Available online at www.sciencedirect.com
ScienceDirect
Analytical Biochemistry 357 (2006) 289–298
ANALYTICAL BIOCHEMISTRY
www.elsevier.com/locate/yabio

Thermofluor-based high-throughput stability optimization of proteins for structural studies

Ulrika B. Ericsson [a], B. Martin Hallberg [a,b], George T. DeTitta [c], Niek Dekker [d], Pär Nordlund [a,b,*]

[a] *Department of Biochemistry and Biophysics, Stockholm University, Stockholm SE-109 51, Sweden*
[b] *Department of Medical Biochemistry and Biophysics, Karolinska Institutet, Stockholm SE-171 77, Sweden*
[c] *Hauptman–Woodward Institute, Buffalo, NY 14203, USA*
[d] *AstraZeneca, R&D Mölndal, SE-431 83 Mölndal, Sweden*

Received 11 January 2006
Available online 10 August 2006

Abstract

Production of proteins well suited for structural studies is inherently difficult and time-consuming. Protein sample homogeneity, stability, and solubility are strongly correlated with the proteins' probability of yielding crystals, and optimization of these properties will improve success rates of crystallization. In the current study, we applied the thermofluor method as a high-throughput approach for identifying optimal protein formulation for crystallization. The method also allowed optimal stabilizing buffer compositions to be rapidly identified for each protein. Furthermore, the method allowed the identification of potential ligands, physiological or non-physiological, that can be used in subsequent crystallization trials. For this study, the thermally induced melting points were determined in different buffers as well as with additives for a total of 25 *Escherichia coli* proteins. Crystallization trials were set up together with stabilizing and destabilizing additives identified using thermofluor screening. A twofold increase in the number of crystallization leads was observed when the proteins were cocrystallized with stabilizing additives as compared with experiments without these additives. This suggests that thermofluor constitutes an efficient generic high-throughput method for identification of protein properties predictive of crystallizability.

Keywords: Thermofluor; High-throughput screening; Crystallization

An unprecedented number of purified proteins are being produced through the efforts of structural genomics projects, and thus an increasing number of proteins are available in sufficient amounts for structural and functional studies. X-ray crystallography is the most powerful method for determining three-dimensional structures of proteins, and it is evident that the production of high-quality protein samples, allowing useful crystals to be grown, is emerging as the main rate-limiting step for structure determination.

Protein homogeneity, stability, and solubility are factors that are expected to be highly correlated with the probability of a protein forming useful crystals. The treatment of the protein samples during scale-up production is important, where the buffer providing the local solvent environment keeps the protein stable and minimizes aggregation. Similarly, the solvent environment of the protein in a crystallization experiment critically controls the events potentially leading to crystal growth. The number of variables affecting crystallization, such as pH, ionic strength, additives, precipitants, and protein concentration, is large, and a vast number of conditions usually are attempted to identify a crystallization lead condition [1]. Both structural and functional studies require stable proteins, but many proteins are found to be unstable under a typical buffer condition. Although buffer conditions can be optimized using functional assays, the emerging view is that optimization on biophysical properties, such as homogeneity, solubility, and stability, is more likely to be predictive of

* Corresponding author. Fax: +46 8 52486850.
E-mail address: par.nordlund@mbb.ki.se (P. Nordlund).

0003-2697/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.ab.2006.07.027

successful crystallization. For structural genomics, the biological function of the target proteins is often not known prior to the structural analysis and therefore a generic biophysical characterization method is required.

Many proteins have been biophysically characterized in individual studies, but few systematic studies have been made on multiple proteins to establish how biophysical properties are correlated with crystallizability. Zulauf and D'Arcy reported that monodispersity of proteins measured using dynamic light scattering (DLS)[1] correlated well with crystallizability [2]. Although recent developments allow DLS-based buffers and ligands screening to be performed in multiwell formats, the correlation with ligand and buffer screen effects have not yet been investigated.

In the current study, we applied the fluorescence-based thermal stability assay developed by Pantoliano and coworkers (US patent 6,020,141) as a high-throughput screen for buffer optimization and ligand-induced stabilization of proteins. The method originally was developed for drug discovery to allow rapid identification of ligands of target proteins from compound libraries [3]. The conceptual basis for the method is that folded and unfolded proteins can be distinguished through exposure to a hydrophobic fluoroprobe. The probe is quenched in aqueous solution but will preferentially bind to the exposed hydrophobic interior of an unfolding protein leading to a sharp decrease in quenching so that a readily detected fluorescence emission can be studied as a function of temperature. Thermally induced unfolding is an irreversible unfolding process following a typical two-state model with a sharp transition between the folded and unfolded states, where T_m is defined as the midpoint of temperature of the protein-unfolding transition. Melting temperatures obtained with the thermofluor method have been shown to correlate well for several proteins with values determined by other biophysical methods such as circular dichroism (CD), turbidity measurements, and differential scanning calorimetry (N. Dekker, unpublished data).

Lo and coworkers used the thermal shift assay to study 1 protein, β-amyloid precursor protein-cleaving enzyme 1 (BACE1), combined with 13 different inhibitors [4]. The data obtained were used to calculate the ligand association constants, and the results were found to be consistent with ligand association constants measured by isothermal titration calorimetry. Good agreement between results obtained with the thermofluor method and the binding affinities acquired with several other methods, such as radioactive competition and fluorescence polarization assays, has also been shown [3].

