins include the γ-crystallins,15 bovine pancreatic trypsin inhibitor,16 ferritin,17 urate oxidase,18 and glucose isomerase.19

Another model protein is thaumatin. The thaumatins are a family of single-domain, highly homologous globular proteins (207 amino acids, eight disulfide bonds, molecular weight approximately 22 kDa) found in the fruit of Thaumatococcus daniellii.20,21 The two most abundant components are thaumatin I and thaumatin II, which differ by five amino acids.22 The thaumatins are among the sweetest compounds known—about 100 000 times sweeter than sugar on a molar basis23—and their physical and chemical properties (in particular, those related to taste) have been studied extensively since thaumatin I and II were first isolated in 1972.20,24–26

The rise of thaumatin as a model protein began soon after a discovery by Ko et al.: the addition of L-tartrate ions to a solution of thaumatin leads to the rapid formation of bipyramidal protein crystals.27 Since this discovery, numerous investigators have used the thaumatin that is commercially available from Sigma-Aldrich in combination with tartrate ions to examine different aspects of protein crystallization and X-ray structure determination. The topics that have been studied include crystallization in microgravity,28 in gels,29 and under controlled hydrostatic pressure,30,31 crystallization mechanisms,32,33 new crystallization
platforms;34–38 radiation damage;59 cryocoolling;30 and optimization of X-ray data collection.41
The most important quantity in the crystallization of any protein is its solubility, i.e., the concentration at which the protein solution and a specific crystal phase are in equilibrium under a given set of conditions. Unless the solubility is exceeded, the protein will not crystallize. Although the crystallization of thaumatin has been studied for over a decade, the results for its solubility are contradictory. It has been reported that the solubility of thaumatin increases with temperature,30 decreases with temperature,42 or is essentially zero53.

In a recent communication,43 we reported our results for the solubility of purified thaumatin with tartrate and showed that the temperature dependence of the protein solubility varies with the chirality of the tartrate ion added. In this article, we provide a detailed description of our findings. We describe the purification and characterization of the thaumatin we use, which was obtained from Natex UK Limited. This is a different source from the one that has been used in almost all other studies, where the supplier was Sigma-Aldrich. We also examine the effect of precipitant stereochemistry on the crystallization of thaumatin by crystallizing the proteins with L, D, and DL (racemic) tartrate ions. We find that the solubility of thaumatin crystals increases with temperature in L-tartrate, while it decreases with temperature in D-tartrate. Thus, we suggest that the inconsistencies among previous results for the solubility may stem from two factors: the use of a heterogeneous preparation of thaumatin and the different behavior of the protein with the L and D enantiomers of tartrate.

2. Experimental Section

2.1. Materials. Thaumatin (batch DBDS004) was provided by Natex UK Limited, Sandy, UK, and was also purchased from Sigma-Aldrich, St. Louis, MO (cat. no. T7638; lot no. 108F0299). Acetic acid (cat. no. BP1185), sodium azide (cat. no. S227), monobasic sodium phosphate (cat. no. BP330), dibasic sodium phosphate (cat. no. BP332), sodium acetate (cat. no. BP334), Tris base (cat. no. BP512), sodium chloride (cat. no. BP512), and sodium hydroxide (cat. no. BP559) were purchased from Fisher Scientific, Pittsburgh, PA. L-Tartaric acid (cat. no. T3807, lot no. 035K3650 and 115K3712), D-tartaric acid (cat. no. T206, lot no. 10121AD), and sodium potassium L-tartrate (cat. no. S2377, lot no. 062K0010; cat. no. S6170, lot no. 035K01981) were purchased from Sigma-Aldrich. Deionized water was obtained from an E-pure 4-Module deionization system (Barnstead International, Dubuque, IA). All materials were used without further purification, except for the Natex thaumatin, which was purified as described below. Solutions were filtered through a Nalgene disposable 0.22 µm filter unit (Nalgene Nunc International, Rochester, NY) prior to use.

2.2. Protein Purification. Crude thaumatin from Natex UK Limited was dissolved at high concentration (about 100 mg/mL) in 275 mM sodium acetate buffer (pH 4.5, conductivity σ = 8.0 mS/cm) with 0.02% sodium azide. The solution was centrifuged at 31,000 g for 30 min to sediment any undissolved solids. The supernatant was purified with an AKTAprime-plus preparative scale low-pressure size-exclusion chromatography system (GE Healthcare, Piscataway, NJ) by isotropic elution at 1.5 mL/min from an XK26/100 column packed with Sephacryl S-200 HR resin (GE Healthcare, Piscataway, NJ). The buffer used was 275 mM sodium acetate (pH 4.5, σ = 8.0 mS/cm) with 0.02% sodium azide. The monomer, shoulder, and aggregate fractions were collected, but only the monomer fraction was used in the solubility studies.

