Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Optimum solubility (OS) screening: an efficient method to optimize buffer conditions for homogeneity and crystallization of proteins

Jarmila Jancarik,[a] Ramona Pufan,[a] Connie Hong,[a] Sung-Hou Kim[a,b] and Rosalind Kim[a]*

[a]Berkeley Structural Genomics Center, Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA, and [b]Department of Chemistry, University of California, Berkeley, California 94720, USA

Correspondence e-mail: r_kim@lbl.gov

Received 15 February 2004
Accepted 4 May 2004

One of the most critical steps in the preparation of protein samples for structural studies by X-ray crystallography is to obtain biochemically pure and conformationally homogenous protein samples. Very often, the purified sample does not meet these qualifications and therefore does not crystallize. A screening method, Optimum Solubility Screen, has been developed that consists of two steps. The first step selects a better buffer than that used during purification. 24 different buffers ranging from pH 3 to pH 10 are screened using a vapor-diffusion method and very small amounts of protein. The solubility of the protein is first determined by visual examination using a light microscope and those drops that remain clear after 24 h are further evaluated using dynamic light scattering. If the results from the first step are still not satisfactory, a second step explores a variety of chemical additives in order to improve the monodispersity of the protein sample. In 64% of the cases, crystallization was successful from proteins that had initially shown high levels of aggregation. This screen can be configured to perform in an automated high-throughput mode and can be expanded for additional buffers and additives.

1. Abbreviations

ADA, *N*-(2-acetamido)-iminodiacetic acid; BME, β-mercaptoethanol; bis-tris, 2,2-bis-(hydroxymethyl-2,2′,2′′-nitrilotriethanol; CAPS, 3-(cyclohexylamino)-1-propanesulfonic acid; CHAPS, [3-(3-cholamidopropyl)dimethyl-ammonio]-1-propanesulfonate; CHES, 2-(cyclohexylamino)ethanesulfonic acid; DLS, dynamic light scattering; DTT, 1,4-dithio-DL-threitol; EPPS, 4-(2-hydroxyethyl)piperazine-1-propanesulfonic acid; HEPES, *N*-2-hydroxy-ethylpiperazine-*N*-2-ethanesulfonic acid; MES, 2-(*N*-morpholino)ethanesulfonic acid; MOPS, 3-(*N*-morpholino)propanesulfonic acid; PIPPS, 4-(2-hydroxyethyl)piperazine-1-propanesulfonic acid; Tris, [tris(hydroxymethyl)amino-methane]; TCEP, tris(2-carboxyethyl) phosphine.

2. Introduction

Advances in X-ray crystallography have provided three-dimensional structures of thousands of proteins. In spite of these advances, protein aggregation continues to be a common problem that can lead to unsuccessful crystallization of proteins. This problem is becoming more prominent as attempts to crystallize many different proteins and protein complexes on a structural genomics scale continue to expand.

At the Berkeley Structural Genomics Center (BSGC), a purified protein sample is obtained after one or more chromatography steps [immobilized metal-affinity chromatography (IMAC), ion-exchange chromatography and size-exclusion chromatography] and the sample is analyzed by SDS–PAGE (Laemmli, 1971) to determine the purity of the protein. In the process of developing techniques to automate protein purification, one tries as much as possible to use a set of generic buffers. Very little is known about the properties of the proteins we are working with *a priori*, except for their theoretical pI, molecular weight and amino-acid composition. The general practice is to use one or two favorite buffers, where pH and salt concentration are some of the variables. However, a protein has complex properties and its condition and behavior depend very much on the environment it is in. In the past, when faced with difficulties of aggregation and precipitation we would try to change purification parameters, add or remove fusion tags and/or test some additives (DTT, glycerol *etc.*); after purification, the protein would be concentrated and the presence of aggregates would be assessed using the dynamic light-scattering method. This method, as described by Zulauf & D'Arcy (1992), has shown that the presence of aggregates in the protein solution may inhibit crystal nucleation or growth (D'Arcy, 1994; Ferré-D'Amare & Burley, 1997; Habel *et al.*, 2001).

