Available online at www.sciencedirect.com
SCIENCE DIRECT®
Methods 34 (2004) 354–363
METHODS
www.elsevier.com/locate/ymeth

The use of recombinant methods and molecular engineering in protein crystallization

Zygmunt S. Derewenda*

Department of Molecular Physiology and Biological Physics, University of Virginia School of Medicine, Charlottesville, VA 22908, USA

Accepted 24 March 2004

Abstract

Recombinant techniques are routinely used for the preparation of protein samples for structural studies including X-ray crystallography. Among other benefits, these methods allow for a vast increase in the amount of obtained protein as compared to purification from source tissues, ease of purification when fusion proteins containing affinity tags are used, introduction of SeMet for phasing, and the opportunity to modify the protein to enhance its crystallizability. Protein engineering may involve removal of flexible regions including termini and interior loops, as well as replacement of residues that affect solubility. Moreover, modification of the protein surface to induce crystal growth may include rational engineering of surface patches that can readily mediate crystal contacts. The latter approach can be used to obtain proteins of crystals recalcitrant to crystallization or to obtain well-diffracting crystals in lieu of wild-type crystals yielding data to limited resolution. This review discusses recent advances in the field and describes a number of examples of diverse protein engineering techniques used in crystallographic investigations.

1. Introduction

Crystallization is the key limiting step in macromolecular X-ray diffraction analysis, ironically more so today than during the few decades after J.D. Bernal and D. Hodgkin recorded the first X-ray diffraction from a protein crystal [1]. The legacy of Sumner, Northrop, and other pioneers of protein biochemistry, who crystallized a multitude of enzymes during the 1920s and 1930s, kept the crystallographers busy for much of the subsequent half-century. Hemoglobin, pepsin, insulin, lysozyme, and many other early targets of X-ray analysis had all been crystallized well before full structural analysis became technically possible, and the availability of crystals was not an issue. However, as the number of crystal structures tackled by crystallographers escalated in the early 1980s, it became obvious that the prospects for progress are dramatically limited by the availability of target proteins, and—more importantly—by their propensity to form good quality diffracting crystals. Protein crystallographers no longer had the luxury of picking up the next available crystal from the shelf. As there was no way to predict crystallization conditions, and no way to alter the nature of the purified protein—except, perhaps for the level of purity—the only way to crystallize a protein was to screen a wide range of possible conditions. Some of the ideas that even today constitute the core of high-throughput crystallization screens, such as the incomplete factorial method, date from that period [2].

With inefficient and labor-intensive purification protocols, and uncertain prospects for crystallization, only the most abundant and stable medium-size proteins, such as those found in bodily fluids or muscle, were considered to be feasible targets. The whole field might have stalled prematurely, had it not been for the advent of contemporary recombinant methods and heterologous overexpression. Shortly after the first successful expression of recombinant proteins in *Escherichia coli*, including insulin [3] and somatostatin [4], protein crystallographers turned to recombinant methods as the means to obtain samples for crystallization. Among the very first proteins crystallized using recombinant samples were

* Fax: 1-804-982-1616.
E-mail address: zsd4n@virginia.edu.

1046-2023/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.ymeth.2004.03.024

insulin [5], human leukocyte interferon A [6,7], murine interferon-β [8], and eglin C [9],

With the dawn of recombinant methods, modification of protein and using protein as another variable in the crystallization experiments became also possible. Prior to that point, the only way to use the protein as a variable was to screen homologues from various species. The concept was originally used by Kendrew [10] in his original selection of sperm whale myoglobin as a target for X-ray analysis. Later on, Campbell et al. [11] reiterated the idea and suggested screening of various species as a routine tool in search for a crystallizable protein homologue. The approach has been successfully used in many studies, and is still often exploited in membrane protein crystallography, where finding a suitable and crystallizable variant of the target molecule is even more of a challenge.

In spite of its obvious advantages, homologue screening is cumbersome, and not useful if a specific structure is needed. In contrast, recombinant methods offer a range of possibilities for modifying the target protein to help in the expression, purification, and crystallization. As a consequence, most protein crystallography groups have incorporated molecular biology techniques into their standard arsenal of methodologies, and often invest much more time at the bench producing the protein and crystallizing it, than solving the structure once crystals are in hand.

The aim of this review is to provide the reader with an overview of the contemporary protein engineering approaches used in a protein crystallography laboratory to facilitate the preparation of high-quality crystals. The subject is broad, and a truly comprehensive review is not possible within the limitations of this chapter. Thus, we highlight the selected important aspects and provide representative examples of case applications for globular proteins. Particular attention is given to the advances in crystallization enhancement via surface mutagenesis. We do not discuss protein engineering of membrane proteins, as the options there are very limited due to the much less advanced stage of heterologous overexpression.

2. Use of recombinant proteins in crystallization

The advantages of using recombinant proteins for crystallization typically include, among others, a dramatic increase in the yield of protein, ease of purification using affinity tags and fusion proteins, possibility of protein modification or use of isolated domains, and ease of labeling with SeMet for phasing purposes. A survey of crystallographic studies published last year in high-impact journals (excluding membrane proteins) indicates that over 90% of crystal structures are based on recombinant material (Bielnicki, J., Janda, I., Derewenda, Z.S., unpublished).

Bacteria, and in particular *E. coli*, are preferred as protein producers for biophysical studies [12]. It is the best characterized host with a diversity of strains for specific applications, such as expression of proteins with non-optimal codon bias, protease-deficient strains, strains containing chaperones for enhanced folding, etc [13]. Other advantages include ease of handling compared to eukaryotic cell cultures, multitude of vectors for expression of numerous fusion proteins, and overall low cost. However, in spite of these advantages, not all proteins of interest will be expressed with an adequate yield, in a properly folded soluble form, or at all.

A simple method to increase the yield of protein production is to increase the culture density from the typical OD of 0.6–2.0 or higher. Very high density can be achieved using disposable 2-L plastic bottles, which get well aerated in shakers and which also minimize the overall effort by disposing with sterilization [14]. The media composition may also be altered and both minimal media and rich media have been shown to be advantageous in specific cases. To further increase the yield for those proteins for which expression levels are particularly low, it is necessary to resort to fermentors, although the optimization of large-scale expression is often a significant effort in itself [15].