To compare our results for the Sigma-Aldrich protein with those of other investigators, we adopted the common procedure of using this protein without further purification.28,30,33,42 Concentration measurements were carried out by UV–vis extinction spectroscopy on a Beckman-Coulter DU800 spectrophotometer. The extinction coefficient of thaumatin at 280 nm was taken to be ε0.1% = 1.27 mg mL−1 cm−1.28,42 Conductivity and pH measurements were performed using an Orion 4-Star conductivity and pH meter with a DuraProbe conductivity cell and a RossSure-Flow pH electrode (Thermo Fisher Scientific, Waltham, MA).

2.3. Size-Exclusion and Cation-Exchange Chromatography. Both size-exclusion and cation-exchange high-performance liquid chromatography (SE-HPLC and CE-HPLC, respectively) were carried out on a Beckman System Gold apparatus at a flow rate of 1 mL/min (Beckman Coulter, Fullerton, CA). SE-HPLC measurements were carried out by isotropic elution on a Superdex 75 10/300 GL column (GE Healthcare, Piscataway, NJ). The buffer used was 100 mM sodium phosphate buffer (pH 7.1, σ = 11.4 mS/cm) with 0.02% sodium azide. CE-HPLC measurements were carried out on a Bakerbond Wide-Pore (5 µm particle size) CBX column (Mallinkrodt Baker, Phillipsburg, NJ). The proteins were eluted in 20 mM Tris acetate containing 0.02% sodium azide (pH 6.5, σ = 1.5 mS/cm) with a salt gradient of 0–100% 5 M sodium acetate (pH 6.5, σ = 30.5 mS/cm) for 41 min.

2.4. Quasielastic Light Scattering. Quasielastic light scattering (QELS) was performed on a home-built optical system using a 35 mW He–Ne laser (Coherent, Santa Clara, CA), a custom-made scattering cell (Precision Detectors, Bellingham, MA), and a PD2000DLS5/18 256 channel correlator (Precision Detectors, Bellingham, MA). The scattering angle was 90° and all measurements were carried out at 20 ± 0.3 °C. The measured correlation functions were analyzed by a constrained regularization method as implemented in the PrecisionDeconvolve software (version 5.4) provided by Precision Detectors. This software computes the distribution of scattered intensity as a function of the diffusion coefficient. To convert the diffusion coefficient to a hydrodynamic radius,45 we took the viscosity of the solution to be that of water (1.002 mPa s). The QLS measurements were done in 275 mM acetate buffer (pH 4.5, σ = 8.0 mS/cm) with 0.02% sodium azide, and all samples were filtered through an Anotop 10 0.02 µm filter (Whatman, Maidstone, UK).

2.5. Electrospray Ionization Mass Spectroscopy. Samples for mass spectrometry were dialyzed into deionized water using Ultrafree-MC centrifugal concentrators (5 kDa membrane; Millipore, Billerica, MA) and then filtered through an Ultra-AC centrifugal filter (0.22 µm; Millipore, Billerica, MA). Mass spectrometry was performed at the Laboratory for Macromolecular Analysis and Proteomics at the Albert Einstein College of Medicine of Yeshiva University.

2.6. Precipitation Studies. Thaumatin (Sigma or purified Natex) was dialyzed in an Amicon ultrafiltration cell with a 10 kDa membrane (Millipore, Billerica, MA) into the buffer of interest. It was then filtered through a 0.22 µm filter (Millipore, Billerica, MA) and then filtered through an Ultra-AC centrifugal filter (0.22 µm; Millipore, Billerica, MA). Mass spectrometry was performed at the Laboratory for Macromolecular Analysis and Proteomics at the Albert Einstein College of Medicine of Yeshiva University.

To determine the solubility curves, crystals were placed at a fixed temperature (within 0.1 °C) in fresh stocks (without protein) of the solution of interest. As the crystals dissolved, the concentration of protein in solution was monitored by periodically removing aliquots of the supernatant. The solution was judged to be in equilibrium when the concentration reached a constant value (typically in 48 h). The
solutions crystal system was stirred continuously to ensure thorough mixing of the components. To check that our results were self-consistent, some solubility data was collected starting with supersaturated protein solutions, as well as undersaturated ones. For the solubility measurements, the concentration of protein could not be determined accurately by direct UV extinction spectroscopy at 280 nm, as the tartrate ions also absorb around 280 nm. Instead, SEC-HPLC was performed on each aliquot and the concentration of the protein was determined from the area of the well-resolved protein peak. By using stock solutions of known protein concentrations (range: 4 × 10−3 to 9 mg/ml), we confirmed that the concentration was proportional to the peak area. We found the same proportionality constant (to within 1%) for protein solutions without tartrate, with 0.5 M sodium L-tartrate, and with 0.5 M sodium D-tartrate. There are also low-molecular-weight contaminants past 270 min; yield about 26% for protein solutions without tartrate, with 0.5 M sodium L-tartrate, and with 0.5 M sodium D-tartrate.