The thermofluor assay can be performed in a commercially available real-time PCR instrument [4], where thermal melting curves of buffer/ligand conditions can be screened quickly in high-throughput mode (96 samples in 1 h). Only relatively small quantities of protein, some 1–15 μg per well, are needed. These features make the method very well suited for high-throughput biophysical characterization, and the method potentially constitutes an efficient technique to search for optimal buffers and ligands. The method is generic in the sense that no prior knowledge of the protein is required to screen for stabilizing conditions.

Here we present an application of the thermofluor method as a generic approach for identifying stabilizing buffers and ligands. The buffer screen allows fast identification of stabilizing buffer conditions. Different additives, found to be stabilizing or destabilizing in the thermofluor-based additive screen, were used in crystallization experiments together with 10 of the 25 different *Escherichia coli* proteins used in this study. Using stabilizing additives, identified with the thermofluor additive screen, in the crystallization screens significantly improved the proteins' probability of yielding crystals. Hence, the thermofluor additive screen constitutes a fast approach to identify stabilizing additives to obtain increased success rates in crystallization experiments.

Materials and methods

Sample preparation

All of the proteins used in the thermal shift assay are expressed from genes from *E. coli* (Table 1). The molecular weights of these proteins range from 19 to 79 kDa. Recombinant proteins were expressed with either a six-residue His fusion tag both N-terminally and C-terminally, a C-terminal His fusion tag (STHHHHHH-C) and an N-terminal Flag fusion tag (N-MDYKDDDDKGSTSLYKKAGSTELYIQG), or only an N-terminal His fusion tag (N-MHHHHHHGSTSLYKKAGFEDRT) using a PT73.3HisGW, PT73.3FlagGW, or modified pET15b expression vector [5,6] in *E. coli* BL21 (DE3) as described previously [7]. Purification of the proteins was performed using an immobilized metal affinity column (IMAC) followed by gel filtration chromatography using a buffer consisting of 20 mM Hepes (pH 7.4), 150 mM NaCl, 10 mM β-mercaptoethanol, 10% glycerol, and 2 mM ethylenediaminetetraacetic acid (EDTA). Mass spectrometry was used to confirm that the proteins were of the expected masses.

Thermal shift assay

Solutions of 7.5 μl of 300 × Sypro Orange (Molecular Probes), 12.5 μl of 2 × test compound (buffer or additive, see Tables 2 and 3), and 5 μl of 2.5 mg/ml protein were added to the wells of a 96-well thin-wall PCR plate (Bio-Rad). Water was added instead of test compound in the control samples. The plates were sealed with Optical-Quality Sealing Tape (Bio-Rad) and heated in an iCycler iQ Real Time PCR Detection System (Bio-Rad) from 20 to

[1] *Abbreviations used:* DLS, dynamic light scattering; CD, circular dichroism; BACE1, β-amyloid precursor protein-cleaving enzyme 1; IMAC, immobilized metal affinity column; EDTA, ethylenediaminetetraacetic acid; CCD, charge-coupled device; PDB, Protein Data Bank; AdoMet, *S*-adenosyl-L-methionine.

Table 1
Protein targets used in the study

Target	Swiss–Prot number	Predicted function	Fusion tag[a]
AC02	P42624	Hypothetical protein yhaK	H/H
AC04	P39173	Protein yeaD	H/H
AC05	P75864	Hypothetical protein ycbY	H/F
AC06	P76055	Hypothetical protein ydaO	H/H
AC07	Q57261	tRNA pseudouridine synthase D	H/H
AC08	P46850	Protein rtcB	H/F
AC11	P32695	tRNA-dihydrouridine synthase A	H/H
AC12	P26615	3-Octaprenyl-4-hydroxybenzoate carboxy-lyase	H/H
AD01	P06989	Histidine biosynthesis bifunctional protein hisIE	H/–
AD17	P31057	3-Methyl-2-oxobutanoate hydroxymethyltransferase	H/–
AD21	P00373	Pyrroline-5-carboxylate reductase	H/–
AD28	P37757	Hypothetical protein yddE	H/–
AE15	P09126	Uroporphyrinogen III synthase	H/F
AE30	P36553	Coproporphyrinogen III oxidase	H/H
AE37	P16384	IPP transferase	H/H
AF75	P76273	Hypothetical protein yebU	H/H
EC0942	P33601	NADH dehydrogenase I chain E	H/–
EC1027	P27863	Protoporphyrinogen oxidase	H/–
EC1192	P64588	Hypothetical protein yqjI	H/–
EC1221	P27110	Peptide methionine sulfoxide reductase	H/–
EC4022	P76174	Hypothetical protein ynfI	H/–

[a] The proteins are either His-tagged both N- and C-terminally (H/H), Flag-tagged N-terminally and His-tagged C-terminally (F/H), or His-tagged only N-terminally (H/–).

Table 2
List of the buffers and their pH values used in the screen

Number	Buffer (100 mM)	pH
1	Sodium acetate	4.5
2	Sodium citrate	4.7
3	Sodium acetate	5.0
4	Potassium phosphate	5.0
5	Sodium phosphate	5.5
6	Sodium citrate	5.5
7	Mes	5.8
8	Potassium phosphate	6.0
9	Mes	6.2
10	Sodium phosphate	6.5
11	Sodium cacodylate	6.5
12	Mes	6.5
13	Potassium phosphate	7.0
14	Hepes	7.0
15	Ammonium acetate	7.3
16	Sodium phosphate	7.5
17	Tris	7.5
18	Imidazole	8.0
19	Hepes	8.0
20	Tris	8.0
21	Bicine	8.0
22	Tris	8.5
23	Bicine	9.0

90 °C in increments of 0.2 °C. Fluorescence changes in the wells of the plate were monitored simultaneously with a charge-coupled device (CCD) camera. The wavelengths for excitation and emission were 490 and 575 nm, respectively.