Protein aggregation and inclusion body formation are often a problem [16]. To monitor the solubility of the expressed protein, one can use a fusion with a downstream located GFP (green fluorescent protein). Under those circumstances, increase in fluorescence indicates proper folding of the target protein and can be used as a diagnostic [17]. If the target protein expresses well, but goes into inclusion bodies, the difficulties can often be resolved using a number of alternative strategies. The simplest remedy is to lower the IPTG inducer concentration to ~0.2 mM and reduce the post-induction temperature [18], so that the translation machinery slows down allowing the polypeptide chain enough time to fold. For example, reduction of the temperature from the typical 30–37 °C to 25 °C resulted in a 2-fold increase in the soluble fraction of apoglobin in *E. coli* [19]. A similar result was obtained for the phage K11 RNA polymerase, which required a temperature of 25 °C for proper folding and solubility [20]. In the case of the hepatitis C virus NS3 serine protease, maximum solubility was achieved when the temperature was dropped to 15 °C [21]. In our laboratory, we routinely bring the temperature to 10 °C with very good results (Derewenda, U., unpublished).

If the temperature does not help, it may be necessary to use an *E. coli* strain co-expressing molecular chaperones belonging to the DnaK–DnaJ–GrpE and GroEL–GroES systems [22]. The 50 kDa *E. coli* trigger factor, a peptidyl-prolyl *cis*/*trans* isomerase, can also significantly help to prevent protein aggregation, with or without the assistance of chaperones [23]. Finally, it should not be forgotten that some carrier proteins used in fusion

constructs for expression (see below) significantly increase the amount of folded, soluble protein in expression experiments.

Low expression levels in *E. coli* may be caused by the presence of rare codons within the target gene [24]. This can be dealt with in two ways. The first option is to use strains adjusted with additional copies of rare tRNA genes. Among the most popular are BL21(DE3)RIL cells containing extra copies of *argU*, *ileY*, and *leuW* genes, and the RP variant of the same strain which contains extra copies of the *argU* and *proL* tRNA to overcome expression problems of genes containing either AGG/AGA or CCC codons in CG-rich genes [25]. In some cases, the yield of recombinant protein obtained using these cells can be over 100 times more than using conventional BL21(DE3) cells [26]. The second option is to use wholly or partly synthetic genes, with optimized codons. For example, in the case of the gene encoding the herpes simplex virus 1 protease the use of a synthetic gene with optimized codon usage leads to a 20-fold increase in the expression yield [27].

Interestingly, rare codons may not always be the reason behind low expression related to the cDNA sequence. In their studies of the poorly expressing S1 dehydrofolate reductase (DHFR), Dale et al. [28] found that the replacement of all rare codons in the first 66 bases did little to help expression. However, replacement of the N-terminal 18 out of 22 amino acids by the corresponding codons from the SaDHFR, which is 80% to the type S1 enzyme, led to a dramatic increase in expression levels. It is possible that subtle issues involving RNA secondary structure are at play.

Finally, increasing intrinsic protein solubility via protein mutagenesis of surface residues (see below) can also significantly boost expression levels and yields. For example, four different mutations of interferon-γ increased the relative amount of the protein in the soluble fraction by 1-fold [29]. A similar result was reported for the S1 DHFR, as we discuss below [28].

In spite of all the available options, many eukaryotic proteins cannot be produced in bacterial hosts due to a multitude of factors, including but not limited to diverse post-translational modifications. The solution is to use more labor-intensive and time-consuming eukaryotic hosts such as insect cells, baker's yeast, *Pichia pastoris* or CHO (Chinese hamster ovary) cells. These systems are significantly more labor-intensive, and a limited scope of this review precludes their detailed presentation.

Finally, progress is being made on cell-free systems for large-scale protein production [30]. Although at the time when this review was written, these systems did not yet constitute an alternative. Due in part to a very high cost, they are expected—when optimized—to revolutionize the field in the possibly not too distant future.

3. The use of fusion proteins for expression and purification

Our literature survey indicated that approximately 75% of proteins for crystallization are expressed as fusion constructs [31] using a number of small proteins and tags for affinity purification as well as for solubility enhancement. A wide choice of systems is available for bacterial hosts, and a smaller but ever expanding selection is available for eukaryotic cells. Among the fusion partners particularly popular among structural biologists are the hexaHis tag [32], GST (glutathione-*S*-transferase) [33], and MBP (maltose-binding protein) [34]. Some of the less commonly used include thioredoxin [35], Z-domain from protein A [36], NusA [37] GB1-domain from protein G [38], and a number of others [13]. None of the tags is universally superior, and often parallel expression experiments determine what the best strategy is. Families of vectors allowing for parallel assays of this kind have been developed [39,40].

A survey of recent crystallographic studies reveals that nearly 60% used one of the several incarnations of the polyHis tag, making it by far the first choice among crystallographers. The popularity is due to the relative simplicity of Ni-affinity chromatography and the fact that the tag itself need not necessarily pose a major obstacle in crystallization trials. While some groups advocate the need to remove proteolytically the tag prior to crystallization, others argue against it or use a two-tier screening process whereby the first screen uses tagged proteins, and the second screen is performed after tag removal only for those proteins that failed to crystallize during the first attempt. Many examples of proteins crystallized with the His-tag left intact are reported in the literature. The His-tag works well with soluble, small to medium size proteins, but is known in many cases to lower the solubility. Interestingly, a comparative study of several N-terminal fusion tags and 32 human proteins revealed that most protein tags are superior to the hexaHis tag with respect to conferring solubility, and two—thioredoxin and MBP (maltose-binding protein)—appear to be particularly powerful [41]. Thus, if the His-tagged protein is not adequately expressed as a soluble protein, the investigator should immediately consider changing the fusion protein.

A distant second in popularity among crystallographers as a fusion tag is GST. This tag is useful as a solubility enhancer, but it also creates some problems due to its dimeric state, and is not appropriate for expression of homo-oligomers because of the possibility of forming aggregates. Finally, MBP has recently received a lot of attention due to its potential to yield soluble proteins otherwise difficult or impossible to express [42]. However, MBP is not very effective as a purification tag, because MBP-fusion proteins often fail to bind to the amylose resin [43]. Thus, the use of supplementary affinity tags is

advantageous [42,44–46]. A particularly effective combination includes a biotin acceptor peptide (BAP) on the N-terminus of MBP and/or a hexaHis tag on the C-terminus of the passenger proteins [44].

Following purification of the fusion proteins, the tags are normally removed from the protein by proteolytic cleavage. Due to stringent requirements for homogeneity of the protein samples going into crystallization trials, this proteolytic cleavage is typically done using specific viral proteases, such as recombinant tobacco etch virus proteinase (rTEV) [47–50], or the recombinant human rhinovirus 3C protease [46], also known as PreScission protease. Other proteases, such as thrombin and Factor X, are becoming less popular due to frequent non-specific secondary cleavage observed in many proteins [51].