3. Results and Discussion

3.1. Source of Protein. When we began our investigation of the phase behavior of thaumatin, we turned to the same source used by almost all other studies over the past decade, Sigma-Aldrich. To our dismay, we obtained irreproducible results. We discovered that the thaumatin from Sigma-Aldrich (henceforth “Sigma thaumatin”) that we and others had been using all came from the same batch (lot no. 108F0299), which was produced in October 1988.

As the Sigma protein is almost 20 years old, we decided against purifying it. Instead, we searched for an alternative source of thaumatin and found that fresh, industrial preparations of the protein are available from Natex UK Limited (henceforth “Natex thaumatin”); the batch we used (DBDS004) was produced in November 2003. By purifying the Natex thaumatin (as described in section 2.2), we were able to obtain pure, monomeric protein that gives reproducible results for the phase behavior. We show below that Natex thaumatin is more homogeneous than Sigma thaumatin.

3.2. Protein Purification and Characterization. Figure 1 shows a low-pressure size-exclusion chromatogram for a solution of Natex thaumatin. The large peak around 220 min is the thaumatin monomer whose yield is about 38% by weight. The shoulder fraction (approximately 110–210 min; yield about 26% by weight) contains monomer and protein oligomers, whereas the aggregate peak around 90 min (yield about 11% by weight) contains both protein and nonprotein material (most likely carbohydrate from the gel-coated seeds that contain thaumatin). There are also low-molecular-weight contaminants past 300 min, but we did not collect these. Only the monomer fraction was used in the solubility studies, so it was the only fraction to be characterized extensively.

Electrospray ionization mass spectrometry (ESIMS) of six separate preparations of Natex thaumatin gave an average mass of $M_r = 22190 ± 2$ Da for the monomer fraction. This result is consistent with the mass of thaumatin I ($M_r = 22188$ Da) as deduced from the published sequences, assuming that all disulfide bonds are intact. To further characterize the monomer fraction, we examined our samples by quasielastic light scattering (QLS), size-exclusion high performance liquid chromatography, and cation-exchange high performance liquid chromatography (SE-HPLC and CE-HPLC, respectively).

Figure 2 shows QLS results for purified Natex and unpurified Sigma thaumatin. The proteins were measured at the same concentration (2.6 mg/ml) and solution conditions (20 °C; 275 mM acetate buffer at pH 4.5 with 0.02% sodium azide). The average hydrodynamic radius for purified Natex thaumatin is 2.19 ± 0.02 nm (Figure 2A), which is consistent with the molecular mass of the monomeric protein. The purified protein is stable in solution at 4 °C for at least 3 months.

By contrast, the average hydrodynamic radius of Sigma thaumatin (Figure 2B) is larger (2.36 ± 0.05 nm) and the size distribution is broader than for Natex thaumatin. These QLS results suggest that Sigma thaumatin is less homogeneous than Natex thaumatin; this suggestion is confirmed by SE-HPLC (Figure 3A). Although Natex thaumatin (solid line) is essentially monomeric, Sigma thaumatin (dashed line) contains small amounts of oligomeric and low-molecular-weight contaminants. Another source of heterogeneity in Sigma thaumatin is revealed by CE-HPLC. As can be seen in Figure 3B, Natex thaumatin (solid line) is predominantly (> 80%) a single-charge species (major peak at 27 min). Sigma thaumatin (dashed line) contains two large peaks: one at 27 min, as in Natex thaumatin, and another at 31 min, where there is only a very small peak (less than 1%) in Natex thaumatin. ESIMS shows that the peak at 27 min has a molecular mass of 22190 ± 2 Da (Natex) and 22189 ± 2 Da (Sigma), consistent with that of thaumatin I (22 188 Da); the peak in Sigma thaumatin at 31 min corresponds to a mass of 22 272 ± 2 Da, consistent with that of thaumatin II (22 272 Da).

Finally, there is a heterogeneity that can be seen by eye. Sigma thaumatin is contaminated by a pigment that makes the solutions golden-yellow to reddish-brown; as the protein solution-crystal system was stirred continuously to ensure thorough mixing of the components. To check that our results were self-consistent, some solubility data was collected starting with supersaturated protein solutions, as well as undersaturated ones. For the solubility measurements, the concentration of protein could not be determined accurately by direct UV extinction spectroscopy at 280 nm, as the tartrate ions also absorb around 280 nm. Instead, SEC-HPLC was performed on each aliquot and the concentration of the protein was determined from the area of the well-resolved protein peak. By using stock solutions of known protein concentrations (range: 4 × 10−3 to 9 mg/ml), we confirmed that the concentration was proportional to the peak area. We found the same proportionality constant (to within 1%) for protein solutions without tartrate, with 0.5 M sodium L-tartrate, and with 0.5 M sodium D-tartrate. There are also low-molecular-weight contaminants past 270 min; yield about 26% for protein solutions without tartrate, with 0.5 M sodium L-tartrate, and with 0.5 M sodium D-tartrate.