Table 3
List of the additives and their concentrations used in the screen

Number	Additive	Final concentration (mM)
1	FAD	2.5
2	NAD	5
3	Lysine	5
4	β-Octylglucoside	5
5	NADP	5
6	Proline	5
7	AMP–PNP	2.5
8	$FeCl_2$	15
9	GDP	5
10	MgAc	15
11	$MnCl_2$	15
12	Glucose	5
13	$CuCl_2$	15
14	ATP	5
15	$CoCl_2$	15
16	Mannose	5
17	Fructose	5
18	Maltose	5
19	CaAc	15
20	$NiCl_2$	15
21	NADH	5
22	Haemin	2.5
23	UTP	5
24	DM	10
25	DDM	10
26	Cholic acid	10
27	Chaps	10
28	Glycerol	10%
29	Vanilic acid	10
30	$ZnCl_2$	10
31	Glycine	10
32	Phenylalanine	10
33	4-Hydroxy benzoic acid	10
34	Protoporphyrin	10
35	Coproporphyrin	10
36	Trimethanine	10
37	Heptanetriol	10
38	Benzamidine	10
39	NADPH	10
40	Galactose	10

Note. The first 23 additives were used in the first screen. The last 17 additives and a subset of the first 23 were used in the second screen. MgAc, magnesium acetate; CaAc, calcium acetate; DM, decyl maltoside; DDM, dodecyl maltoside.

Analysis of thermal shift data

To obtain the temperature midpoint for the protein unfolding transition, T_m, a Boltzmann model was used to fit the fluorescence imaging data obtained by the CCD detector using the curve-fitting software XLfit 4 (ID Business Solutions):

$$I = \left(A + \frac{(B - A)}{1 + e^{(T_m - T)/C}}\right), \quad (1)$$

where I is the fluorescence intensity at temperature T, A and B are pretransitional and posttransitional fluorescence intensities, respectively, and C is a slope factor.

Data points after the fluorescence intensity maximum were excluded from fitting. In the absence of ligands (con-

trol), $T_m = T_0$, and the ligand-dependent changes in midpoint temperature, $\Delta T_m = T_m - T_0$, could be calculated for each well.

Crystallization

Protein samples were shipped on dry ice to the high-throughput crystallization screening facility at the Hauptman–Woodward Medical Research Institute to identify initial crystallization conditions [8].

Crystallization experiments were set up using the microbatch under oil technique [9]. Each of the 1536 experiments contained 200 nl of crystallization cocktail solution combined with 200 nl of protein solution (US patent 6,296,673) under 5 µl of paraffin oil in a 1536-well plate (Greiner Bio-One). The protein solution contained protein at 8 mg/ml and stabilizing or destabilizing additives. Control plates without stabilizing or destabilizing additives were set up for all proteins. The final concentration of the additives in the protein/cocktail drop was 10 mM, and drops were dispensed using a Tango Liquid Handling System. The protein samples were screened against 1536 crystallization cocktails, including 552 from Hampton Research. Plates were stored at 23 °C and periodically imaged weekly over a 1-month time course. (Data are available on request from the corresponding author.)

Results

Initial protein screen

An initial experiment was performed with the fluorescent dye Sypro Orange and 25 proteins (Fig. 1 and Table 1). Sypro Orange has low quantum yield in aqueous solution but is highly fluorescent in nonpolar environments with low dielectric constants such as hydrophobic sites in proteins. When a protein starts to unfold or melt, the dye binds to exposed hydrophobic parts of the protein, resulting in a significant increase in fluorescence emission. The fluorescence intensity reaches a maximum and then starts to decrease, probably due to precipitation of the complex of the fluorescent probe and the denatured protein. No clear melting curves were given by four of the proteins, and these proteins were not used in the remaining experiments. It is noteworthy that difficulties with concentrating the protein after purification were observed for these four proteins.

Buffer screens

Structural genomics projects are searching for automated and robust methods for protein purification, commonly by using the same buffer conditions for all individual protein targets. Using the buffer screen, one or a couple of buffers that are likely to stabilize a majority of the proteins can be identified. Another application of the buffer screen is to find a suitable buffer for a particular protein, especially if the protein aggregates or precipitates from solution in the ordinary buffer of choice given that this might be correlated with protein stability.

The buffer screen, consisting of a set of 23 different buffers (Table 2), each at a concentration of 100 mM and with a pH range from 4.5 to 9.0, was performed with 17 *E. coli* proteins (Fig. 2).

The T_m values measured in buffer were compared with the T_m values for the control experiments, and the change in unfolding temperature, ΔT_m, could be calculated. A positive ΔT_m can be coupled to an increase in structural order and a reduced conformational flexibility, whereas a decrease in stability, negative ΔT_m, indicates that the buffer induces protein structural changes toward a more disordered conformation or it can be a sign of misfolding.