Inspired by a screen for NMR studies (LePre & Moore, 1999), we have developed a screen in which we test a panel of buffers and many additives in order to obtain the most homogeneous and monodisperse protein conditions for proteins that usually aggregate and cannot be concentrated prior to setting up crystallization screens. A panel of 24 buffers is tested using the hanging-drop method and vapor-diffusion equilibrium. After monitoring precipitation, the conditions leading to clear drops are selected for dynamic light-scattering (DLS) characterization. For this part of the screen, only 24 µl of protein (of concentration at least 3 mg ml^{-1}) is required. If the DLS results are not optimal, a series of additives are tested in the presence of the best buffer selected from the initial screen and DLS is again used to determine the best condition. We have tested 14 poorly behaving proteins so far: 11 of the proteins had highly improved DLS results and were able to concentrate well after exchanging the buffer. Nine of these have crystallized.

3. Experimental

3.1. Protein samples

The proteins that we have worked on are expressed from genes from *Mycoplasma pneumoniae*, *M. genitalium* or their homologs from other organisms. The protein-identification numbers listed in Table 3 are targets from BSGC (http://strgen.org). These proteins have a His_6 tag or a His_6-maltose-binding protein (MBP)-tobacco etch virus (TEV) protease cleavage site at the N-terminus. Purification of the proteins with a His_6 tag involved an immobilized metal-affinity column (IMAC) followed by either ion-exchange chromatography and/or size-exclusion chromatography. For those proteins with a His_6-MBP fusion, a TEV cleavage step was included in the purification protocol after the IMAC step.

3.2. Optimum solubility (OS) screen

A set of 24 buffers each at a concentration of 100 m*M* with a pH range from 3 to 10 was prepared using Fluka chemicals (Table 1). The screening was performed by pipetting 0.5 ml of each of the buffers into each reservoir of a 24-well Linbro plate. A 1 µl aliquot of protein solution (concentrated to as high a concentration as possible) was pipetted onto a siliconized glass cover slip and to it was added 1 µl of 100 m*M* reservoir buffer. A layer of silicone grease was applied to the top of each of the wells and the cover slips were then inverted and sealed onto each of the wells. The plates were incubated at room temperature for 24 h. During this period, vapor diffusion takes place and, depending on the stability of the protein in a given buffer, clear drops or drops with different degrees of precipitation can be observed under a light microscope. The protein concentration ranged from 3.0 to 30 mg ml^{-1}.

Table 1
Optimum solubility screen buffers.

No.	Buffer (100 m*M*)	pH
1	Glycine	3
2	Citric acid	3.2
3	PIPPS	3.7
4	Citric acid	4
5	Sodium acetate	4.5
6	Sodium/potassium phosphate	5
7	Sodium citrate	5.5
8	Sodium/potassium phosphate	6
9	Bis-tris	6
10	MES	6.2
11	ADA	6.5
12	Bis-tris propane	6.5
13	Cacodylate	6.5
14	Ammonium acetate	7
15	MOPS	7
16	Sodium/potassium phosphate	7
17	HEPES	7.5
18	Tris	7.5
19	EPPS	8
20	Imidazole	8
21	Tris	8.5
22	CHES	9
23	CHES	9.5
24	CAPS	10

3.3. Dynamic light-scattering (DLS) analysis

The drops that remained clear gave us an indication that the protein sample was more soluble in those buffers. The clear drops were then diluted into the same reservoir solution at a ratio of 1:14 (protein:buffer) and DLS was performed using a DynaPro-99 (Proterion Corp., Piscataway, NJ, USA) to assess the homogeneity/monodispersity of the sample. If the protein sample appeared to be monodisperse in a particular buffer, then the protein was exchanged into that buffer at a final concentration of 20 m*M*.