3. Results and Discussion

3.1. Source of Protein. When we began our investigation of the phase behavior of thaumatin, we turned to the same source used by almost all other studies over the past decade, Sigma-Aldrich. To our dismay, we obtained irreproducible results. We discovered that the thaumatin from Sigma-Aldrich (henceforth “Sigma thaumatin”) that we and others had been using all came from the same batch (lot no. 108F0299), which was produced in October 1988.

As the Sigma protein is almost 20 years old, we decided against purifying it. Instead, we searched for an alternative source of thaumatin and found that fresh, industrial preparations of the protein are available from Natex UK Limited (henceforth “Natex thaumatin”); the batch we used (DBDS004) was produced in November 2003. By purifying the Natex thaumatin (as described in section 2.2), we were able to obtain pure, monomeric protein that gives reproducible results for the phase behavior. We show below that Natex thaumatin is more homogeneous than Sigma thaumatin.

3.2. Protein Purification and Characterization. Figure 1 shows a low-pressure size-exclusion chromatogram for a solution of Natex thaumatin. The large peak around 220 min is the thaumatin monomer whose yield is about 38% by weight. The shoulder fraction (approximately 110–210 min; yield about 26% by weight) contains monomer and protein oligomers, whereas the aggregate peak around 90 min (yield about 11% by weight) contains both protein and nonprotein material (most likely carbohydrate from the gel-coated seeds that contain thaumatin). There are also low-molecular-weight contaminants past 300 min, but we did not collect these. Only the monomer fraction was used in the solubility studies, so it was the only fraction to be characterized extensively.

Electrospray ionization mass spectrometry (ESIMS) of six separate preparations of Natex thaumatin gave an average mass of $M_r = 22190 ± 2$ Da for the monomer fraction. This result is consistent with the mass of thaumatin I ($M_r = 22188$ Da) as deduced from the published sequences, assuming that all disulfide bonds are intact. To further characterize the monomer fraction, we examined our samples by quasielastic light scattering (QLS), size-exclusion high performance liquid chromatography, and cation-exchange high performance liquid chromatography (SE-HPLC and CE-HPLC, respectively).

Figure 2 shows QLS results for purified Natex and unpurified Sigma thaumatin. The proteins were measured at the same concentration (2.6 mg/ml) and solution conditions (20 °C; 275 mM acetate buffer at pH 4.5 with 0.02% sodium azide). The average hydrodynamic radius for purified Natex thaumatin is 2.19 ± 0.02 nm (Figure 2A), which is consistent with the molecular mass of the monomeric protein. The purified protein is stable in solution at 4 °C for at least 3 months.

By contrast, the average hydrodynamic radius of Sigma thaumatin (Figure 2B) is larger (2.36 ± 0.05 nm) and the size distribution is broader than for Natex thaumatin. These QLS results suggest that Sigma thaumatin is less homogeneous than Natex thaumatin; this suggestion is confirmed by SE-HPLC (Figure 3A). Although Natex thaumatin (solid line) is essentially monomeric, Sigma thaumatin (dashed line) contains small amounts of oligomeric and low-molecular-weight contaminants. Another source of heterogeneity in Sigma thaumatin is revealed by CE-HPLC. As can be seen in Figure 3B, Natex thaumatin (solid line) is predominantly (> 80%) a single-charge species (major peak at 27 min). Sigma thaumatin (dashed line) contains two large peaks: one at 27 min, as in Natex thaumatin, and another at 31 min, where there is only a very small peak (less than 1%) in Natex thaumatin. ESIMS shows that the peak at 27 min has a molecular mass of 22190 ± 2 Da (Natex) and 22189 ± 2 Da (Sigma), consistent with that of thaumatin I (22 188 Da); the peak in Sigma thaumatin at 31 min corresponds to a mass of 22 272 ± 2 Da, consistent with that of thaumatin II (22 272 Da).

Finally, there is a heterogeneity that can be seen by eye. Sigma thaumatin is contaminated by a pigment that makes the solutions golden-yellow to reddish-brown; as the protein
Stereochemistry and Thaumatin Crystallization

Figure 3. HPLC results. The absorbance at 280 nm as a function of the retention time is shown. (A) Size-exclusion HPLC chromatograms for purified Natex thaumatin (6.5 mg/mL; solid line) and unpurified Sigma thaumatin (5.2 mg/mL; dashed line). The arrows indicate the additional components found in Sigma thaumatin. (B) Cation-exchange HPLC chromatograms for purified Natex thaumatin (12.3 mg/mL; solid line) and unpurified Sigma thaumatin (11.7 mg/mL; solid line). The large peak at 27 min is thaumatin I; the arrow indicates the thaumatin II peak.

concentration is increased, the color darkens. Liquid–liquid phase separation experiments show that this pigment associates with the protein.° Crude Natex thaumatin is also colored: it forms a dark-pink solution. The solutions of the purified protein, however, are colorless.