4. Truncations and related modifications for crystallization

Many proteins contain highly flexible or even completely unfolded fragments that dramatically interfere with crystallization. This is particularly true of large multidomain signaling proteins, in which the unstructured linker regions often account for more than 50% of the molecule. To gain structural information about the stable fragments of such proteins, it is necessary to isolate them in pure form. This can be accomplished in three ways: elimination of flexible N- and/or C-terminal polypeptides, removal of internal flexible Ω-loops, or extraction of individual domains from a multidomain protein. The boundaries of folded segments used to be identified almost exclusively using proteolytic digestion followed by either mass spectrometry or chromatographic analysis of the products. The method was pioneered with studies of antibodies [52]. This method is still useful, and a battery of commercially available proteases with diverse specificities is available for that purpose. Limited proteolysis assisted in the crystallization of a plethora of proteins, including the Gtα subunit of transducin [53], and the Gtβγ dimer [54]. More recently, the eukaryotic nuclear cap-binding complex (CBP) was crystallized after mild trypsination. This complex is a heterodimer comprising a 790 residue-long CBP80 and a CBP20, made up of 156 residues, and it crystallizes poorly. Limited proteolysis resulted in the removal of the first 19 residues, as well as the 671–684 loop of CBP80, and also of the removal of the N-terminal 21 residues and an internal dipeptide 77–78 from CBP20. This truncated complex yielded crystals diffracting to 2.0 Å resolution [55].

In spite of its undisputed potential, limited proteolysis may introduce heterogeneity into the samples, and for proteins that can be overexpressed it is better to engineer the necessary truncations at the cDNA level. Proteolysis can provide useful guidelines, but recent advances in our knowledge of protein domains, and in structure prediction, allow for ab initio design of expression clones. Domain boundaries within large proteins can be predicted with approximately 63% success rate using sequence information alone [56]. Biochemical and functional studies also help to identify regions of the protein that interfere with crystallization.

Truncations and—to a lesser extent—deletions constitute the most frequently used protein engineering tool in a contemporary crystallographic laboratory, and there are numerous examples of this approach in the literature. For example, C-terminal truncation is a routine way to express and crystallize cytosolic GTPases of the Ras and Rho families. These proteins contain an extended C-terminal which undergoes post-translational modification including prenylation or acylation of the cysteine, proteolytic removal of the terminal tripeptide, and finally carboxymethylation [57]. These modifications are not carried out in *E. coli*, and substantially reduce the protein's solubility. In contrast, a C-terminally truncated version can be easily overexpressed in bacteria, as was demonstrated in the case of Ras [58]. Subsequently, the same protocol was applied to Rac1 [59], RhoA [60], and others. Another example is the 24 kDa fragment of the DNA gyrase B from *Staphylococcus aureus*. The apoenzyme form does not crystallize. Based on the structure of the complex, the authors identified a flexible loop encompassing residues 105–127 and deleted it at the level of the cDNA [61]. The modified protein yielded crystals diffracting to 2.2 Å. Finally, perhaps one of the most dramatic examples of successful protein engineering designed to obtain crystalline protein was that of the HIV-1 gp120 glycoprotein [62]. The variant which was crystallized in complex with the CD4 receptor and a neutralizing human antibody had deletions of 52 and 19 residues at the N and C termini, respectively, and Gly-Ala-Gly substitutions for 61 V1/V2 loop residues and 32 V3 loop residues, in addition to extensive deglycosylation. This construct was based on extensive limited proteolysis experiments which defined the boundaries of the stable core [63].

5. Enhancing protein solubility by protein engineering

One of the common problems in crystallization experiments is the poor solubility of the target protein. Even if the protein is expressed predominantly in the soluble fraction (see above), particularly if it is expressed as a fusion protein, it may precipitate at concentrations required for crystallization after it is cleaved from the carrier protein. Provided that the protein is stable and properly folded, solubility is the function of surface hydrophobicity and can be dramatically altered by mutational modification of selected surface residues. In the absence of a plausible model, this can be achieved by scanning mutagenesis targeting hydrophobic residues.

Such strategies lead to the crystallization and structure determination of the F185K mutant of the HIV-1 integrase [64,65]. Similarly, a soluble and crystallizable form of the Moloney murine leukemia virus reverse transcriptase was obtained using the L435K mutant [66]. If a structure of a homologous protein is available, it can dramatically simplify the problem by allowing for direct identification of hydrophobic residues exposed to solvent. Dale et al. [28] followed this rational protocol in search of a more soluble variant of a type S1 hydrofolate reductase, by replacing exposed amide-containing side chains with carboxylates. Of the four prepared mutants three had elevated solubility and one easily crystallized.

An alternative, rarely used approach is to introduce a solubilizing motif. Bianchi et al. [67] designed N- and C-terminal tri-Lys tags for a poorly soluble (10 μM) synthetic minibody protein (61-residues). They found that either tag was able to increase the solubility by two orders of magnitude, without affecting the stability of the protein. A similar effect should be expected for a poly-arginine variant, which could also serve as an affinity tag [68,69].

6. Removal of glycosylation sites and removal of unpaired cysteines

Expression of proteins in eukaryotic hosts may lead to N-glycosylation. As the glycan moieties typically interfere with crystallization, it is advisable to remove as many as possible. This can be accomplished by the replacement of glycosylated asparagines with other residues, typically glutamines or aspartates. Human butyrylcholinesterase is heavily N-glycosylated at nine sites, and in its wild-type form was found to be recalcitrant to crystallization. Nachon et al. [70] removed four out of nine sites by replacing the modified Asn residues with Gln, and obtained a stable mutant that formed high-quality crystals diffracting to 2.0 Å resolution. Similarly, the extracellular erythropoietin receptor was crystallized using a N52Q mutant which removed the only glycosylation site expressed in *Pichia pastoris* [71]. This type of engineering is possible provided the non-glycosylated mutant protein folds properly. If the protein requires glycosylation for proper folding, the expressed samples can be deglycosylated using recombinant glycosidases [72]. To further facilitate enzymatic deglycosylation, one can use specific cell lines unable to process *N*-glycans beyond the endo-H sensitive Man5GlcNAc2 [73]. This approach helped to crystallize samples of the human co-stimulatory molecule B7-1 [74].

Exposed, easily oxidized sulfhydryl groups constitute another frequent source of problems leading to sample heterogeneity or aggregation, typically associated with non-native, intermolecular disulfide bridges. The use of reducing agents is often ineffective, and it is better to replace the culprit Cys residue by mutagenesis. For example, Stover et al. [75] mutated two free cysteines in *E. coli* chorismate lyase to serines, thus overcoming serious problems of protein aggregation. The mutant yielded three high-quality crystal forms, including one diffracting to atomic (1.1 Å) resolution. Finally, it is worth mentioning that even in the absence of aggregation and oxidation, replacing Cys with Ser can dramatically alter crystallization kinetics, as was shown in the case of bovine γB crystallin [76].

