electronic reprint

Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449
Editors: **E. N. Baker and Z. Dauter**

A pivotal role for reductive methylation in the *de novo* crystallization of a ternary complex composed of *Yersinia pestis* virulence factors YopN, SycN and YscB

Florian D. Schubot and David S. Waugh

Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

A pivotal role for reductive methylation in the *de novo* crystallization of a ternary complex composed of *Yersinia pestis* virulence factors YopN, SycN and YscB

Florian D. Schubot and David S. Waugh*

Protein Engineering Section, Macromolecular Crystallography Laboratory, Center for Cancer Research, National Cancer Institute at Frederick, PO Box B, Frederick, MD 21702-1201, USA

Correspondence e-mail: waughd@ncifcrf.gov

Received 30 August 2004
Accepted 15 September 2004

Structural studies of a ternary complex composed of the *Yersina pestis* virulence factors YopN, SycN and YscB were initially hampered by poor solubility of the individual proteins. Co-expression of all three proteins in *Escherichia coli* yielded a well behaved complex, but this sample proved to be recalcitrant to crystallization. As crystallization efforts remained fruitless, even after the proteolysis-guided engineering of a truncated YopN polypeptide, reductive methylation of lysine residues was employed to alter the surface properties of the complex. The methylated complex yielded crystals that diffracted X-rays to a maximal resolution of 1.8 Å. The potential utility of reductive methylation as a remedial strategy for high-throughput structural biology was further underscored by the successful modification of a selenomethionine-substituted sample.

1. Introduction

It is not uncommon to encounter a well expressed, highly soluble and stably folded protein that is recalcitrant to crystallization. Indeed, the latest data from large-scale structural genomics initiatives indicate that fewer than half of the soluble proteins that can be purified to homogeneity actually yield crystals (O'Toole *et al.*, 2004). In such cases, the only recourse may be to alter the sequence of the protein itself. However, beyond the proteolysis-guided generation of truncation mutants, there are no general rules that enable one to predict which truncation, amino-acid substitution or ortholog will crytallize. For this reason, alternative methods that could be used to promote the growth of protein crystals would be of considerable utility.

Surface-entropy-reduction mutagenesis, a strategy pioneered by Zygmunt Derewenda and coworkers (Derewenda, 2004; Longenecker *et al.*, 2001; Mateja *et al.*, 2002), has shown some promise in this regard. In this method, linear clusters of presumably surface-located amino-acid side chains with high conformational entropy (*e.g.* Lys and Glu) are replaced by methyl groups (Ala) in an effort to create new epitopes that will facilitate crystallization. Several proteins have been crystallized in this manner, including *Yersinia pestis* LcrV (Derewenda *et al.*, 2004), suggesting that the method may be of general utility. Yet, because it is impossible to predict which cluster mutant(s) will crystallize, the probability of a successful outcome is proportional to the number of mutants that are screened. Consequently, surface-entropy-reduction mutagenesis can be a very labor-intensive undertaking.

An alternative strategy, which has received relatively little attention thus far, is to modify the wild-type protein by reductive alkylation with formaldehyde and dimethylamine–

borane complex (Rayment *et al.*, 1993; Rypniewski *et al.*, 1993). The net result of this reaction is the dimethylation of all accessible lysine side chains and the N-terminal amino group. Although this does not change the intrinsic charge of a protein, it may alter its isoelectric point slightly. The rationale behind this strategy is that dimethylation of lysine side chains will reduce their interaction with solvent, thereby causing them to adopt more ordered conformations that may facilitate crystallization.

Here, we present a case study in which reductive methylation played a critical role in the crystallization of a ternary complex composed of the *Yersinia pestis* virulence factors YopN, SycN and YscB. We also demonstrate that selenomethionine- (SeMet) substituted proteins can be safely derivatized by this method, which is particularly important in a high-throughput setting. Because it can be carried out in a very short period of time, we suggest that reductive methylation may be worth considering as a standard remedial strategy for the *de novo* crystallization of proteins and even multimeric complexes prior to pursuing more involved approaches such as surface-entropy-reduction mutagenesis.