In summary, purified Natex thaumatin is essentially a homogeneous preparation of thaumatin I, whereas unpurified Sigma thaumatin is a mixture of thaumatin I and II (approximately 3:1 molar ratio) with oligomeric and low-molecular-weight contaminants.

3.3. Precipitation. As we explored various conditions for crystallizing thaumatin, we observed precipitation of the protein (Figure 4). Indeed, we were able to map out a coarse precipitation curve as a function of protein concentration, temperature, pH and ionic strength. This curve marks the boundary of where solid phases form; below the precipitation curve, the solution remains supersaturated without the formation of a solid phase even after long periods of time.° Although the precipitation curve is not a thermodynamic boundary like the solubility curve, it has been shown that knowledge of the location of the precipitation curve can help to obtain crystals.° The underlying idea is that precipitation indicates that the solution is supersaturated, a necessary condition for crystallization. Indeed, under conditions where precipitation is seen, the crystal is often the most stable phase, but nucleates too slowly to be observed. However, by fine-tuning the conditions, it should be possible to favor crystallization kinetically as well. In our case, we did so by adding the tartrate precipitant.

Figure 4. Precipitation of thaumatin. Amorphous structures formed in the absence of tartrate (C_{pro} (= 28 mg/mL) at 20 °C as the protein is concentrated. The solution contained 10 mM sodium phosphate (pH 7.3) with 0.002% sodium azide. The scale bars represents 50 μm.

We find that the concentration at which Natex and Sigma thaumatin precipitate decreases as either the pH or the ionic strength (I) increase. For example, Natex thaumatin can be concentrated up to 350 mg/mL in 275 mM sodium acetate with 0.02% sodium azide (pH 4.5, I = 120 mM) without forming any solid phase (for at least 48 h) over the temperature range –12 to 22 °C. The addition of 2 M sodium chloride at room temperature (20 ± 2 °C), however, leads to precipitation at 1 mg/mL. Similarly, if the pH is raised to 7.3, precipitation occurs at around 40 mg/mL despite the low ionic strength (10 mM sodium phosphate with 0.002% sodium azide; I = 24 mM).

These results are consistent with those of other investigators and can be explained in terms of the high isoelectric point (PI) of the protein.° As the pH or ionic strength is increased, the electrostatic repulsion between the proteins diminishes, leading to an increased attraction between the proteins and hence precipitation.° Similar behavior has been studied extensively for lysozyme, which also has a high PI.°

Though our results for Natex and Sigma thaumatin are qualitatively similar, there are important quantitative differences. For example, in 100 mM sodium phosphate with 0.02% sodium azide (pH = 7.1; I = 240 mM), the Natex protein precipitates at about 30 mg/mL at room temperature. Under the same conditions, we can often concentrate the Sigma protein to around 200 mg/mL before any precipitation occurs. This is consistent with the common observation that heterogeneities in proteins affect the phase behavior, often retarding the formation of a solid phase by reducing nucleation rates or hindering growth.°

3.4. Crystallization. We chose to perform our crystallization experiments on Natex thaumatin in 10 mM sodium phosphate with 0.002% sodium azide (pH 7.3 and I = 24 mM). This combination of low ionic strength and moderate pH allowed us to reach high concentrations of protein (100 mg/mL) while minimizing the formation of precipitates; thus we were able to produce large quantities of crystals for the solubility experiments. Additional factors also motivated our choice of conditions. First, we used a phosphate buffer because the pH varies very little with temperature.° Also, this buffer does not absorb significantly at 280 nm—unlike N-(2-acetamido)iminodiacetic acid, which is commonly used in thaumatin studies—and so does not interfere with the measurement of protein concentration. Second, because we add tartrate ions to the protein solution, the pH had to be sufficiently far from the second pK_a of the tartrate ion (at 25 °C, pK_{a2} = 4.34 for L- and D-tartaric acid
and $pK_a = 4.37$ for DL-tartaric acid\textsuperscript{56} to prevent crystallization of the poorly soluble, alkali metal hydrogen tartrates.\textsuperscript{57,58}

Figure 5 shows the crystalline phases that formed in our experiments with Natex thaumatin. We obtained bipyramidal protein crystals in solutions with 0.5 M sodium L-tartrate (Figure 5A). These are the crystals which have made thaumatin a popular model system for they form rapidly: crystals around $50 \mu m$ in size form within a few hours and sometimes even faster. When we used 0.5 M sodium D-tartrate as a precipitant, we obtained crystal with a different habit (Figure 5B). In this case, the crystals are initially prismatic and then grow wider to become stubby crystals. The crystals in D-tartrate can form as rapidly as those in L-tartrate, but they usually take longer to appear and to reach a comparable size. Both types of crystals are birefringent; the bipyramids are more strongly birefringent than the prismatic crystals. X-ray diffraction shows that the protein packs differently in the two habits: the unit cell in the bipyramidal crystal is tetragonal, whereas the unit cell in the prismatic and stubby crystals is orthorhombic.\textsuperscript{43} To the best of our knowledge, this is the first case in which a change in the chirality of a precipitant produces substantial differences in the morphology and packing of a protein crystal.