7. The use of molecular scaffolds for protein crystallization

When the target protein fails to succumb to crystallization efforts, one of the possible options is to use another, easily crystallizable protein as a scaffold for the crystal lattice. It is well established that some proteins crystallize well as a part of a complex, but not alone. This principle can be exploited further, even if the complex at hand has no biological significance, other than to promote the crystallization of one of its components.

Potentially the simplest way to create a scaffold is to use the affinity tag on the target protein, in the hope that it will generate lattice-forming interactions. This is normally very difficult due to the intrinsic flexibility of fusion proteins conferred on them by the linker segments. The presence of a fusion carrier protein in the crystal would not constitute a problem per se. On the contrary, a stable, easily crystallizable module rich in methionines would be very beneficial to crystallographic studies.

Recently, several structures of fusion proteins containing MBP as an affinity tag were successfully crystallized and characterized by X-ray diffraction: the human T cell leukemia virus I gp21 ectodomain [77]; the SarR protein [78]; and the MATa2 homeodomain [79]. The key to success in these cases was the use of short linkers, such as AAA, instead of the typical oligopeptides containing proteolytic sites (reviewed by Smyth et al. [80]). In another interesting case, diffracting crystals of the ribosomal L30 protein in fusion with MBP were obtained without any modifications to the linker region [81]. Preliminary crystallization reports have also been published for the DNA-binding domain of a replication-related element-binding protein DREF in fusion with GST [82]; VanH, a D-lactate dehydrogenase in fusion with thioredoxin [83]; and the extracellular domain of CD38 in fusion with MBP [83].

Using fusion proteins to generate scaffolds is an interesting alternative, particularly if the passenger protein is properly folded but insoluble in the absence of the carrier MBP, as is often the case. An additional advantage is that structure determination is facilitated by the fact that the molecular replacement method can be directly applied to native data, since an accurate model of MBP

has been determined. In the known examples, MBP mediates most of the crystal contacts in the crystal structures, and is without doubt responsible for the formation of the crystal. However, so far all proteins successfully crystallized by this method are small, and it is not known if the method will be generally applicable to larger molecules.

An alternative way to create scaffolds is to use Fab fragments of antibodies raised against the protein of interest, or to fuse the protein of interest to an antigen that binds a Fab fragment [84,85]. The former approach is inefficient, as it requires raising antibodies and obtaining recombinant Fab fragments for each target protein. Nonetheless, it does have considerable potential in the field of membrane proteins (see the Chapter by M.C. Wiener in this volume).

8. Preparation of recombinant oligomeric complexes for crystallographic studies

Many key physiological phenomena are regulated by signaling cascades which involve transient oligomeric protein complexes. Structural studies of such complexes are of high importance as they help understand the molecular basis of signal transduction. In the simplest possible case, the components can be expressed separately, and mixed prior to crystallization, as was the case with the 53BP1 BRCT domains crystallized in complex with the p53 tumor suppressor [86]. The complex may be purified to homogeneity after the individual components are mixed using gel filtration chromatography, or a stoichiometry mixture may alternatively be used directly. However, in many cases the problem is much more difficult, particularly if one of the components is unstable in the unbound form. Similar problems arise when an oligomeric protein, such as a multisubunit enzyme complex, is being studied.

One of the ways to overcome this problem is to co-express the two components using independent vectors. A study of two large multiprotein complexes, the basal transcription factors TFIID and TFIIH, showed that coexpression of appropriate combinations of subunits leads to increase of solubility and stability of the proteins [23]. The small GTPase ARL2 and its effector protein, the δ subunit of human cGMP phosphodiesterase (hPDE δ), were co-expressed using individual plasmids, and the complex was purified by gel filtration and successfully crystallized; it was noted that co-expression significantly increased the otherwise low yield of the PDE δ expression in *E. coli* [87].

An alternative approach is to use polycistronic vectors, which code for all the subunits within the target complex. The glycine decarboxylase (P-protein) of *Thermus thermophilus*, an $\alpha_2\beta_2$ tetrameric complex of a molecular mass of 200 kDa, was successfully expressed in *E. coli* using a bicistronic vector based on the pET11 system; the complex was subsequently crystallized [88]. The mammalian AMP-activated protein kinase is a heterotrimeric complex composed of one catalytic and two regulatory components. Successful expression was accomplished in *E. coli* using a tricistronic expression plasmid, with a polyHis tag on the catalytic subunit alone [89]. A bicistronic vector was also a key to expression and crystallization of the catalytic α_1/α_2 heterodimer of the brain plasma-activating factor acetylhydrolase [90]. The individually expressed subunits form tightly associated homodimers, and no exchange of subunit occurs when they are mixed. The only way to prepare the heterodimer was to synthesize it in bacteria in situ, and use temperature to optimize the yield compared to the two possible homodimers.

Finally, it is possible to use the aforementioned method of detecting folded protein using GFP, to visualize the folding of components of heterooligomeric complexes [91].

9. Enhancing protein propensity for crystallization by surface mutagenesis

In a classic study of the impact of point mutations on the protein's surface on the propensity of the protein to form crystals, Villafranca and co-worker [92] studied 12 single site mutants of thymidylate synthase and assessed their behavior in crystallization screens. They reported that the mutations dramatically affected the protein's solubility and its ability to crystallize, although there was no clear correlation between the two. These observations were subsequently corroborated by other similar studies, including the study of nine mutants of the 24-kDa fragment of the *E. coli* DNA gyrase B subunit [93]. More recently, the crystallization of aspartyl-*t*-RNA synthase (AspRS-1) was studied using seven mutants affecting intermolecular interactions [94]. It was concluded, as expected, that crystallization is very sensitive to the integrity of the crystal contacts.

In addition to these test studies, a number of other proteins were reported crystallized in mutated form when wild-type crystals were unavailable, but in most cases by serendipity. For example, GroEL was crystallized using a variant with two mutations accidentally introduced by PCR [95,96]. Rational engineering of crystal contacts is a much rarer approach. The first use of site-directed mutagenesis for crystal engineering was reported by Lawson et al. [97] who reproduced crystal contacts from rat L ferritin in human ferritin H-chain. A replacement of Lys 86, found in the human sequence, with Glu which occurs in rat, recreated a Cd^{2+}-binding bridge which mediates crystal contacts in the rat orthologue. More recently, Wingren et al. [98] proposed to engineer 'cassettes' of packing motifs into the external

β-strands of antibodies to create suitable crystal contacts. However, neither of the approaches proposed to date seems to be ubiquitously applicable to all classes of proteins.