2. Materials and methods

2.1. Expression and purification of YopN–SycN–YscB-His$_6$

A multicistronic expression vector encoding a truncated form of YopN (residues 32–277), SycN and YscB-His$_6$ was assembled by Gateway multisite recombinational cloning. Ribosome-binding sites and the appropriate *att* recombination sites were added to each gene by PCR. *yopN*$^{32–277}$, *sycN* and *yscB*-His$_6$ were inserted into pDONR208A, pDONR213B and pDONR214C, respectively (Invitrogen, Carlsbad, CA, USA). The genes were sequence-verified and subsequently recombined into pDEST42 (Invitrogen) to create the multicistronic expression vector pFS1468.

Single ampicillin-resistant colonies of *Escherichia coli* BL21(DE3) CodonPlus-RIL cells (Stratagene, La Jolla, CA, USA) containing pFS1468 were used to inoculate 100 ml Luria broth supplemented with 2 g l^{-1} glucose, 100 µg ml^{-1} ampicillin and 30 µg ml^{-1} chloramphenicol. The culture was grown with shaking (225 rev min^{-1}) to saturation overnight at 310 K and then diluted 66-fold into 6 l of fresh medium. When the cells reached early log phase ($OD_{600\,nm}$ = 0.5), the temperature was reduced to 303 K and isopropyl-β-D-thiogalactopyranoside (IPTG) was added to a final concentration of 1 m*M*. 4 h later, the cells were recovered by centrifugation at 5000*g* for 15 min.

The cell paste was resuspended in 400 ml of 50 m*M* HEPES pH 7.6, 150 m*M* NaCl, 25 m*M* imidazole (buffer *A*) along with four tablets of Complete EDTA-free protease-inhibitor cocktail (Roche Molecular Biochemicals, Indianapolis, IN, USA). The cells were lysed with an APV Gaulin Model G1000 homogenizer at 69 MPa and centrifuged at 30 000*g* for 30 min at 277 K. The supernatant was filtered through a 0.45 µm polyethersulfone membrane and applied onto a 25 ml Ni–NTA Superflow affinity column (Qiagen, Valencia, CA, USA) equilibrated with buffer *A*. The column was washed with five volumes of buffer *A* and then eluted with a linear gradient from 25 to 250 m*M* imidazole. The peak fractions were pooled, concentrated using an Amicon YM10 membrane (Millipore, Billerica, MA, USA) and then applied onto a 26/60 HiLoad Superdex 75 prep-grade column (Amersham Biosciences, Piscataway, NJ, USA) equilibrated with 25 m*M* Tris–HCl pH 7.4, 150 m*M* NaCl (buffer *B*). The peak fractions were pooled and concentrated to 26 mg ml^{-1}. Aliquots were flash-frozen with liquid nitrogen and stored at 193 K until use. The final product was judged to be >95% pure by SDS–PAGE (Fig. 1). The molecular weights were confirmed by electrospray mass spectrometry (ESMS), which revealed a 10 Da discrepancy in the molecular weight of SycN. This discrepancy was later explained by the presence of a spontaneous mutation in the expression vector that altered the final residue of SycN from proline to serine. The SeMet-substituted YopN$^{32–277}$–SycN–YscB-His$_6$ complex was produced as described in Doublié (1997) and handled in the same manner as the native protein complex; ES-MS indicated that selenomethionine substitution was >99%.

2.2. Reductive methylation of the YopN$^{32-277}$–SycN–YscB-His$_6$ complex

The reductive-methylation reaction of the ternary complex closely followed the protocol published by Ivan Rayment (Rayment, 1997). 2 µl of the YopN$^{32–277}$–SycN–YscB-His$_6$ complex in buffer *B* (26 mg ml^{-1}) was dialyzed into 50 m*M* HEPES pH 7.6, 150 m*M* NaCl (buffer *C*). 1 *M* solutions of dimethylamine–borane complex (DMAB; Aldrich, St Louis, MO, USA) and methanol-free formaldehyde (Polysciences Inc., Warrington, PA, USA) were prepared in the dialysis buffer. The protein sample was placed in a 15 ml Falcon tube and 3 ml of dialysis buffer was added to bring the protein