Several investigators have used the racemic mixture of sodium DL-tartrate to produce thaumatin crystals.\textsuperscript{29,30,44} Therefore, we crystallized the protein in 0.5 M of sodium DL-tartrate (to prepare the racemic precipitant, we mixed 0.5 M sodium L-tartrate with an equal volume of 0.5 M sodium D-tartrate). In the experiments with the racemate, we see bipyramids, prisms, and amorphous precipitate (Figure 6); the variability in the structures that form may reflect the heterogeneous nature of the racemic solution. Although the proportions of these structures vary from run to run, there is a correlation between the temperature and the habits formed: the lower the temperature, the more bipyramids we observed (This observation suggests an application: to use thaumatin crystals for the chiral separation of the racemic tartrate.).

The crystallization results for Natex thaumatin are robust: the phenomena described above were reproduced many times (five runs for the racemate, nine runs with L- and D-tartrate, and 21 runs with no tartrate) with only minor variations among runs. In contrast, we did not find such a high degree of reproducibility when crystallizing Sigma thaumatin. For example, sometimes when Sigma thaumatin was concentrated, the protein precipitated, but other times it did not. When we could reach high concentrations (at least 100 mg/mL) with Sigma thaumatin without precipitation occurring, crystals occasionally formed even though no tartrate had been added. Sometimes these were prisms, other times a mixture of prisms and bipyramids. This unpredictable behavior suggested that results for the solubility of Sigma thaumatin may not be reliable. Therefore, we focused our efforts on measuring the solubility curves on Natex thaumatin and made fewer measurements on Sigma thaumatin.

### 3.5. Solubility. 3.5.1. Natex Thaumatin

Figure 7A shows the solubility of Natex thaumatin; at each temperature we plot the protein concentration C, which is in equilibrium with the crystal phase. The solubilities in 0.5 M L-tartrate and 0.5 M D-tartrate are very different. With L-tartrate (where bipyramidal crystals form), the solubility is normal (direct), i.e., it increases with temperature. In D-tartrate (where prismatic and stubby crystals form), the solubility is retrograde, i.e., it decreases as the temperature increases. To extract the enthalpies $\Delta H$ and entropies $\Delta S$ of crystallization, we fit our solubility data using the van’t Hoff equation.\textsuperscript{43,59} The results are listed in Table 1. As expected, the energetic contribution to the free energy dominates for the normal solubility (L-tartrate), whereas the entropic contribution dominates for the retrograde solubility (D-tartrate).\textsuperscript{60}

The solubility results are unusual for three reasons. First, most globular proteins display normal solubility;\textsuperscript{63} retrograde solubility is less common.\textsuperscript{60–65} Second, if a protein crystallizes in more than one habit, the solubilities are usually of the same type.\textsuperscript{52,64,65} It is rare for a given protein to display both types of solubility because the underlying crystallization mechanisms are different: normal solubility is dominated by energetic effects, whereas retrograde solubility is driven by entropic interactions.\textsuperscript{60} Finally, our results are unusual because they involve the chirality of the precipitant. It is known that a single point mutation\textsuperscript{66} or a modification in the chemical composition of the precipitant\textsuperscript{16} can lead to a change in solubility from normal to retrograde, but our results provide the first example of an inversion in protein solubility due to a change in the chirality of the precipitant. We do not yet understand how this inversion occurs, and we are currently conducting X-ray diffraction experiments to elucidate this phenomenon.
3.5.2. Sigma Thaumatin. Because of the difficulty of obtaining reproducible results, our data on Sigma thaumatin is less complete than that on Natex thaumatin. The limited solubility measurements for bipyramidal crystals and prismatic crystals are given in Figure 7B. We also show other investigators’ results for bipyramidal crystals.\textsuperscript{28,30,42} To properly compare the various Sigma thaumatin results, we have recalculated the solubilities reported by Juárez-Martínez et al.\textsuperscript{42} and Kadri et al.\textsuperscript{30} with the extinction coefficient that we and Lorber et al.\textsuperscript{28} used. Table 1 gives the enthalpies and entropies of crystallization, and Table 2 summarizes the solution conditions used for each set of results.