10. Protein crystallization by local reduction in surface conformational entropy

Recently, we proposed yet another approach to crystallization which involves mutational engineering of surface patches designed to mediate crystal contacts [114]. Kwong et al. [63] argued that the probability of obtaining a good quality crystal is dependent on the presence of conformationally homogeneous patches on the surface of the protein. We postulated that such patches can be created if selected residues with high conformational entropy are replaced with alanines. Such replacement removes surface side chains that would otherwise be immobilized at the crystal contact with thermodynamically unfavorable loss of entropy. One should remember that the ΔG of crystallization always contains a negative component of the $T\Delta S$ term, due to the ordering of protein molecules in the lattice. An additional loss of entropy due to the ordering of side chains at the crystal contacts could render crystallization thermodynamically impossible, unless there is a compensating release of water molecules from the interacting surfaces [99,100]. Consequently, proteins containing a high proportion of surface amino acids with large side chains may show lower propensity to form crystals. We hypothesized that the key residues impeding formation of crystal contacts are lysines and glutamates [101,102].

Lysines are predominantly located on the surface, with 68% exposed, 26% partly exposed, and only 6% buried [103]. The high conformational entropy of solvent-exposed lysines, ~2 kcal/mol [104], is likely to impair the formation of protein–protein contacts. Due to the large size of the side chain, lysine accounts for 12–15% of the solvent accessible surface in globular proteins [105]. It occurs less frequently in protein–protein interfaces, than on the exposed surface. As shown by Conte et al. [105], lysines constitute only 5.4% of interface surface in oligomeric proteins, compared to 14.9% of the total solvent accessible surface. In contrast, for arginine, in spite of its equally high conformational entropy, the percentages are 9.9 and 8.4, while for leucine, they are 10.5 and 3.8%, respectively. Thus, protein–protein interactions discriminate against lysines, and the same is likely to apply to crystal contacts.

Glutamate (Glu) is another potential target: like lysine, it occurs most frequently at the surface, with only 12% buried. The conformational entropy of Glu is estimated to be between 1.55 and 1.75 kcal/mol, depending on the secondary structure context [104]. Protein interfaces discriminate against Glu almost as much they do against Lys. In a selection of oligomeric proteins, the percentage of interface surface attributed to Glu was 4.1%, in contrast to 10.3% of the exposed surface [105].

To test the hypothesis, we studied the behavior of a number of single and multiple Lys → Ala and Glu → Ala mutants of the human protein RhoGDI [101,102]. This cytosolic, highly soluble protein has a large combined Lys/Glu content (~20%) and is difficult to crystallize in its wild-type form. The crystal structure of its immunoglobulin domain obtained by proteolytic cleavage was originally refined using data extending to 2.5 Å from relatively poor quality crystals [106]. We showed that most of the mutations replacing Lys and Glu with Ala critically affected the crystallization properties, yielding a significantly higher rate of success than wild-type protein. To obtain new crystal forms of the protein, it was often necessary to mutate more than one site, but in close proximity. This allowed for the use of single primers for double and triple mutations. Further, in most crystal structures of RhoGDI mutants, the crystal contacts involved the mutated patches. Thus, a causal relationship was observed between the type of surface mutagenesis and the crystallization properties of the protein. In addition, some of the novel crystal forms were shown to diffract to much higher resolution than the wild-type crystal. For example, the double mutant E → A/E → A yielded high-quality crystals which allowed for data collection to 1.2 Å resolution [102].

The approach has been used in our laboratory to crystallize and solve the structures of several new proteins. The RGSL domain of PDZRhoGEF was crystallized using a triple mutant K183A/E185A/E186A, and its structure was determined at 2.2 Å resolution [107,108]. The LcrV antigen from *Yersinia pestis* was also crystallized using a triple mutant, K40A/D41A/K42A, in which, in addition to two lysines, an aspartate separating them was also mutated; the structure was solved and refined at 2.2 Å resolution [115]. The *YkoF* gene product from *Bacillus subtilis* was crystallized using a double mutant K33A/K34A, and the structure was solved and refined at 1.6 Å resolution (unpublished data). In this case, the refined structure revealed that the exposed main chain carbonyls within the mutated patch created a calcium-binding site between two molecules in the asymmetric unit, explaining the dependence of the crystallization on Ca^{2+} ions. The YdeN gene product was crystallized using a double mutant K88A/Q89A, and the structure was solved and refined at 1.8 Å resolution [116]. In this case, the lack of suitable K/E-rich motifs prompted the mutagenesis of a glutamine adjacent to a lysine. Hurley and co-workers [109] used a double mutant (K435A/K436A) of the CUE domain to obtain crystals of a complex with ubiquitin. In all of these cases, the mutated patches appear to play a critical role in formation of crystal contacts. Although the database is limited, two distinct mechanisms can be identified:

(a) the mutated surface patch directly participates in crystal contacts making direct H-bonds to symmetry adjacent molecules; or (b) the patch makes no direct contacts, but modeling of the wild-type sequence suggests that the wild-type amino acids would have caused steric clashes or electrostatic repulsion in the mutant's crystal lattice.

11. Improving crystal quality by site-directed mutagenesis

It is quite typical of crystallographic investigations to focus on the first crystal form of reasonable quality obtained from screens. For soluble proteins, the norm is to use crystals diffracting to better than ≈3.0 Å. It is uncommon to search for better diffracting crystals, unless several crystal forms appear after the first screening. If the only available crystals diffract poorly, an investigator can pursue one of two possibilities: either attempt to improve the diffraction quality of the existing crystals or—if this fails—to make modifications to the protein. The former can frequently be accomplished by special treatment of crystals, such as thermal annealing [110] or dehydration [111,112]. Modifying the protein is very likely to change its physicochemical properties in a way sufficiently significant to alter the crystallization process and to warrant re-initiating the screens from scratch. This has both drawbacks and benefits: the drawback is that the investigator should no longer expect the crystallization observed for the previous samples to be reproducible. On the other hand, new crystal forms may appear with superior diffraction qualities. The modifications are essentially the same as those typically considered if a protein does not crystallize at all, e.g., truncations, changing the construct boundaries or surface mutagenesis. As indicated above, the rational surface mutagenesis strategy can yield crystals of superior quality as compared to wild-type protein [102]. This can be particularly important when an accurate structure of a potential drug target is sought. Munshi et al. [113] used this approach to improve the quality of the crystals of the kinase domain of the insulin-like growth factor receptor-1. In the latter case, a double mutant E1067A/E1069A yielded crystals diffracting to 1.5 Å resolution, in contrast to 2.7 Å resolution reported by the same group for the wild-type crystals.

12. Conclusions

As the easily crystallizable proteins, also known as the 'low-hanging fruit,' are effectively harvested by Structural Genomics Centers and similar initiatives, the projects facing crystallographers today are increasingly difficult at the stage of expression, purification, and crystallization. The use of eukaryotic cells and in vitro translation systems may solve the former problems, but crystallization difficulties will require separate and unique solutions. Rational modification of protein surface and the use of scaffolds to achieve crystallization appear to constitute a promising and exciting alternative to extensive and unpredictable screens. As the number of successful applications of these protocols increases, it will become possible to optimize and make them generally applicable.