3.4. Additive screen

If none of the clear drops provided an ideal monodisperse sample (radius < 5 nm/polydispersity < 25%), the buffer that gave the best DLS reading was selected and an additive screen was tested. The protein sample was exchanged into the best buffer from the OS screen using an Ultrafree unit (Millipore Corp., Bedford, MA, USA) to a final concentration of at least 2–3 mg ml^{-1}. In 0.5 ml microfuge tubes, individual additives were added to 15 µl of protein sample (Table 2). Selection of additives was performed by choosing the most common compounds that have been found in the past to be beneficial for protein solubility. After 2 h incubation at room temperature, DLS readings were performed. The condition that gave the best DLS reading was selected and the protein was exchanged into the selected buffer and additive before setting up the crystal screens.

Table 2
Additive screen.

These are the final concentrations of additives used in a final volume of 15 µl.

20, 50, 100 m*M* sodium chloride
5 and 10% glycerol
2 m*M* CHAPS (CMC = 6–10 m*M*)
0.1, 1% octylglucoside (CMC = 0.53%)
0.1, 1% dodecyl maltoside (CMC = 0.0087%)
10 m*M* BME
1, 5 m*M* DTT
30 m*M* TCEP

4. Results

4.1. Protein 1371B

The OS screen was performed on 14 samples of cytoplasmic proteins that had aggregated as measured by DLS, had precipitated upon concentration or could not be concentrated. One of the proteins, 1371B, was purified through a HiTrap Chelating HP column (Amersham Biosciences, Piscataway, NJ, USA) and a Superdex 75 size-exclusion chromatography (Amersham Biosciences, Piscataway, NJ, USA) column. The sample in 20 m*M* Tris pH 7, 0.3 *M* NaCl could be concentrated to 27 mg ml^{-1} only upon removal of salt, but the DLS showed large aggregation. The protein underwent OS screening and some of the drops remained clear. However, as shown in Figs. 1(*a*) and 1(*b*), although both buffers gave a clear drop, the sample was more monodisperse in Tris buffer pH 8.5. Even in this buffer, a higher molecular-weight peak could be observed, so the additive screen was tested and 1% octylglucoside was shown to make the 1371B sample more monodisperse, with a radius of 3.15 nm and 21% polydispersity. The buffer was exchanged to 20 m*M* Tris pH 8.5, 1% octylglucoside (Fig. 1*c*) and crystallization screens were performed. The protein crystallized (Fig. 1*d*) and the structure was determined (unpublished results).

4.2. Examples of other proteins that underwent OS screening

Table 3 lists 14 protein samples that have been screened through the OS screen. These proteins varied in molecular weight over a

Table 3
Optimization results.

Protein ID	Molecular weight (kDa)/pI†	DLS after purification (nm‡/PD§)/purification buffer	DLS in best buffer (nm‡/PD§)/ best buffer	Additive	Crystals
1008B	34.9/9.1	Protein precipitated during concentration/Tris pH 7.5	Precipitation in all drops		No
1135B	59.8/5.9	7.6/35%; protein precipitated during concentration/HEPES pH 8.0	5.8/9%/PIPPS pH 3.7		Yes¶
1139B	16.7/6.7	6.1/37%/Tris pH 8.0	4.2/27%/MES pH 6.5	5% glycerol	Yes
1142B	16.8/6.8	39/19.3%; protein precipitated/HEPES pH 8.0	4.9/20%/sodium acetate pH 5.0	50 m*M* NaCl, 1% OG, 0.1 *M* TCEP	Yes
1149B	10.8/9.2	7.36/66%/HEPES pH 7.0	3.5/15%/potassium phosphate pH 5.0		Yes
1154B	24.9/8.5	8.8/53%/HEPES pH 7.5	3.41/13.8%/potassium phosphate pH 7.5		No
1166B	30.1/4.9	Protein aggregated with time/Tris pH 7.5	No good buffer found		No
1227B	35.6/6.8	Protein precipitated during concentration/HEPES pH 8.0	Crystallized in OS screen/MES pH 6.2		Yes¶
1275B	44.1/6.4	Could not read the DLS/HEPES pH 8.0	3.78/28%/MES pH 6.0		Yes
1334B	59.9/4.3	11.6/45%/HEPES pH 7.0	5.1/14%/HEPES pH 7.5		No
1335B	53.5/4.6	19/46%/HEPES pH 7.0	No good buffer found		No
1349B	23.1/9.1	2.05/8.6%, part of large aggregate/HEPES pH 7.5	Crystallized in OS screen/MES pH 5.8		Yes¶, solved
1368B	24.6/6.3	4.22/58%/Tris pH 7.5	3.9/12%/sodium acetate pH 5.0	25 m*M* NaCl	Yes
1371B	27.1/5.3	Too aggregated/Tris pH 7.0	3.15/21%/Tris pH 8.5	0.1% OG	Yes, solved