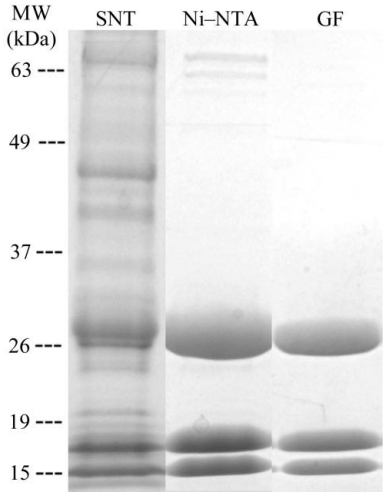

Figure 1
Composite picture of lanes from SDS–polyacrylamide gels illustrating the two-step purification process for the YopN$^{32–277}$–SycN–YscB-His$_6$ complex. Shown are the lanes for the initial supernatant (SNT), the peak fractions eluted from the Ni–NTA affinity column (Ni–NTA) and the final sample after gel filtration (GF).

concentration to approximately 10 mg ml^{-1}. The tube was wrapped in aluminium foil and placed on a shaker at 277 K. 100 µl of DMAB solution and 200 µl of 1 *M* formaldehyde were then added to the sample with very slow shaking. After letting the reaction shake for 2 h, this last step was repeated. After another 2 h period, a final 50 µl of DMAB was added and the reaction was allowed to shake slowly for 18 h at 277 K. The reaction was quenched by the addition of 500 µl 1 *M* $(NH_4)_2SO_4$ and the progress of the reaction was then checked by ES-MS. The modified sample was subsequently loaded onto an S-100 gel-filtration column that had been pre-equilibrated with buffer *B*. The peak fractions were pooled and concentrated to about 10 mg ml^{-1}.

2.3. Crystallization

Crystallization screening of the methylated sample was conducted in Vapor Batch Plates (Hampton Research, Aliso Viejo, CA, USA) using the modified microbatch technique (Chayen, 1997) in which the drop is covered with 2 ml of a 50:50 mixture of paraffin and silicone oil (Hampton Research). The sample was initially screened with commercially available crystallization matrices and crystals were obtained from condition G12 of the Index HT Crystallization Matrix (Hampton Research). Optimized conditions, consisting of 100 m*M* HEPES pH 6.8, 400 m*M* $MgCl_2$, 22% PEG 3350 and the protein complex at 10 mg ml^{-1} in 25 m*M* Tris–HCl pH 7.4, 150 m*M* NaCl, yielded tetragonal crystals at 291 K that were used for structure solution (Fig. 2). These crystals diffracted X-rays to a maximal resolution of 1.8 Å. The methylation and crystallization procedures for the SeMet-substituted YopN$^{32–277}$–SycN–YscB-His$_6$ complex were identical.

2.4. X-ray data collection

After cryosoaking in 100 m*M* HEPES pH 6.8, 400 m*M* $MgCl_2$, 22% PEG 3350 and 5% glycerol, the crystals were mounted in a loop and flash-frozen with liquid nitrogen. The data sets obtained from crystals of the native and SeMet-labeled complex that were used to solve the structure were collected at the SER-CAT beamline 22-ID (Advanced Photon Source, Argonne National Laboratory) equipped with a MAR CCD 225 detector. Data processing was carried out with the *HKL*2000 program suite (Otwinowski & Minor, 1997). Details of the data-collection and processing statistics are presented in Table 1.

Table 1
Data-collection statistics.

Values in parentheses relate to the highest resolution shell.