Acknowledgments

The research in the author's laboratory referred to in this paper is supported by NIGMS (GM62615). Author thanks Dr. Urszula Derewenda for critical reading of the manuscript.

References

[1] J.D. Bernal, D. Crowfoot, Nature 133 (1934) 794–795.
[2] C.W. Carter Jr., C.W. Carter, J. Biol. Chem. 254 (1979) 12219–12223.
[3] D.V. Goeddel, D.G. Kleid, F. Bolivar, H.L. Heyneker, D.G. Yansura, R. Crea, T. Hirose, A. Kraszewski, K. Itakura, A.D. Riggs, Proc. Natl. Acad. Sci. USA 76 (1979) 106–110.
[4] K. Itakura, T. Hirose, R. Crea, A.D. Riggs, H.L. Heyneker, F. Bolivar, H.W. Boyer, Science 198 (1977) 1056–1063.
[5] R.E. Chance, E.P. Kroeff, J.A. Hoffmann, B.H. Frank, Diabetes Care 4 (1981) 147–154.
[6] D.L. Miller, H.F. Kung, S. Pestka, Science 215 (1982) 689–690.
[7] D.L. Miller, H. Kung, T. Staehelin, S. Pestka, Methods Enzymol. 79 (1981) 3–7.
[8] S. Matsuda, G. Kawano, S. Itoh, Y. Mitsui, Y. Iitaka, J. Biol. Chem. 261 (1986) 16207–16209.
[9] M.G. Grutter, W. Marki, H.P. Walliser, J. Biol. Chem. 260 (1985) 11436–11437.
[10] J.C. Kendrew, R.G. Parrish, J.R. Marrack, E.S. Orlans, Nature 174 (1954) 946–949.
[11] J.W. Campbell, E. Duee, G. Hodgson, W.D. Mercer, D.K. Stammers, P.L. Wendell, H. Muirhead, H.C. Watson, Cold Spring Harb. Symp. Quant. Biol. 36 (1972) 165–170.
[12] C.W. Goulding, L.J. Perry, J. Struct. Biol. 142 (2003) 133–143.
[13] P. Balbas, Mol. Biotechnol. 19 (2001) 251–267.
[14] C.S. Millard, L. Stols, P. Quartey, Y. Kim, I. Dementieva, M.I. Donnelly, Protein Expr. Purif. 29 (2003) 311–320.
[15] P. Neubauer, K. Hofmann, O. Holst, B. Mattiasson, P. Kruschke, Appl. Microbiol. Biotechnol. 36 (1992) 739–744.
[16] A.B. Cubitt, R. Heim, S.R. Adams, A.E. Boyd, L.A. Gross, R.Y. Tsien, Trends Biochem. Sci. 20 (1995) 448–455.
[17] G.S. Waldo, B.M. Standish, J. Berendzen, T.C. Terwilliger, Nat. Biotechnol. 17 (1999) 691–695.
[18] M.J. Weickert, D.H. Doherty, E.A. Best, P.O. Olins, Curr. Opin. Biotechnol. 7 (1996) 494–499.
[19] M.J. Weickert, M. Pagratis, S.R. Curry, R. Blackmore, Appl. Environ. Microbiol. 63 (1997) 4313–4320.
[20] K.G. Han, S.S. Lee, C. Kang, Protein Expr. Purif. 16 (1999) 103–108.
[21] G.X. Du, L.H. Hou, R.B. Guan, Y.G. Tong, H.T. Wang, World J. Gastroenterol. 8 (2002) 1088–1093.
[22] J.G. Thomas, A. Ayling, F. Baneyx, Appl. Biochem. Biotechnol. 66 (1997) 197–238.