There is qualitative agreement between our Natex and Sigma thaumatin results: bipyramids have normal solubility, and prisms have retrograde solubility. There is poor quantitative agreement, however, as can be seen from Figure 7B and Table 1. Furthermore, there are substantial differences among the various Sigma thaumatin results.

The differences between the Natex and Sigma thaumatin results are probably due, at least in part, to the differences between the proteins (see section 3.2). Natex thaumatin is essentially a homogeneous preparation of thaumatin I, whereas Sigma thaumatin is a mixture of thaumatin I and II with oligomeric and low-molecular-weight contaminants. The perturbing effect of contaminants on protein crystallization is well-known and has been studied in detail for lysozyme.\textsuperscript{67–72} Also, although thaumatin I and II are close homologues which cocrystallize (results not shown; see also Lorber et al.\textsuperscript{29}), this does not imply that their solubilities are the same. For example, \(\beta\)-lactoglobulin A and B differ only by two amino acids and cocrystallize in all proportions. Nevertheless, their solubilities differ significantly, and the solubilities of the mixtures fall between those of the two pure proteins.\textsuperscript{73}

It is less clear why there is a large spread for the results on Sigma thaumatin. The solution conditions used by various investigators are not identical (see Table 2), which may contribute to variation in solubility. The conditions are close enough, however, that the variation in solubility should be less than that observed. For example, Kadri et al.\textsuperscript{30} report that when the pH is changed from 7.0 to 6.5, the solubility increases by 7\%, while the results in Figure 7B change by more than a factor of 2 (the bowties are the results at pH 7.0; the cross is at pH 6.5).

![Figure 7. Solubility of thaumatin.](image)

**Table 1. Thermodynamic Parameters for the Crystallization of Natex and Sigma Thaumatin**

| Source | Precipitant | \(\Delta H\) (kJ/mol) | \(\Delta S\) (kJ/mol) | \(T_0\) | Ref.
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Natex 0.50 M Na l-tartrate</td>
<td>-43 ± 4</td>
<td>-24 ± 1</td>
<td>this work</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Natex 0.50 M Na D-tartrate</td>
<td>19 ± 1</td>
<td>33 ± 2</td>
<td>this work</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sigma 0.75 M Na K l-tartrate</td>
<td>-28 ± 4</td>
<td>-12 ± 4</td>
<td>this work</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sigma No tartrate</td>
<td>9 ± 2</td>
<td>20 ± 2</td>
<td>this work</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sigma 0.53 M Na K l-tartrate</td>
<td>18</td>
<td>32</td>
<td>Juárez-Martínez et al.\textsuperscript{42}</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sigma 0.71 M Na K l-tartrate</td>
<td>8</td>
<td>23</td>
<td>Juárez-Martínez et al.\textsuperscript{42}</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sigma 0.73 M Na DL-tartrate</td>
<td>-26 ± 5</td>
<td>-14 ± 5</td>
<td>Kadri et al.\textsuperscript{30}</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\textsuperscript{a} The precipitants are abbreviated as follows: sodium l-tartrate (Na l-tartrate); sodium D-tartrate (Na D-tartrate); sodium potassium l-tartrate (Na K l-tartrate); and sodium DL-tartrate (Na DL-tartrate). \(T_0\) is 293 K.

**Table 2. Solution Conditions for the Solubility Measurements of Natex and Sigma Thaumatin**

| Source | Precipitant | Buffer | pH | Symbol | Ref.
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Natex 0.50 M Na l-tartrate</td>
<td>10 mM Na PO\textsubscript{4}</td>
<td>7.3</td>
<td>open squares; this work</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Natex 0.50 M Na D-tartrate</td>
<td>10 mM Na PO\textsubscript{4}</td>
<td>7.3</td>
<td>open triangles; this work</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sigma 0.75 M Na K l-tartrate</td>
<td>100 mM Na PO\textsubscript{4}</td>
<td>7.1</td>
<td>solid squares; this work</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sigma No tartrate</td>
<td>100 mM Na PO\textsubscript{4}</td>
<td>7.1</td>
<td>solid triangles; this work</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sigma 0.53 M Na K l-tartrate</td>
<td>Na K PO\textsubscript{4}</td>
<td>7.0</td>
<td>diamonds; Juárez-Martínez et al.\textsuperscript{42}</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sigma 0.71 M Na K l-tartrate</td>
<td>Na K PO\textsubscript{4}</td>
<td>7.0</td>
<td>bowties; Juárez-Martínez et al.\textsuperscript{42}</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sigma 0.73 M Na DL-tartrate</td>
<td>100 mM ADA*</td>
<td>6.5</td>
<td>circles; Kadri et al.\textsuperscript{30}</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sigma 0.71 M Na DL-tartrate</td>
<td>100 mM ADA</td>
<td>6.5</td>
<td>cross; Lorber et al.\textsuperscript{28}</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\textsuperscript{a} The buffers are abbreviated as follows: sodium phosphate (Na PO\textsubscript{4}); disodium hydrogen phosphate/ potassium dihydrogen phosphate (Na K PO\textsubscript{4}); N-(2-acetamido)iminodiacetic acid (ADA); and ADA with 0.3\% (m/v) agarose gel (ADA*). \(\Delta\) The symbols refer to those used in Figure 7A (Natex) and 7B (Sigma).
The differences in tartrate concentration seem to be more important: the solubility of Sigma thaumatin decreases as the tartrate concentration increases.\textsuperscript{28,42} Nevertheless, this observation is insufficient to explain the results: the data in Figure 7B at approximately 20 °C and 0.71–0.75 M tartrate spans a factor of almost ten in concentration.