† pI, isoelectric point. ‡ nm, hydrodynamic radius in nanometers. § PD, polydispersity. ¶ Protein crystallized while performing the OS screen.

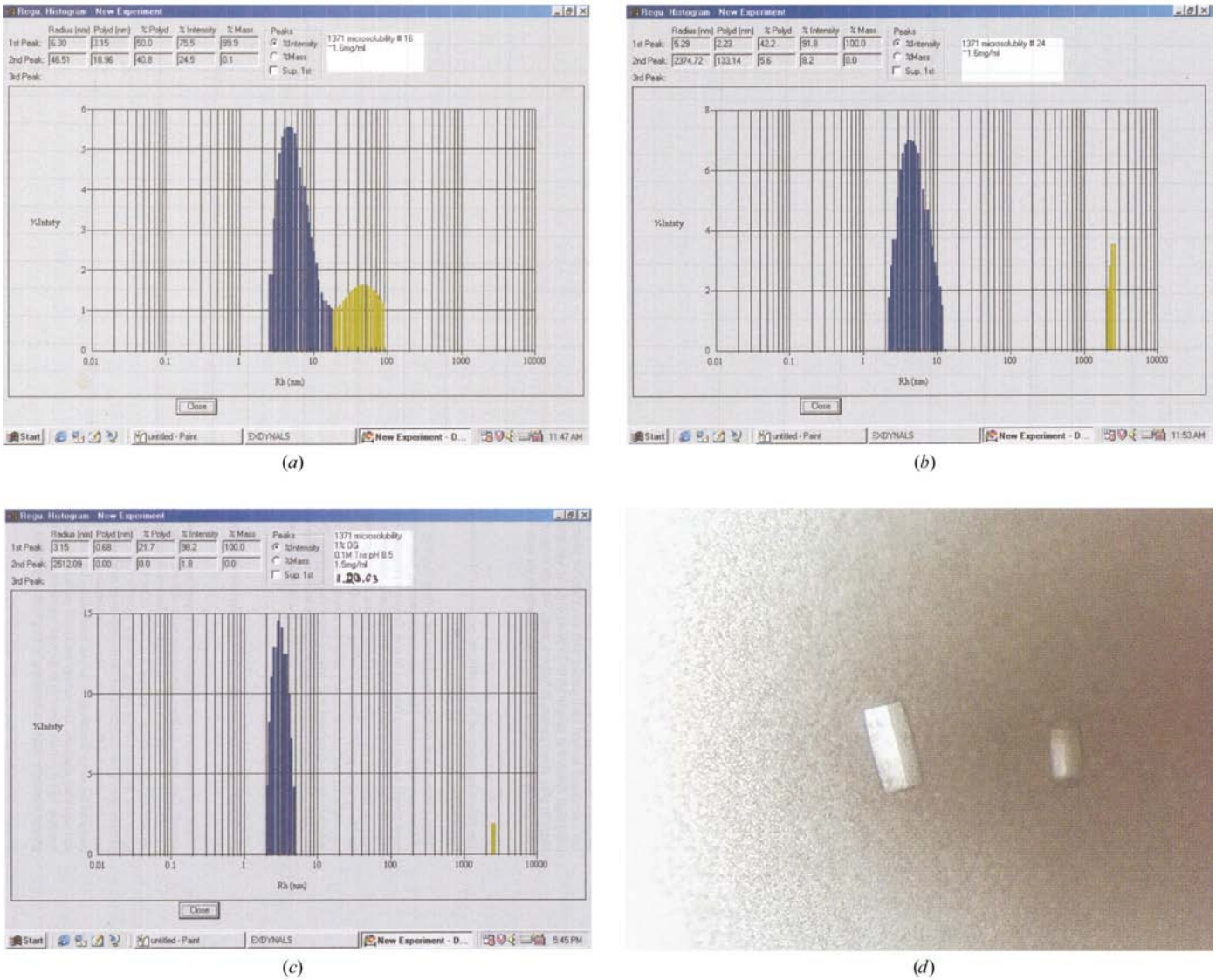

Figure 1
Dynamic light-scattering scans of protein 1371B in different OS screen buffers. (*a*) Protein in 100 m*M* CHES pH 9.0. (*b*) Protein in 100 m*M* Tris pH 8.5. (*c*) Protein in 100 m*M* Tris pH 8.5, 1% octylglucoside. (*d*) The buffer for 1371B was exchanged to 20 m*M* Tris pH 8.5, 1% octylglucoside, crystals were obtained from 0.1 *M* sodium acetate pH 4.6, 6.0 *M* ammonium nitrate and the structure was solved (unpublished results).

wide range, showed aggregation or precipitation upon concentration or had non-ideal DLS results. The OS screen indicated that out of the 14 protein samples, the DLS of 11 of them could be improved by exchanging the buffer and in some cases an additive further improved the DLS. Nine of these proteins could be crystallized. It is also interesting to note that in three cases (targets 1135B, 1227B and 1349B) the proteins crystallized while the OS screen was being performed. Crystals were improved by testing different molecular-weight polyethylene glycols (400, 1500, 3350, 4000, 8000) as precipitants with the particular buffer that was selected as being the optimum one for solubility, but the protein was left in the original buffer. This generated much improved crystals. The structures of two of the proteins that crystallized have been solved (1349B, 1371B).

5. Discussion

In order to grow crystals, the conditions under which proteins will precipitate out of solution must be identified. The technique of vapor diffusion is used to analyze this controlled precipitation and by using a sparse-matrix approach (Jancarik & Kim, 1991), further expanded by Hampton Research (Laguna Niguel, CA, USA), a large number of crystallization conditions can be tested. This assumes that the starting