[23] K. Nishihara, M. Kanemori, H. Yanagi, T. Yura, Appl. Environ. Microbiol. 66 (2000) 884–889.
[24] U. Brinkmann, R.E. Mattes, P. Buckel, Gene 85 (1989) 109–114.
[25] C.P. Carstens, Methods Mol. Biol. 205 (2003) 225–233.
[26] T. Kleber-Janke, W.M. Becker, Protein Expr. Purif. 19 (2000) 419–424.
[27] H. Apeler, U. Gottschalk, D. Guntermann, J. Hansen, J. Massen, E. Schmidt, K.H. Schneider, M. Schneidereit, H. Rubsamen-Waigmann, Eur. J. Biochem. 247 (1997) 890–895.
[28] G.E. Dale, C. Broger, H. Langen, A. D'Arcy, D. Stuber, Protein Eng. 7 (1994) 933–939.
[29] R. Wetzel, L.J. Perry, C. Veilleux, Biotechnology (NY) 9 (1991) 731–737.
[30] Y. Endo, T. Sawasaki, Biotechnol. Adv. 21 (2003) 695–713.
[31] M. Uhlen, G. Forsberg, T. Moks, M. Hartmanis, B. Nilsson, Curr. Opin. Biotechnol. 3 (1992) 363–369.
[32] M.C. Smith, T.C. Furman, T.D. Ingolia, C. Pidgeon, J. Biol. Chem. 263 (1988) 7211–7215.
[33] D.B. Smith, K.S. Johnson, Gene 67 (1988) 31–40.
[34] C. di Guan, P. Li, P.D. Riggs, H. Inouye, Gene 67 (1988) 21–30.
[35] E.R. LaVallie, E.A. DiBlasio, S. Kovacic, K.L. Grant, P.F. Schendel, J.M. McCoy, Biotechnology (NY) 11 (1993) 187–193.
[36] B. Nilsson, T. Moks, B. Jansson, L. Abrahmsen, A. Elmblad, E. Holmgren, C. Henrichson, T.A. Jones, M. Uhlen, Protein Eng. 1 (1987) 107–113.
[37] G.D. Davis, C. Elisee, D.M. Newham, R.G. Harrison, Biotechnol. Bioeng. 65 (1999) 382–388.
[38] J.R. Huth, C.A. Bewley, B.M. Jackson, A.G. Hinnebusch, G.M. Clore, A.M. Gronenborn, Protein Sci. 6 (1997) 2359–2364.
[39] H.T. Ding, H. Ren, Q. Chen, G. Fang, L.F. Li, R. Li, Z. Wang, X.Y. Jia, Y.H. Liang, M.H. Hu, Y. Li, J.C. Luo, X.C. Gu, X.D. Su, M. Luo, S.Y. Lu, Acta Crystallogr. D Biol. Crystallogr. 58 (2002) 2102–2108.
[40] P. Sheffield, S. Garrard, Z. Derewenda, Protein Expr. Purif. 15 (1999) 34–39.
[41] M. Hammarstrom, N. Hellgren, S. van Den Berg, H. Berglund, T. Hard, Protein Sci. 11 (2002) 313–321.
[42] R.B. Kapust, D.S. Waugh, Protein Sci. 8 (1999) 1668–1674.
[44] K.M. Routzahn, D.S. Waugh, J. Struct. Funct. Genom. 2 (2002) 83–92.
[45] J.D. Fox, R.B. Kapust, D.S. Waugh, Protein Sci. 10 (2001) 622–630.
[46] A. Alexandrov, K. Dutta, S.M. Pascal, Biotechniques 30 (2001) 1194–1198.
[47] J. Phan, A. Zdanov, A.G. Evdokimov, J.E. Tropea, H.K. Peters III, R.B. Kapust, M. Li, A. Wlodawer, D.S. Waugh, J. Biol. Chem. 277 (2002) 50564–50572.
[48] R.B. Kapust, J. Tozser, T.D. Copeland, D.S. Waugh, Biochem. Biophys. Res. Commun. 294 (2002) 949–955.
[49] R.B. Kapust, J. Tozser, J.D. Fox, D.E. Anderson, S. Cherry, T.D. Copeland, D.S. Waugh, Protein Eng. 14 (2001) 993–1000.
[50] R.B. Kapust, D.S. Waugh, Protein Expr. Purif. 19 (2000) 312–318.
[51] A.P. Doskeland, A. Martinez, P.M. Knappskog, T. Flatmark, Biochem. J. 313 (Pt. 2) (1996) 409–414.
[52] R.R. Porter, Science 180 (1973) 713–716.
[53] J.P. Noel, H.E. Hamm, P.B. Sigler, Nature (London) 366 (1993) 654–663.
[54] J. Sondek, A. Bohm, D.G. Lambright, H.E. Hamm, P.B. Sigler, Nature (London) 379 (1996) 369–374.
[55] C. Mazza, A. Segref, I.W. Mattaj, S. Cusack, Acta Crystallogr. D Biol. Crystallogr. 58 (2002) 2194–2197.
[56] O.V. Galzitskaya, B.S. Melnik, Protein Sci. 12 (2003) 696–701.
[57] P. Adamson, C.J. Marshall, A. Hall, P.A. Tilbrook, J. Biol. Chem. 267 (1992) 20033–20038.
[58] E.F. Pai, W. Kabsch, U. Krengel, K.C. Holmes, J. John, A. Wittinghofer, Nature 341 (1989) 209–214.
[59] M. Hirshberg, R. Stockley, G. Dodson, R. Webb, Nat. Struct. Biol. 4 (1997) 147–151.
[60] Y. Wei, Y. Zhang, U. Derewenda, X. Liu, W. Minor, R.K. Nakomoto, A.V. Somlyo, A.P. Somlyo, Z.S. Derewenda, Nat. Struct. Biol. 4 (1997) 699–702.
[61] G.E. Dale, D. Kostrewa, B. Gsell, M. Stieger, A. D'Arcy, Acta Crystallogr. D Biol. Crystallogr. 55 (1999) 1626–1629.
[62] P.D. Kwong, R. Wyatt, J. Robinson, R.W. Sweet, J. Sodroski, W.A. Hendrickson, Nature 393 (1998) 648–659.
[63] P.D. Kwong, R. Wyatt, E. Desjardins, J. Robinson, J.S. Culp, B.D. Hellmig, R.W. Sweet, J. Sodroski, W.A. Hendrickson, J. Biol. Chem. 274 (1999) 4115–4123.
[64] F. Dyda, A.B. Hickman, T.M. Jenkins, A. Engelman, R. Craigie, D.R. Davies, Science 266 (1994) 1981–1986.
[65] T.M. Jenkins, A.B. Hickman, F. Dyda, R. Ghirlando, D.R. Davies, R. Craigie, Proc. Natl. Acad. Sci. USA 92 (1995) 6057–6061.
[66] D. Das, M.M. Georgiadis, Protein Sci. 10 (2001) 1936–1941.
[67] E. Bianchi, S. Venturini, A. Pessi, A. Tramontano, M. Sollazzo, J. Mol. Biol. 236 (1994) 649–659.
[68] J.C. Smith, R.B. Derbyshire, E. Cook, L. Dunthorne, J. Viney, S.J. Brewer, H.M. Sassenfeld, L.D. Bell, Gene 32 (1984) 321–327.
[69] S.J. Brewer, B.L. Haymore, T.P. Hopp, H.M. Sassenfeld, Bioprocess Technol. 12 (1991) 239–266.
[70] F. Nachon, Y. Nicolet, N. Viguie, P. Masson, J.C. Fontecilla-Camps, O. Lockridge, Eur. J. Biochem. 269 (2002) 630–637.
[71] H. Zhan, B. Liu, S.W. Reid, K.H. Aoki, C. Li, R.S. Syed, C. Karkaria, G. Koe, K. Sitney, K. Hayenga, F. Mistry, L. Savel, M. Dreyer, B.A. Katz, J. Schreurs, D.J. Matthews, J.C. Cheetham, J. Egrie, L.B. Giebel, R.M. Stroud, Protein Eng. 12 (1999) 505–513.
[72] S.J. Davis, E.A. Davies, A.N. Barclay, S. Daenke, D.L. Bodian, E.Y. Jones, D.I. Stuart, T.D. Butters, R.A. Dwek, P.A. van der Merwe, J. Biol. Chem. 270 (1995) 369–375.
[73] S.J. Davis, M.J. Puklavec, D.A. Ashford, K. Harlos, E.Y. Jones, D.I. Stuart, A.F. Williams, Protein Eng. 6 (1993) 229–232.
[74] S.J. Davis, S. Ikemizu, A.V. Collins, J.A. Fennelly, K. Harlos, E.Y. Jones, D.I. Stuart, Acta Crystallogr. D. Biol. Crystallogr. 57 (2001) 605–608.
[75] C. Stover, M.P. Mayhew, M.J. Holden, A. Howard, D.T. Gallagher, J. Struct. Biol. 129 (2000) 96–99.
[76] N. Asherie, J. Pande, A. Pande, J.A. Zarutskie, J. Lomakin, A. Lomakin, O. Ogun, L.J. Stern, J. King, G.B. Benedek, J. Mol. Biol. 314 (2001) 663–669.
[77] B. Kobe, R.J. Center, B.E. Kemp, P. Poumbourios, Proc. Natl. Acad. Sci. USA 96 (1999) 4319–4324.
[78] Y. Liu, A. Manna, R. Li, W.E. Martin, R.C. Murphy, A.L. Cheung, G. Zhang, Proc. Natl. Acad. Sci. USA 98 (2001) 6877–6882.
[79] A. Ke, C. Wolberger, Protein Sci. 12 (2003) 306–312.
[80] D.R. Smyth, M.K. Mrozkiewicz, W.J. McGrath, P. Listwan, B. Kobe, Protein Sci. 12 (2003) 1313–1322.
[81] J.A. Chao, G.S. Prasad, S.A. White, C.D. Stout, J.R. Williamson, J. Mol. Biol. 326 (2003) 999–1004.
[82] M. Kuge, Y. Fujii, T. Shimizu, F. Hirose, A. Matsukage, T. Hakoshima, Protein Sci. 6 (1997) 1783–1786.
[83] V.S. Stoll, A.V. Manohar, W. Gillon, E.L. MacFarlane, R.C. Hynes, E.F. Pai, Protein Sci. 7 (1998) 1147–1155.
[84] E.A. Stura, M. Graille, N.G. Housden, M.G. Gore, Acta Crystallogr. D. Biol. Crystallogr. 58 (2002) 1744–1748.
[85] E.A. Stura, M.J. Taussig, B.J. Sutton, S. Duquerroy, S. Bressanelli, A.C. Minson, F.A. Rey, Acta Crystallogr. D. Biol. Crystallogr. 58 (2002) 1715–1721.
[86] D.J. Derbyshire, B.P. Basu, T. Date, K. Iwabuchi, A.J. Doherty, Acta Crystallogr. D Biol. Crystallogr. 58 (2002) 1826–1829.
[87] L. Renault, M. Hanzal-Bayer, R.C. Hillig, Acta Crystallogr. D Biol. Crystallogr. 57 (2001) 1167–1170.
[88] T. Nakai, N. Nakagawa, N. Maoka, R. Masui, S. Kuramitsu, N. Kamiya, Acta Crystallogr. D Biol. Crystallogr. 59 (2003) 554–557.

