Acta Crystallographica Section F
Structural Biology and Crystallization Communications
ISSN 1744-3091

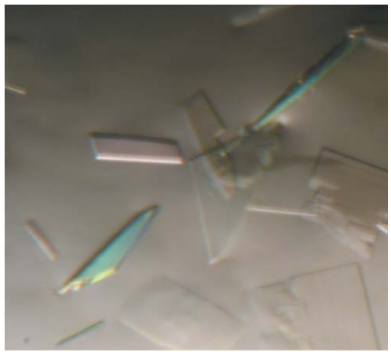
Expression, purification and preliminary structural analysis of the head domain of *Deinococcus radiodurans* RecN

Simone Pellegrino,[a] Jens Radzimanowski,[a,b] Sean McSweeney[a] and Joanna Timmins[a,c]*

[a]Structural Biology Group, European Synchrotron Radiation Facility, 6 Rue Jules Horowitz, 38043 Grenoble CEDEX 9, France, [b]Unit of Virus Host–Cell Interactions, UJF–EMBL–CNRS, UMI3265, 38043 Grenoble CEDEX 9, France, and [c]Institut de Biologie Structurale J.-P. Ebel, 41 Rue Jules Horowitz, 38027 Grenoble CEDEX, France

Correspondence e-mail: joanna.timmins@ibs.fr

Received 18 October 2011
Accepted 16 November 2011



Deinococcus radiodurans is well known for its extreme tolerance to harsh conditions and for its extraordinary ability to repair DNA. Double-strand breaks (DSBs) are the most hazardous lesions that can be induced by ionizing radiation, and homologous recombination (HR) is the principal mechanism by which the integrity of the DNA is restored. In *D. radiodurans* the RecFOR complex is the main actor in HR and the RecN protein is believed to play an important role in DSB recognition. Here, SAXS and preliminary X-ray diffraction studies are presented of the head domain, which is the globular region formed upon interaction of the N- and C-terminal domains of RecN. The crystal structure of this domain was solved using the single-wavelength anomalous dispersion method. Model building and refinement are in progress.

1. Introduction

Deinococcus radiodurans is an extremophilic bacterium that is able to survive high doses of ionizing radiation, which cause lesions in the DNA, especially in the form of double-strand breaks (DSBs). Many factors contribute to the innate resistance of *D. radiodurans* to DNA damage, such as a high genome copy number (between four and ten in the logarithmic phase) and a very efficient DNA-repair pathway (Battista, 1997).

DNA repair in *D. radiodurans* occurs through extended synthesis-dependent strand annealing (ESDSA) followed by homologous recombination (HR) mediated by RecFOR (Bentchikou *et al.*, 2010). In other bacteria, such as *Escherichia coli* for instance, DSBs are mainly repaired by the RecBCD pathway. However, in *D. radiodurans* RecBC homologues are missing from the genome. On the other hand, RecFOR homologues are encoded and are predicted to accomplish recombinational repair, mediating the loading of RecA protein onto the damaged DNA.

What remains largely unknown is how the DSBs are recognized and how the RecFOR machinery is recruited to the damaged DNA in order to trigger homologous recombination. In eukaryotes, the MRN complex (Mre11–Rad50–Nbs1) has been shown to be responsible for DSB recognition and sister chromatid tethering (Weitzman *et al.*, 2010). In prokaryotes, although homologues of Mre11 and Rad50 can be found, at present the best candidate for DSB recognition is the RecN protein (Ayora *et al.*, 2011; Kidane *et al.*, 2004).

RecN belongs to the structural maintenance of chromosomes (SMC) family, which plays a critical role in DNA replication and repair. In these proteins, the N- and C-terminal domains form a unique globular domain called the head domain and are connected by a long antiparallel coiled-coil region. The organization of the head domain creates a functional ATP-binding pocket, which then allows the protein to carry out nucleotide hydrolysis through a mechanism similar to the ATP-binding cassette (ABC) proteins (Hirano, 2002).

Sequence annotation predicted a potential nucleotide-binding region in RecN. We designed a construct based on previous crystal structures obtained for the pfRad50 ATPase domain (PDB entry 1ii8; Hopfner *et al.*, 2001) and the tmSMC head domain (PDB entry 1e69; Löwe *et al.*, 2001). The predicted RecN head domain was cloned, expressed, purified and crystallized. Structure determination will be relevant for initial understanding of the mechanism of DSB recognition in *D. radiodurans*.