The ΔT_m values for all 17 proteins were determined for each of the 23 buffers, and the average ΔT_m could be calculated. Relatively large variations in stability could be observed when varying the buffers (Fig. 2), and a few

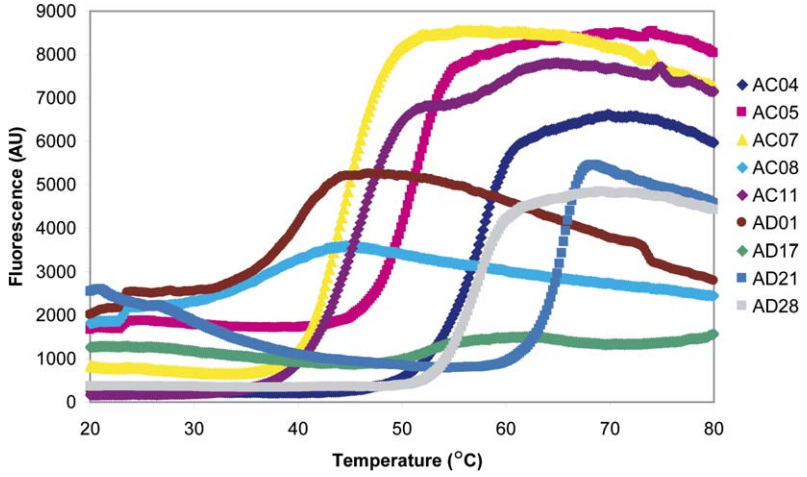

Fig. 1. Thermal shift assay result for the initial screen of nine different proteins. A significant difference in the shape of the curves could be observed between the proteins that have crystallized (e.g., AC04, AC07, AD28) and the proteins that have not crystallized (e.g., AC08, AD01, AD17).

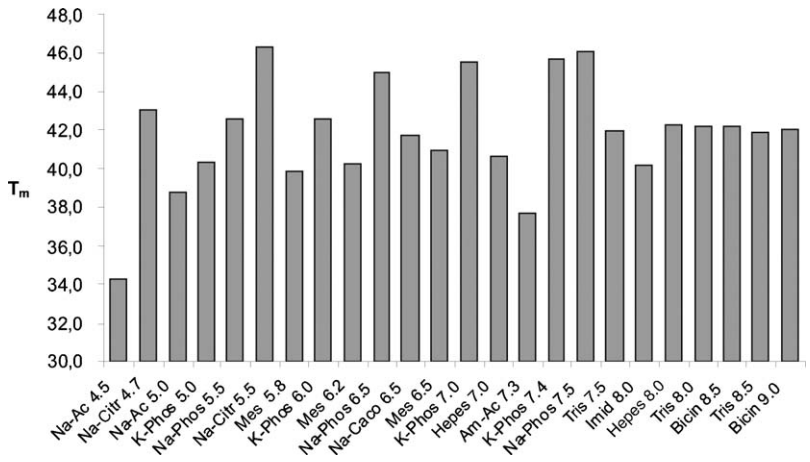

Fig. 2. Midpoint temperatures of the protein-unfolding transition (T_m) for AC07 in the presence of the buffers listed in Table 2. Na-Ac, sodium acetate; Na-Citr, sodium citrate; K-Phos, potassium phosphate; Na-Phos, sodium phosphate; Na-Caco, sodium cacodylate; Am-Ac, ammonium acetate; Imid, imidazole; Bicin, bicine.

buffers appear to be generally more favorable for protein stabilization. Potassium phosphate pH 6.0 and 7.0, sodium phosphate pH 6.5 and 7.5, and sodium citrate pH 5.5 were the most stabilizing buffers, whereas sodium acetate pH 4.5 and 5.0, sodium citrate pH 4.7, and Mes pH 5.8 significantly destabilized several of the proteins (Fig. 3). The destabilizing effect might be due to the acidic pH and therefore not an intrinsic property of the buffer.

Several of the 23 buffers investigated in the buffer screen gave no measurable transitions in combination with one or a few of the proteins, possibly caused by destabilization/partial unfolding and potentially aggregation of the proteins. For example, a clear thermal transition could be detected with only 11 of the 17 proteins in combination with sodium acetate pH 4.5. However, for some of the buffers, among them the overall most stabilizing buffers, a measurable transition could be recorded together with all of the proteins.

Sodium citrate pH 5.5 gave the highest average ΔT_m value for all 17 proteins, and it was also the buffer that gave the highest positive shift in melting temperature for 6 of the individual proteins. Even though sodium acetate pH 5.0 was one of the buffers that reduced the thermally induced melting temperatures the most in general, it was also the buffer that gave the highest ΔT_m value for 1 of the proteins. This emphasizes the need to screen a wide set of conditions to identify the best conditions for each individual protein. The three most stabilizing buffers were determined for each individual protein, and it is interesting that sodium citrate pH 5.5 was ranked among the top three in 8 of the 17 cases. Other buffers also performing well were potassium phosphate pH 6.0 and 7.0, occurring seven times each among the top three buffers. Together, sodium citrate pH 5.5 and potassium phosphate pH 6.0 and 7.0 were the three most stabilizing buffers for 11 of the 17 proteins and in fact included the most stabilizing buffer for nine of the

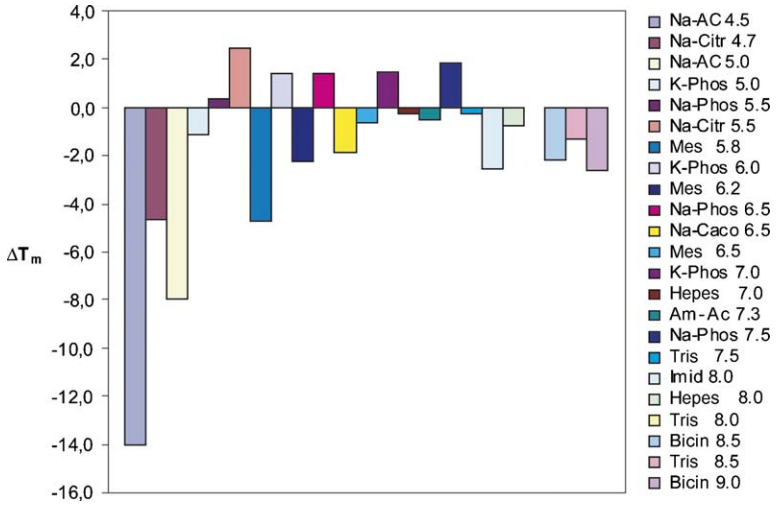

Fig. 3. Changes in the unfolding transition temperature (ΔT_m) were calculated for measurements on 17 proteins in 23 buffers. The bars represent the median ΔT_m values. A negative median ΔT_m value signifies that the buffer destabilizes the proteins, and a positive ΔT_m value indicates that the buffer has a stabilizing effect. For abbreviations, see Fig. 2 legend.

proteins. Therefore, these three buffers could constitute a suitable minimal buffer screen in high-throughput protein production projects.