[89] D. Neumann, A. Woods, D. Carling, T. Wallimann, U. Schlattner, Protein Expr. Purif. 30 (2003) 230–237.
[90] P.J. Sheffield, T.W. McMullen, J. Li, Y.S. Ho, S.M. Garrard, U. Derewenda, Z.S. Derewenda, Protein Eng. 14 (2001) 513–519.
[91] H. Wang, S. Chong, Proc. Natl. Acad. Sci. USA 100 (2003) 478–483.
[92] H.H. McElroy, G.W. Sisson, W.E. Schottlin, R.M. Aust, J.E. Villafranca, J. Crystal Growth 122 (1992) 265–272.
[93] A. D'Arcy, M. Stihle, D. Kostrewa, G. Dale, Acta Crystallogr. D Biol. Crystallogr. 55 (1999) 1623–1625.
[94] B. Lorber, A. Theobald-Dietrich, C. Charron, C. Sauter, J.D. Ng, D.W. Zhu, R. Giege, Acta Crystallogr. D Biol. Crystallogr. 58 (2002) 1674–1680.
[95] K. Braig, Z. Otwinowski, R. Hegde, D.C. Boisvert, A. Joachimiak, A.L. Horwich, P.B. Sigler, Nature 371 (1994) 578–586.
[96] A. Horwich, Nat. Struct. Biol. 7 (2000) 269–270.
[97] D.M. Lawson, P.J. Artymiuk, S.J. Yewdall, J.M. Smith, J.C. Livingstone, A. Treffry, A. Luzzago, S. Levi, P. Arosio, G. Cesareni, et al., Nature 349 (1991) 541–544.
[98] C. Wingren, A.B. Edmundson, C.A. Borrebaeck, Protein Eng. 16 (2003) 255–264.
[99] P.G. Vekilov, A.R. Feeling-Taylor, S.T. Yau, D. Petsev, Acta Crystallogr. D Biol. Crystallogr. 58 (2002) 1611–1616.
[100] P.G. Vekilov, Methods Enzymol. 368 (2003) 84–105.
[101] K.L. Longenecker, S.M. Garrard, P.J. Sheffield, Z.S. Derewenda, Acta Crystallogr. D Biol. Crystallogr. 57 (2001) 679–688.
[102] A. Mateja, Y. Devedjiev, D. Krowarsch, K. Longenecker, Z. Dauter, J. Otlewski, Z.S. Derewenda, Acta Crystallogr. D Biol. Crystallogr. 58 (2002) 1983–1991.
[103] F. Baud, S. Karlin, Proc. Natl. Acad. Sci. USA 96 (1999) 12494–12499.
[104] F. Avbelj, L. Fele, J. Mol. Biol. 279 (1998) 665–684.
[105] L.L. Conte, C. Chothia, J. Janin, J. Mol. Biol. 285 (1999) 2177–2198.
[106] N.H. Keep, M. Barnes, I. Barsukov, R. Badii, L.Y. Lian, A.W. Segal, P.C.E. Moody, G.C.K. Roberts, Structure 5 (1997) 623–633.
[107] S.M. Garrard, K.L. Longenecker, M.E. Lewis, P.J. Sheffield, Z.S. Derewenda, Protein Expr. Purif. 21 (2001) 412–416.
[108] K.L. Longenecker, M.E. Lewis, H. Chikumi, J.S. Gutkind, Z.S. Derewenda, Structure (Camb.) 9 (2001) 559–569.
[109] G. Prag, S. Misra, E.A. Jones, R. Ghirlando, B.A. Davies, B.F. Horazdovsky, J.H. Hurley, Cell 113 (2003) 609–620.
[110] S. Kriminski, C.L. Caylor, M.C. Nonato, K.D. Finkelstein, R.E. Thorne, Acta Crystallogr. D Biol. Crystallogr. 58 (2002) 459–471.
[111] R.M. Esnouf, J. Ren, E.F. Garman, D.O. Somers, C.K. Ross, E.Y. Jones, D.K. Stammers, D.I. Stuart, Acta Crystallogr. D Biol. Crystallogr. 54 (Pt. 5) (1998) 938–953.
[112] T.H. Tahirov, H. Oki, T. Tsukihara, K. Ogasahara, K. Yutani, C.P. Libeu, Y. Izu, S. Tsunasawa, I. Kato, J. Struct. Biol. 121 (1998) 68–72.
[113] S. Munshi, D.L. Hall, M. Kornienko, P.L. Darke, L.C. Kuo, Acta Crystallogr. D Biol. Crystallogr. 59 (2003) 1725–1730.
[114] Z.S. Derewenda, Structure 12 (2004) 529–535.
[115] U. Derewenda, A. Mateja, Y. Devedjiev, K.M. Routzahn, A.G. Evdokimov, Z.S. Derewenda, D.S. Waugh, Structure 12 (2004) 301–306.
[116] I. Janda, Y. Devedjiev, D. Cooper, M. Chruszcz, U. Derewenda, A. Gabrys, W. Minor, A. Joachimiak, Z.S. Derewanda, Acta Crystallogr. D Biol. Crystallogr. 60 (2004) 1101–1107.