2. Materials and methods

2.1. Sequence amplification and cloning

DNA encoding the RecN head domain (residues 1–195 and 365–564) was amplified from genomic DNA of *D. radiodurans*. A short 14-residue linker with amino-acid sequence ESSKHPTSLVPRGS was introduced between residues 195 and 365 to replace the coiled-coil domain (Löwe *et al.*, 2001). The N-terminal fragment was amplified using the following primers: 5′-CACCATGACCCGCAAGGCCCGTA-3′ as the forward primer and 5′-CACAAGCGACGTTGGATGCTTGCTCGACTCGCTGGCCTGGAGGCGCTCC-3′ as the reverse primer. A second PCR reaction was carried out in order to amplify the C-terminal fragment using the primers 5′-GCCAGCGAGTCGAGCAAGCATCCAACGTCGCTTGTGCCACGAG-3′ and 5′-CCAACGTCGCTTGTGCCACGAGGCAGCGTGGACGCCCTGCACGCCG-3′ as overlapping forward primers and 5′-TTAGCCAGCCAGCAACTCGC-3′ as the reverse primer. The two resulting DNA fragments were then mixed in equimolar concentrations and amplified using the N-terminal forward primer and the C-terminal reverse primer. The head-domain sequence was subcloned into pET151-Topo vector (Invitrogen) and the resulting construct contains a cleavable N-terminal hexahistidine tag.

2.2. Protein expression and purification

The plasmid was transformed into competent *E. coli* BL21* (DE3) cells (Invitrogen) by the heat-shock method. The cells were grown at 310 K in LB (lysogeny broth) medium containing ampicillin (100 µg ml^{-1}). Protein expression was induced using 1 m*M* isopropyl *β*-D-1-thiogalactopyranoside (IPTG) when the OD_{600} reached ∼0.8–1.0 and took place overnight at 293 K. The cells were harvested by centrifugation at 7548*g* for 30 min and were then resuspended in lysis buffer consisting of 50 m*M* Tris pH 8, 1 *M* NaCl, 5 m*M* $MgCl_2$ and 5% glycerol. Lysis was achieved by mechanical force in a cell disrupter operated at 195 MPa at 277 K. The lysate was centrifuged at 48 384*g* at 277 K. The supernatant was collected and loaded onto a 5 ml HisTrap HP column (GE Healthcare) previously equilibrated in buffer *A* (50 m*M* Tris pH 8.0, 300 m*M* NaCl, 5 m*M* $MgCl_2$, 5 m*M* imidazole), followed by a wash with at least five column volumes of buffer *A*. The protein was eluted in a linear gradient of buffer *B* (50 m*M* Tris pH 8.0, 300 m*M* NaCl, 5 m*M* $MgCl_2$, 500 m*M* imidazole) from 0 to 100% over 40 ml at about 145 m*M* imidazole. Fractions containing the target protein (checked by SDS–PAGE) were pooled and dialyzed overnight at 277 K in buffer *C* (50 m*M* Tris pH 8.0, 300 m*M* NaCl, 5 m*M* $MgCl_2$). At the same time, His-tag cleavage was performed by addition of TEV protease (to a final concentration of 0.1 mg ml^{-1}), 0.5 m*M* EDTA and 1 m*M* DTT to the fractions.

The head domain was then concentrated to 10.4 mg ml^{-1} by ultrafiltration (Amicon Ultracel 10K) and was further purified on a Superdex S200 HiLoad column (GE Healthcare) equilibrated in buffer *C*. Protein concentrations were estimated using the Bradford protein assay (Bio-Rad).

2.3. Small-angle X-ray scattering

Sample purity and homogeneity were checked by size-exclusion chromatography (SEC) and dynamic light scattering (DLS) in order to assess the quality of the sample prior to biophysical and structural characterization.