We believe that other factors may explain the divergence of the results. Apart from a single set of measurements (our data on prismatic crystals without tartrate; Figure 7B), all of the Sigma results are nominally for bipyramidal crystals. As our experiments with Natex thaumatin show, however, the racemic Sigma results are nominally for bipyramidal crystals. Thus, the measurements of Kadri et al.\textsuperscript{30} and Lorber et al.\textsuperscript{28} where D-tartrate is used, might be explained similarly if prisms were able to form before the L-tartrate was added.

Another complication that can affect the solubility measurements is the formation of tartrate crystals, as the alkali metal hydrogen tartrates have low solubilities in water.\textsuperscript{48} In particular, potassium hydrogen tartrate (KHT) has low solubility: 26.2 mM at 18.6 °C and 14.1 mM at 4.8 °C.\textsuperscript{55} Therefore, it is potentially problematic to use sodium potassium tartrate (NaKT) as a precipitant, as do Juárez-Martínez et al.\textsuperscript{22} (these investigators report solubilities at 18 and 4 °C). Although the concentration of the hydrogen tartrate ion in solution is small at pH 7.0, it may still be enough to produce tartrate crystals. We estimate that the solubility product of KHT will be exceeded if the molarity of NaKT in solution is above 0.56 M at 18.6 °C or 0.30 M at 4.8 °C. As the molarities of NaKT typically used are 0.50–0.75 M (Table 2), crystals of KHT are thermodynamically stable under the conditions studied. Indeed, we have observed tartrate crystals in our experiments with Sigma thaumatin and NaKT. Usually, the tartrate crystals can be distinguished from protein crystals by their size, shape, and large birefringence. A small tartrate crystal, however, can sometimes be confused with a large, prismatic, thaumatin crystal. To avoid these complications, we eventually abandoned the use of NaKT, switching instead to tartaric acid. Also, we chose buffers that did not contain any potassium.

Finally, additional errors can be introduced by the method used to measure the solubility. It is possible to measure the solubility by starting with an undersaturated solution (where the crystals dissolve until equilibrium is reached) or a super-saturated one (where equilibrium is reached through the formation of new crystals). Ideally, both methods should produce the same result. We checked the Natex thaumatin data, which was obtained principally from undersaturated solutions, and verified that, given enough time, supersaturated solutions produced the same solubility results. It is more difficult, however, to establish equilibrium when starting from supersaturated solutions because as the crystals grow, their surfaces become poisoned with impurities or improperly oriented proteins. As a result, further growth stops, leading to an apparent solubility that is higher than the equilibrium value.\textsuperscript{40} This may be why our solubility results for Sigma thaumatin with L-tartrate (solid squares in Figure 6B) are lower than those of the other investigators who measured the solubility upon growth.

4. Conclusions

We determined the solubility of thaumatin from two commercial sources: Natex UK Limited and Sigma-Aldrich. We purified the Natex protein and demonstrated that it is pure thaumatin I (99.9% by SE-HPLC), in contrast to the Sigma protein, which is a mixture of thaumatin I and II along with oligomers and low molecular weight contaminants. For Natex thaumatin, we found that the chirality of the tartrate precipitant has a dramatic effect on the crystal habit and solubility: with L-tartrate, bipyramidal crystals with normal solubility form; with D-tartrate, prismatic and stubby crystals with retrograde solubility form. The results with Sigma protein show qualitative agreement, but there are quantitative discrepancies due to several factors, including the heterogeneities in the protein, the stereochemical identity of the precipitant, and the method used to measure the solubility. We hope that our results will help others crystallize proteins by promoting the use of both pure protein and stereochemically pure precipitants.

Acknowledgment. We are grateful to Bruce Hrnjez, Fredy Zyppman, Ann Rae Jonas, Jacob Berger, and Marion Giffard for helpful discussions and for comments on the manuscript. We thank Charles Boy of Natex UK Limited for generously providing the thaumatin used in this work. We are especially grateful to Morton Lowengrub and Yeshiva University for their generous support of this work. N.A. thanks the Milton and Miriam Handler Foundation for additional financial support.

References