Data set	Native	SeMet-substituted
Molecules per AU	1	1
Wavelength (Å)	1.0668	0.9797
Space group	$P4_1$	$P4_1$
Unit-cell parameters (Å)	$a = b = 60.7$, $c = 140.4$	$a = b = 60.9$, $c = 140.9$
Resolution (Å)	50–1.84	20–2.07
Total reflections	915853	530445
Unique reflections	43586	29364
Redundancy	5.0 (4.3)	6.5 (2)
Completeness (%)	99.2 (98.4)	92.4 (49.3)
Average $I/\sigma(I)$	26.8 (4.13)	26.5 (2.0)
$R_{\rm merge}$† (%)	4.9 (31.6)	7.2 (38.2)
Twin fraction‡ (%)	17	32

† $R_{\rm merge} = \sum|I| - \langle I\rangle/\sum I$, where I is the observed intensity and $\langle I\rangle$ is the average intensity obtained from multiple observations of symmetry-related reflections after rejections. ‡ Estimated from the Merohedral Twinning Server (http://www.doe-mbi.ucla.edu/Services/Twinning/).

3. Results

As part of an ongoing structural genomics project targeting proteins involved in type III secretion in *Y. pestis*, the causative agent of plague, an effort was made to overproduce 50 different polypeptides in *E. coli*. However, initial attempts to overproduce and purify YopN (293 amino acids), SycN (123 amino acids) or YscB (137 amino acids) individually were hampered by their poor solubility (data not shown). SycN and YscB are known to form a heterodimeric secretion chaperone that binds to YopN (Day & Plano, 1998; Jackson *et al.*, 1998). Reasoning that the ternary complex might be more soluble than the individual polypeptides, we next assembled a multicistronic expression vector to express all three proteins simultaneously. A His_6 tag was added to the C-terminus of YscB to facilitate the purification of the ternary complex.

When expressed simultaneously, all three polypeptides were produced at a high level and in a soluble form in *E. coli*. The ternary complex was readily purified to homogeneity by immobilized metal-affinity chromatography and gel-filtration chromatography. However, although co-expression and co-purification of the YopN–SycN–YscB-His$_6$ complex alleviated the solubility problem, extensive efforts to crystallize the sample were fruitless.

We speculated that the inability to obtain crystals of the ternary complex might be a consequence of the presence of disordered regions in one or more of the polypeptides. Therefore, the complex was subjected to limited proteolysis with thermolysin in an effort to define its domain boundaries. A stable digestion product of YopN lacking 31 N-terminal residues and 16 C-terminal residues was unambiguously identified by N-terminal amino-acid sequencing and ES-MS (data not shown). When co-expressed with SycN and YscB-His$_6$, this truncated YopN polypeptide was also capable of forming a stable 57 kDa complex that could be purified to homogeneity. However, the newly designed complex did not yield crystals either, even though dynamic light-scattering experiments indicated that it was well behaved and existed in solution as monodisperse heterotrimeric species.

Hoping to further improve the 'crystallizability' of the sample, we next subjected it to reductive (di)-methylation. The methylation of accessible lysine side chains by chemical modification is intended to alter the surface properties of a target sample and thereby influence its behavior during crystallization screening. Although in the past this technique has primarily been employed to improve the quality of crystals with inferior diffraction properties, it was originally employed as a *de novo* crystallization tool for determining the structure

Table 2
Molecular masses of YopN$^{32–277}$–SycN–YscB-His$_6$ complex components before and after reductive methylation.

	YopN$^{32–277}$	YscB-His$_6$	SycN
No. lysines	15	6	2
Native sample			
M_r before methylation (predicted)	27276 (27279)	16231 (16232)	13467 (13468)
M_r after methylation (predicted†)	27726 (27727)	16400 (16428)	13551 (13552)
No. methyl groups based on ΔM_r‡	32.1	12.1	6.0
SeMet-substituted sample			
M_r before methylation (predicted§)	27465 (27466)	16419 (16420)	13560 (13561)
M_r after methylation (predicted†)	27914 (27913)	16587 (16615)	13645 (13644)
No. methyl groups based on ΔM_r‡	32.1	12.0	6.0

† Assuming dimethylation of all lysines and the N-terminus. ‡ ΔM_r = M_r after methylation − M_r before methylation. § Assuming full substitution of all methionine residues with selenomethionine.

Figure 2
Crystals (∼50 × 50 × 70 µm) of the methylated YopN$^{32–277}$–SycN–YscB-His$_6$ complex.