Most of the buffer screen experiments in the current study were done in triplicates, and although some variation in the fluorescent intensities could be observed, this did not significantly affect the fitted T_m The standard error was calculated to be less than 0.6 °C, and this discrepancy might be due to differences in excitation intensities across the plate or to the optical edge effects from the CCD detector. In this study, we considered shifts in ΔT_m larger than 2 °C to be significant.

The molecular effect of a buffer solution on a specific protein is expected to be the pH control of titrating groups on the protein, effecting local protein structures and potentially also global ionic effects. However, specific interactions between buffers and proteins are not uncommon; more than 250 structures in the Protein Data Bank (PDB) contain ordered Mes, Hepes, or Tris [10]. Therefore, a buffer molecule that gives a positive ΔT_m in the thermofluor screen might also bind a protein and thus be useful in a crystallization experiment. We have not performed any analysis so far to investigate the correlation between the stabilizing buffers and the buffer and pH where protein actually crystallizes.

Buffer system

A buffer typically has a quite narrow pH range, but the three-component buffer system [10] allows a wide pH range to be screened without altering the composition of the buffering chemicals. The buffer system can be used to replace the ordinary buffers in the commercial crystallization screens. It allows two new chemicals to be sampled in the same experiment, and a wider pH range can be screened in optimization trials.

One of the three-component buffer systems, the citric acid–Ches–Hepes system that covers the pH range from 4 to 10, was used to screen 8 proteins (Fig. 4). An optimal pH range, in which the proteins were most stable, could be identified. It was commonly observed that the thermally induced melting temperatures were low at low and high pH values, suggesting a structural change due to titrating surface groups. However, as expected, the proteins were stable over a relatively broad pH range. The buffer system screen can be used to define the pH range where the proteins have a high thermal stability and a conformation that favors crystallization. Subsequent crystallization experiments could potentially be focused in this region by varying only salts, additives, and precipitants.

The pH dependence for all proteins tested in combination with the buffer system was relatively continuous (Fig. 4). This smooth behavior of the pH dependence was not seen when the pH dependence for the buffer screen described above, where different buffer components are used in different pH regions, was analyzed. Instead, a rather strong dependence on the buffer composition was observed. This suggests that solvent-specific factors other than the pH are equally important for determining whether or not a buffer is stabilizing. It was also observed that some nucleotide-dependent proteins displayed greater stability in phosphate buffers than in other buffers at the same pH values, probably due to specific but relatively weak binding of the anion in the nucleotide binding pocket.

Additive screen

Ligands that are useful in cocrystallization experiments can potentially be identified using the thermal shift assay. Three sets of additive screens were used (Table 3). The first screen contained 23 conditions and was used in combination with 7 proteins (AC04, AC05, AC07, AC11, AD01, AD21, and AD28), and the second screen contained 34 additives and was used together with eight proteins (AC02, AC04, AC06, AC07, AC12, AE15, AE30, AE37, and AF75). Additive concentrations in the millimolar

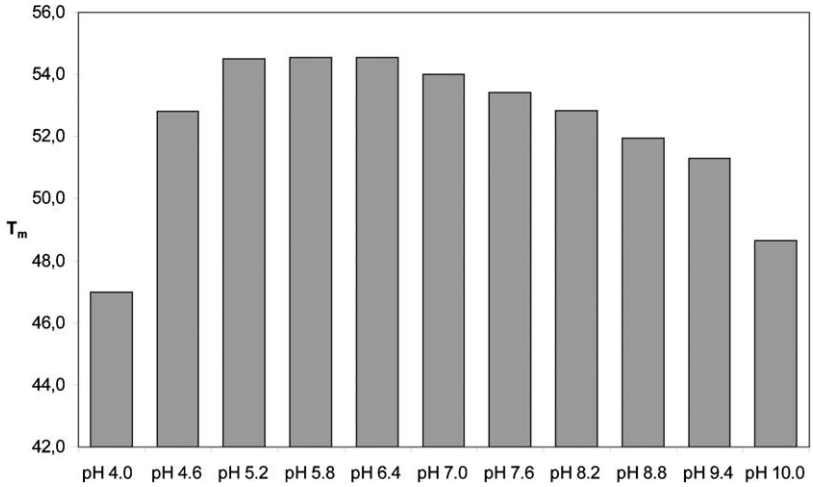

Fig. 4. Midpoint temperatures of the protein-unfolding transition (T_m) for the protein AD28 in the presence of the three-component buffer system.

range were used (Table 3). The third additive screen consisted of 23 heavy atom conditions and was used to screen eight proteins (Table 4).

Interestingly, glycerol, used in the second screen, was the only additive that significantly stabilized all of the proteins, with every other response appearing to be protein specific. Several salts were used in the screen, with a few of them (e.g., $FeCl_2$, $CuCl_2$, $CoCl_2$) appearing to quench the fluoroprobe in most cases. No transition could be recorded for any of the proteins when $FeCl_2$ was used as the additive, and $CuCl_2$ gave a measurable transition only together with one of the proteins. For most proteins, relatively few compounds were found to yield significant stabilization. However, the stabilizing additives were found to be cofactors or substrates for some of the proteins with known function, implying that functional substrates can be identified with the thermofluor method. For example, the protein AD21, which catalyzes the last step in the proline biosynthesis using NAD(P)H, was considerably stabilized by the addition of NADH (Fig. 5).