protein solution is not aggregated or precipitated. In the preparation for both NMR or X-ray crystallography samples, one must start with a protein solution that is homogeneous and monodisperse. Lepre & Moore (1998) developed a modified vapor-diffusion method to efficiently screen solvent conditions for NMR samples in order to optimize solubility. We have adapted this approach for X-ray crystallography in order to increase the solubility of samples that were badly aggregated or showed precipitation and coupled it with DLS in order to determine the aggregation state of a given protein sample. This method requires a very small quantity of protein for the initial OS screen and for testing additives. The ease of use of this screen and the capability of testing many conditions makes this a very efficient means of searching for the best solvent condition for a particular protein. This screen is empirically based and a physical chemical basis for why it has worked well requires the accumulation of many crystal structures obtained from crystals using this screen.

We are grateful to Hisao Yokota and Barbara Gold for cloning, Marlene Henriquez and Bruno Martinez for expression studies and cell-paste preparation, Candice Huang, Andy DeGiovanni, Steve Xu, Yun Lou, Natasha Oganesyan and Vaheh Oganesyan for protein purification. The work described here was supported by the National Institutes of Health grant GM 62412.

References

D'Arcy, A. (1994). *Acta Cryst.* D**50**, 469–471.
Ferré-D'Amare, A. R. & Burley, S. K. (1997). *Methods Enzymol.* **276**, 157–166.
Habel, J. E., Ohren, J. F. & Borgstahl, G. E. O. (2001). *Acta Cryst.* D**57**, 254–259.
Jancarik, J. & Kim, S.-H. (1991). *J. Appl. Cryst.* **24**, 409–411.
Laemmli, U. K. (1971). *Nature (London)*, **22**, 680–685.
Lepre, C. A. & Moore, J. M. (1999). *J. Biomol. NMR*, **12**, 493–499.
Zulauf, M. & D'Arcy, A. (1992). *J. Cryst. Growth*, **122**, 102–106.