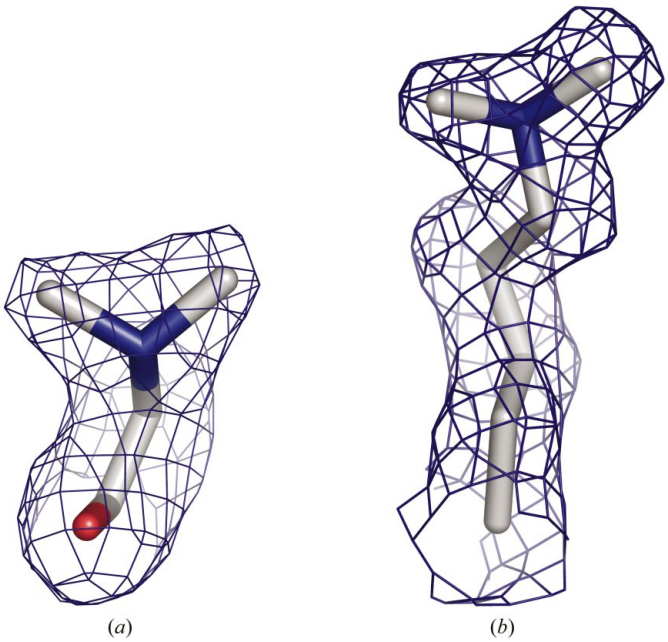

Figure 3
$2F_o - F_c$ Fourier map of electron density contoured at 1.5σ for (*a*) Gly32, the methylated amino-terminus of YopN, and (*b*) YopN residue Lys237. These two residues represent the only places where the added methyl groups were clearly visible. This figure was generated by *PyMOL* (DeLano, 2001).

of myosin subfragment 1 (Rayment *et al.*, 1993). In at least two cases, the reported crystallization conditions for methylated and unmodified samples differed dramatically (Kurinov *et al.*, 2000; Rypniewski *et al.*, 1993), which further supports the notion that one could accomplish a substantial change in the behavior of the target samples.

Reductive methylation of the YopN$^{32–277}$–SycN–YscB-His$_6$ complex was monitored by LC/ES-MS to gauge the extent of the reaction. The data suggested that the standard protocol developed by Ivan Rayment and coworkers (Rayment, 1997) worked very well for this particular complex. A single species was detected for each protein and the shifts in the molecular masses could be accurately explained by the dimethylation of virtually all lysine residues along with the three amino-termini (Table 2). YscB-His$_6$ was the lone exception; the measured mass difference of 169 corresponded to only 12 added methyl groups rather than the predicted 14. After the structure had been solved, the unmodified lysine residue could be readily identified as Lys77, which is involved in hydrogen-bonding interactions with Asp39 of SycN and Glu51 of YopN (Fig. 3). After a polishing step involving gel filtration, which also served to confirm the structural integrity of the methylated complex, the sample was screened for crystallization conditions. Crystals were obtained as described above.

We found no reported examples of reductive methylation of SeMet-substituted proteins. We did not anticipate any problems, however, since the reducing environment and proven specificity of the reaction should leave the SeMet residues unaffected. When the SeMet-substituted protein complex was subjected to reductive methylation, the mass shifts detected by LC/ES-MS corresponded precisely to those observed after the original YopN$^{32–277}$–SycN–YscB-His$_6$ complex had been methylated, confirming not only the reproducibility of the reaction but also that it can be safely employed on SeMet-substituted proteins. Crystals of the methylated SeMet protein complex were obtained under the same conditions as used for the original complex. The details of the structure-solution process and the structure of the ternary complex will be published elsewhere.

4. Discussion

The reasons for the often improved crystallization properties of methylated proteins are not entirely clear. However, lysine side chains (as well as those of glutamic acid and arginine) are notoriously flexible and their electron density often cannot be seen in crystal structures even at a resolution of 2 Å or better. Disordered side chains may have a deleterious effect on the formation of ordered crystal lattices. This notion is supported by the underrepresentation of these residues among amino acids that are actually involved in the formation of inter-

molecular crystal contacts (Lo Conte *et al.*, 1999) and by the observation that replacing linear clusters of lysines and/or glutamate residues with alanines promotes the crystallization of some proteins, a strategy that has been dubbed 'surface-entropy-reduction mutagenesis' (Derewenda, 2004; Longenecker *et al.*, 2001; Mateja *et al.*, 2002).