Phasing by heavy atom derivatization often is the method of choice to solve macromolecular structures. Dozens of reagents are usually tried before a suitable heavy atom derivative can be found, and the preferred technique is soaking the protein crystal in a solution of the reagent. Heavy atoms are also interesting potential additives in crystallization trials because they have an enhanced probability of binding to surface residues of proteins, thereby providing an altered surface for making crystal interactions. Consequently, the identification of binding heavy atoms can be useful both for phasing and for improving crystallization success rates.

Most of the heavy atoms in the screen either destabilized the proteins or quenched the signal, but some of them significantly stabilized the proteins; in one case, the shift in T_m was as large as 13.5 °C. Sodium hexachloroiridate(III) hydrate, ammonium tetrathiotungstate(VI), and lutetium(III) chloride hexahydrate were the only heavy atoms that gave a measurable melting curve with all 8 proteins. Sodium hexachloroiridate(III) hydrate and ammonium tetrathiotungstate(VI) stabilized a majority of the proteins, whereas lutetium(III) chloride hexahydrate destabilized all of the proteins.

Of the proteins that were used together with the additive screen, 10 were chosen for a crystallization study and six of these proteins had been crystallized previously. Three crystallization plates with 1536 conditions on each plate were set up for each protein—one plate with an additive that gave a positive ΔT_m value, one plate with an additive that gave a negative ΔT_m value, and a control plate without any additive—leading to a total of 4608 crystallization experiments. Only one of the stabilizing additives, *S*-adenosyl-L-methionine (AdoMet), was selected outside the screen based on functional knowledge of the protein. However, melting temperatures have been measured for AdoMet in complex with the protein AF75, and this additive was found to be one of the most stabilizing additives for this particular protein. For the purposes of this study, all needles, plates, and three-dimensional crystals were considered to be successful crystallizations. The 1536-condition screen has been optimized to limit the probability of forming salt crystals. If similarly looking crystals were formed in the same crystallization condition in several or all of the plates for all of the proteins, they were assumed to be salt crystals. Magnesium acetate was chosen as the additive for two of the proteins, and several crystals were formed in conditions containing phosphate; these crystals probably were salt crystals. However, none of the presumed salt crystals was included in the statistics. Of the 10 proteins, eight crystallized—three of them in only a single condition—whereas a total of 54 crystals, produced in 40 different conditions, were obtained for one of the proteins (the statistics for the individual proteins are given in Table 5). There were significant differences in the amounts of crystal hits among the different types of plates. The control plates (containing protein without ligand) gave crystals in 36 conditions, crystals were obtained in 62 conditions in the plates with stabilizing additives, and crystals were formed in only 15 conditions in the plates with destabilizing additives.

It cannot be excluded that some of these crystals are of additives. However, the large difference in the numbers of obtained crystals with stabilizing and destabilizing additives, together with the large numbers of crystallization trials analyzed, showed that by enhancing the stability of a protein with an additive, identified with our generic additive screen, we significantly increased the

Table 4
List of the heavy atoms and their concentrations used in the screen

Number	Heavy atom	Final concentration (mM)
1	Mercury(II) acetate	22
2	Baker's dimercurial	2.5
3	Ethylmercuricthiosalicylic acid, sodium salt	11
4	Chloromethylmercury	2.5
5	*para*-Chloromercuribenzoic acid, sodium salt	2.5
6	Thallium(III) chloride hydrate	19
7	Lead(II) acetate trihydrate	2.5
8	Trimethyl lead acetate	2.5
9	Cadmium chloride hydrate	10
10	Sodium hexachloroiridate(III) hydrate	5
11	Potassium hexachloroiridate (IV)	6
12	Potassium osmate(IV) dihydrate	7
13	Ammonium tetrathiotungstate(VI)	7
14	Samarium(III) acetate hydrate	4
15	Europium(III) chloride hexahydrate	3
16	Lutetium(III) chloride hexahydrate	3
17	Potassium hexachlorohenate(IV)	3
18	Ytterbium(III) chloride hydrate	2.5
19	Potassium tetrachloroplatinate(II)	2.5
20	Potassium hexachloroplatinate(IV)	4
21	Di-μ-iodobis(ethylenediamine)diplatinum(II)	4
22	Potassium tetrachloroaurate(III)	4.5
23	Gold(I) potassium	4.5

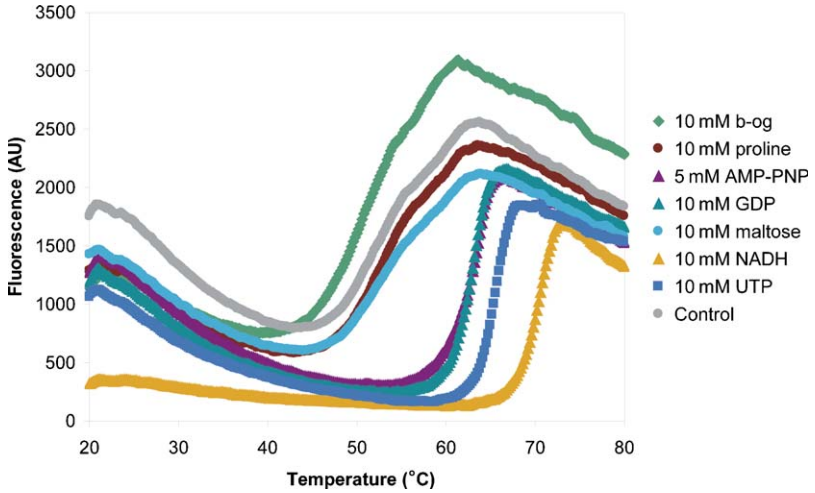

Fig. 5. Thermal shift assay result for the protein AD21 in combination with the additive screen. AD21 catalyzes the third reaction in the proline biosynthetic pathway and uses NAD(P)H in the reaction. For clarity, only selected conditions are presented. b-og, β-octylglucoside.