Obviously, the local changes caused by reductive methylation of lysine side chains are not as dramatic as those arising from surface-entropy-reduction mutagenesis, but the global effects of methylation on the surface properties of the protein are significant and physically measurable. The modified samples usually display reduced solubility and, as also reported by another group (Kurinov *et al.*, 2000), they tend to exhibit a shift towards a larger molecular weight in their elution profiles from gel-filtration columns (data not shown). At some level, the impact of reductive methylation and surface-entropy-reduction mutagenesis may be similar. Although they do not neutralize the positive charge of the lysine side chain, the methyl groups do increase its hydrophobicity, which should reduce its interactions with the aqueous environment. The net result may be a decrease in thermal motion and better ordered side chains (*i.e.* a reduction of surface entropy).

Following this line of reasoning, one might expect to observe one or more dimethylated lysine residues at points of contact between YopN$^{32–277}$–SycN–YscB-His$_6$ complexes in the crystal lattice. After all, in every case in which surface-entropy-reduction mutagenesis has yielded crystals, the mutated epitopes have been observed to engage in intermolecular crystal contacts (Derewenda *et al.*, 2004; Janda *et al.*, 2004; Longenecker *et al.*, 2001). Of the 21 lysines observed in the structure, 16 have reasonably well ordered side chains, while the remaining five are completely disordered. Although the structure was solved at 1.8 Å resolution, there is only one instance (YopN residue Lys237) in which the added methyl groups are clearly visible in the $2F_o - F_c$ Fourier map (Fig. 3). This residue, along with Lys128 and Lys264, is directly involved in the formation of a crystal contact and six other lysines are located within a side-chain rotation of a crystal contact. While these observations are in line with the notion that reductive methylation and surface-entropy-reduction mutagenesis elicit similar effects, with one acting globally and the other acting locally, more structures of methylated proteins will be needed to provide reasonable statistical evidence for distinct distributions of lattice contacts and lysine residues in the context of these lattice contacts.

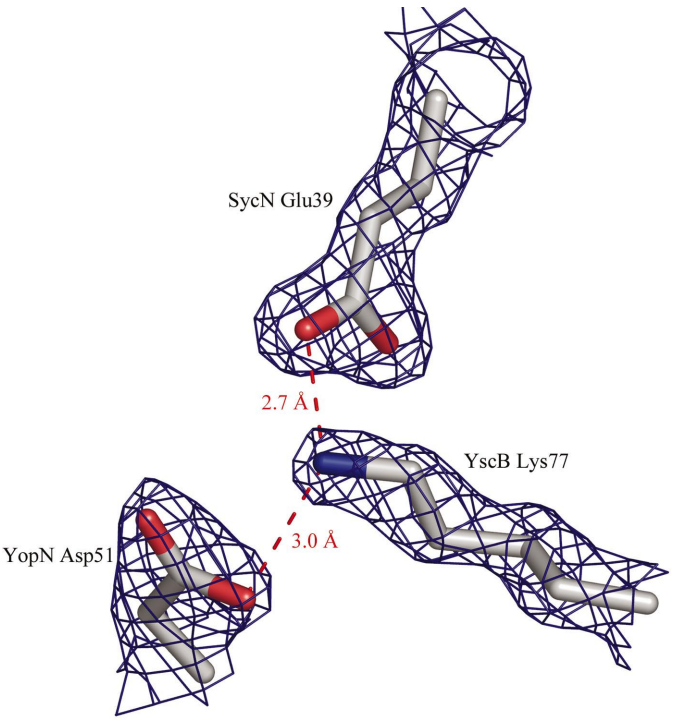

Figure 4
Electron density of a σ-weighted $2F_o - F_c$ Fourier map contoured at 1.5σ for Lys77 of YscB, the only unmethylated lysine residue in the ternary complex, which interacts with acidic residues from both YopN and SycN. This observation is important because it demonstrates that the experimental conditions of reductive methylation leave such intermolecular contact sites unaffected.