Table 5
Summary of results for the crystallization trials

Protein	Control	Destabilizing additive	Stabilizing additive	Total number of crystals/total number of conditions
AC04	22 (10)	7 (5)	25 (12)	54/40
AC05	0	0	0	0/0
AC07	0	0	1	1/1
AC08	0	1	0	1/1
AC11	1	0	0	1/1
AC12	8 (4)	5 (2)	7 (3)	20/13
AD21	3	1	15	19/19
AE15	0	0	0	0/0
AE30	0	0	9	9/9
AF75	2	1	5	8/8
Total	36	15	62	113/92

Additive (destabilizing/stabilizing):

AC04	Lysine/Proline
AC05	β-Octylglucoside/UTP
AC07	ATP/$CoCl_2$
AC08	$MnCl_2$/NADH
AC11	Magnesium acetate/GDP
AC12	Benzamidine/Magnesium acetate
AD21	ATP/NADPH
AE15	β-Octylglucoside/$NiCl_2$
AE30	Glucose/Cholic acid
AF75	Heptanetriol/*S*-Adenosylmethionine

Note. All needles, plates, and three-dimensional crystals were considered to be successful crystallizations. The numbers of unique crystallization conditions are presented in parentheses.

likelihood of crystallization, whereas the addition of a destabilizing additive decreased the number of crystal hits. It was also observed that a larger number of wells with precipitation were found in the plates with destabilizing additives than in the control plates and in the plates with stabilizing additives, stressing that there might be a correlation between protein stability and solubility.

The protein AC04 gave crystals in 22 conditions in the control plate and 25 crystals in the stabilizing additive plate. These 47 crystals were produced in 35 different conditions; hence, 13 new crystallization conditions were recognized with the addition of a stabilizing additive. The crystallization experiment gave a total of 113 crystal leads in 92 different conditions. Crystals produced in various conditions, potentially being other crystal forms, can exhibit different properties regarding, for example, diffraction quality and the possibility of finding suitable cryoprotectants, and a diversity of crystal leads for optimization is a great advantage.

Discussion

In current structural genomics initiatives focused on bacterial proteins, only some 3–15% of the genes entering the pipeline end up as structures. At the same time, more than 50% of proteins are expressed in a soluble form. Many proteins can be purified and yield lead crystals, but the optimization of these crystals often is difficult or too time-consuming, and new proteins are entered into the pipeline to keep productivity instead of solving the structures of the difficult proteins. Many crystal optimization problems probably are due to protein quality problems, and generic high-throughput strategies for optimizing homogeneity, stability, and solubility of protein samples will significantly improve success rates of crystallization and minimize the effort in going from lead crystals to diffraction-quality crystals. However, it is inherently difficult to find a method to predict conditions for successful protein crystal growth a priori. Attempts have been made to correlate NMR data with protein crystallization, but these methods have been suggested to be complementary rather than correlated [11].

Another method commonly used to characterize the protein sample prior to crystallization is DLS. The likelihood of successful crystallization is lower if the DLS size

distribution is polydispersed [2], and DLS can be used for identifying protein samples that are unlikely to yield crystals. However, the method has not been used extensively for larger buffer and ligand screens because it has required relatively large amounts of protein.

The results from our thermofluor-based additive screen indicate that this constitutes a generic approach for increasing the likelihood of protein crystallization. The inclusion of an additive, which gives a positive shift in the transition midpoint in the thermofluor screen, in the crystallization experiments for this test set of proteins increased the probability of obtaining protein crystals to some 70% compared with the control experiments. The addition of a destabilizing additive decreased the likelihood of getting crystals to approximately 50%. Both the stabilizing effect of the additive, yielding more crystals, and the additional orthogonality (i.e., having another version of the protein) are valuable for improving crystallization success rates. The thermofluor method also appears to be able to directly identify heavy atom binding. In a few cases, heavy atoms had very strong stabilizing effects and therefore can be useful in cocrystallization trials, enabling the use of the heavy atom for phasing. However, sodium hexachloroiridate(III) hydrate, ammonium tetrathiotungstate(VI), and lutetium(III) chloride hexahydrate, which were found to bind to the proteins in this study, all are likely to bind in pockets through ionic interaction rather than as covalent-type complexes to side chains such as Cys, His, and Met that often are formed on quite exposed and flat surfaces. Covalent attachments probably are less likely to have dramatic effects on stability and might not be visible in the thermal shifts.

The results from the buffer screen suggest that the most stabilizing buffer for each protein can be rapidly identified with the thermofluor method. A more stable protein generally has a lower tendency for aggregation. However, it is likely that the correlation between intrinsic protein stability, on the one hand, and intrinsic protein solubility, on the other, is not going to be absolute. On the contrary, there are factors that are likely to affect mainly solubility, such as ionic shielding effects, hydrophobic surface effects, and exposed sulfhydryl groups, and these effects might not be observed in the thermal melting curves.