In conclusion, as evidenced by the protection from methylation of YscB residue Lys77, which is involved in crucial intermolecular interactions with Gly39 of SycN and Asp51 of YopN (Fig. 4), the successful attempt to derivatize the YopN$^{32–277}$–SycN–YscB-His$_6$ complex by reductive methylation suggests that the reaction conditions are sufficiently mild so as not to disrupt the structural integrity of protein–protein complexes. Furthermore, the ES-MS data presented here confirm that SeMet-substituted proteins can be safely modified by this technique. We believe that this decade-old method deserves a fresh look, particularly in the context of high-throughput structural biology, as it offers an inexpensive, reliable and rapid means of altering the crystallization properties of single proteins or protein complexes without resorting to site-directed mutagenesis.

We would like to thank the NIGMS-funded Southeast Collaboratory for Structural Genomics, where FS initiated his studies on reductive methylation in 2001. We also thank Dr David Garboczi (NIAID, NIH) for valuable advice on the experimental aspects of the modification reaction. Electrospray mass-spectrometry experiments were conducted on the LC/ES-MS instrument maintained by the Biophysics Resource in the Structural Biophysics Laboratory, Center for Cancer Research, National Cancer Institute at Frederick. X-ray diffraction data were collected at the Southeast Regional Collaborative Access Team (SER-CAT) 22-ID beamline at the Advanced Photon Source, Argonne National Laboratory. Supporting institutions may be found at http://www.ser-cat.org/members.html. Use of the Advanced Photon Source was supported by the US Department of Energy, Office of Science, Office of Basic Energy Sciences under Contract No. W-31–109-Eng-38.

References

Chayen, N. E. (1997). *Structure*, **5** 1269–1274.
Day, J. B. & Plano, G. V. (1998). *Mol. Microbiol.* **30**, 777–788.

DeLano, W. L. (2001). *PyMOL.* DeLano Scientific, LLC, San Carlos, CA, USA.
Derewenda, U., Mateja, A., Devedjiev, Y., Routzahn, K. M., Evdokimov, A. G., Derewenda, Z. S. & Waugh, D. S. (2004). *Structure*, **12**, 301–306.
Derewenda, Z. S. (2004). *Structure*, **12**, 529–535.
Doublié, S. (1997). *Methods Enzymol.* **276**, 523–530.
Jackson, M. W., Day, J. B. & Plano, G. V. (1998). *J. Bacteriol.* **180**, 4912–4921.
Janda, I., Devedjiev, Y., Cooper, D., Chruszcz, M., Derewenda, U., Gabrys, A., Minor, W., Joachimiak, A. & Derewenda, Z. S. (2004). *Acta Cryst.* D**60**, 1101–1107.
Kurinov, I. V., Mao, C., Irvin, J. D. & Uckun, F. M. (2000). *Biochem. Biophys. Res. Commun.* **275**, 549–552.
Lo Conte, L., Chothia, C. & Janin, J. (1999). *J. Mol. Biol.* **285**, 2177–2198.
Longenecker, K. L., Garrard, S. M., Sheffield, P. J. & Derewenda, Z. S. (2001). *Acta Cryst.* D**57**, 679–688.
Mateja, A., Devedjiev, Y., Krowarsch, D., Longenecker, K., Dauter, Z., Otlewski, J. & Derewenda, Z. S. (2002). *Acta Cryst.* D**58**, 1983–1891.
O'Toole, N., Grabowski, M., Otwinowski, Z., Minor, W. & Cygler, M. (2004). *Proteins*, **56**, 201–210.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Rayment, I. (1997). *Methods Enzymol.* **276**, 171–179.
Rayment, I., Rypniewski, W. R., Schmidt-Base, K., Smith, R., Tomchick, D. R., Benning, M. M., Winkelmann, D. A., Wesenberg, G. & Holden, H. M. (1993). *Science*, **261**, 50–58.
Rypniewski, W. R., Holden, H. M. & Rayment, I. (1993). *Biochemistry*, **32**, 9851–9858.