Preliminary data indicate that the shape of the melting curves in the initial protein screen might give information about the condition of the protein. All of the proteins in our small sample set that have crystallized had a clear melting transition combined with a low initial fluorescence intensity, whereas the melting curves for several of the proteins that have not crystallized started at a high-fluorescence intensity and were lacking a clear transition. The high initial fluorescence intensity is likely to be indicative of a high hydrophobicity of the whole protein sample or a fraction of the protein sample. In most cases, no detectable melting curves could be observed when detergents were used in the additive screen, possibly due to detergent interactions with the fluoroprobe. However, balancing the ratio of protein concentration with the total amount of detergent present in the sample enabled Pantoliano and coworkers to conduct thermofluor experiments on the membrane protein bacteriorhodopsin [3].

Several of the proteins used in the current study showed sharp two-state melting transitions even though they were multidomain proteins. This was noted previously by Pantoliano and coworkers, who proposed a cooperative unfolding of these proteins due to energetic coupling between the domains [3]. An alternative explanation for this observation is the preferred and strong binding of the dye to one of the domains during the unfolding process.

In our limited set of proteins, there is no direct correlation between the T_m of a protein and the probability of obtaining crystals when comparing different proteins. However, when the same protein is screened for optimized additives/T_m, a good correlation with crystallizability is seen, as discussed above.

Homogeneity of a protein sample is potentially the most important factor for successful crystallization. The binding of the ligand has the potential to increase the homogeneity of the sample, for example, by minimizing flexibility. This is consistent with the observation that the melting curves often are slightly steeper for the proteins with stabilizing ligands than for the proteins without ligands. This indicates that the unfolding is more cooperative when ligands are present and therefore that the sample has a higher homogeneity. Furthermore, stabilizing buffers are likely to preserve the sample in a more homogeneous form for a longer time than are less stabilizing buffers, thereby suggesting that these buffers should be used during lysis and purification. The three best-performing buffers identified in this study might be a good extension of the single purification buffer approach pursued in most structural genomics initiatives. In the future, when large data sets of thermofluor data are collected using a standardized set of buffers and additives, it might be possible to predict crystallizability from the shape of the curve alone. Thus, by using only a small amount of sample, it would be possible in an early state to select the constructs and ligand complexes most likely to crystallize.

Acknowledgments

We thank the Swedish Research Council, the Wallenberg Consortium North, NIH GM074899, and the EU framework V Network SPINE for financial support. We thank Benita Engvall, Marie Hedren, and Martin Andersson for the initial cloning of the proteins used in this study. We also thank Alexei Savchenko for providing some of the clones and Tove Sjögren (AstraZeneca) for help with preparing ligands.

References

[1] B.W. Segelke, Efficiency analysis of sampling protocols used in protein crystallization screening, J. Cryst. Growth 232 (2001) 553–562.

[2] M. Zulauf, A. D'Arcy, Light-scattering of proteins as a criterion for crystallization, J. Cryst. Growth 122 (1992) 102–106.
[3] M.W. Pantoliano, E.C. Petrella, J.D. Kwasnoski, V.S. Lobanov, J. Myslik, E. Graf, T. Carver, E. Asel, B.A. Springer, P. Lane, F.R. Salemme, High-density miniaturized thermal shift assays as a general strategy for drug discovery, J. Biomol. Screen 6 (2001) 429–440.
[4] M.C. Lo, A. Aulabaugh, G.X. Jin, R. Cowling, J. Bard, M. Malamas, G. Ellestad, Evaluation of fluorescence-based thermal shift assays for hit identification in drug discovery, Anal. Biochem. 332 (2004) 153–159.
[5] D.A. Tobbell, B.J. Middleton, S. Raines, M.R.C. Needham, I.W.F. Taylor, J.Y. Beveridge, W.M. Abbott, Identification of in vitro folding conditions for procathepsin S and cathepsin S using fractional factorial screens, Protein Expr. Purif. 24 (2002) 242–254.
[6] R.G. Zhang, T. Skarina, J.E. Katz, S. Beasley, A. Khachatryan, S. Vyas, C.H. Arrowsmith, S. Clarke, A. Edwards, A. Joachimiak, A. Savchenko, Structure of *Thermotoga maritima* stationary phase survival protein SurE: A novel acid phosphatase, Structure 9 (2001) 1095–1106.
[7] U.B. Ericsson, M.E. Andersson, B. Engvall, P. Nordlund, B.M. Hallberg, Expression, purification, crystallization, and preliminary diffraction studies of the tRNA pseudouridine synthase TruD from *Escherichia coli*, Acta Crystallogr. D 60 (2004) 775–776.
[8] J.R. Luft, R.J. Collins, N.A. Fehrman, A.M. Lauricella, C.K. Veatch, G.T. DeTitta, A deliberate approach to screening for initial crystallization conditions of biological macromolecules, J. Struct. Biol. 142 (2003) 170–179.
[9] N.E. Chayen, P.D.S. Stewart, D.M. Blow, Microbatch crystallization under oil: A new technique allowing many small-volume crystallization trials, J. Cryst. Growth 122 (1992) 176–180.
[10] J. Newman, Novel buffer systems for macromolecular crystallization, Acta Crystallogr. D 60 (2004) 610–612.
[11] A. Savchenko, A. Yee, A. Khachatryan, T. Skarina, E. Evdokimova, M. Pavlova, A. Semesi, J. Northey, S. Beasley, N. Lan, R. Das, M. Gerstein, C.H. Arrowmith, A.M. Edwards, Strategies for structural proteomics of prokaryotes: Quantifying the advantages of studying orthologous proteins and of using both NMR and X-ray crystallography approaches, Proteins 50 (2003) 392–399.