Small-angle X-ray scattering (SAXS) experiments (Fig. 1) were performed on ID14-3 (Pernot *et al.*, 2010) at the ESRF in Grenoble, France using a fixed energy of 13.32 keV (λ = 0.931 Å) and employing a Pilatus 1M pixel detector. Protein samples were tested at three different concentrations (3.43, 2.34 and 1.12 mg ml^{-1}) to exclude inter-particle effects. Ten frames with an exposure time of 10 s were recorded and averaged to improve the signal-to-noise ratio for each of the three samples. In order to avoid radiation damage, the frames

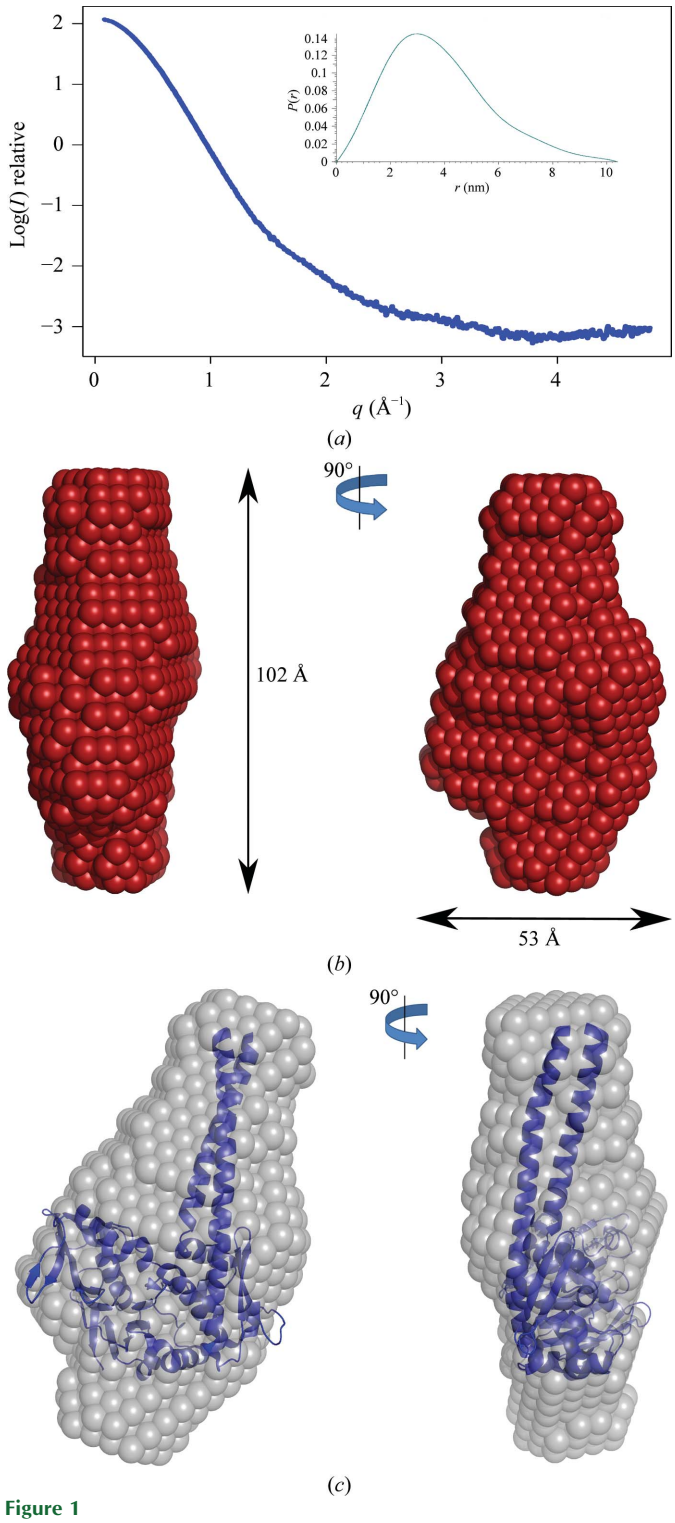
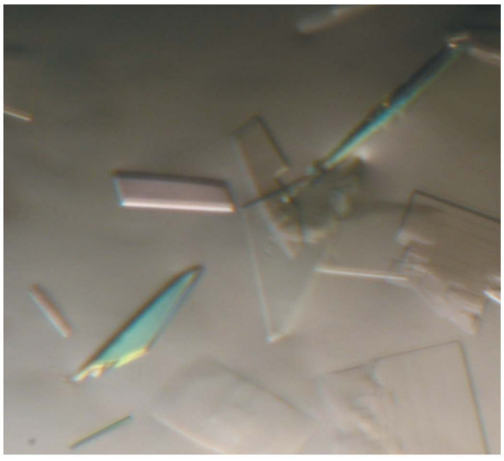


Figure 1
(*a*) Experimental scattering curve of the RecN head domain. The pair distribution function [*P*(*r*)] is shown in the inset. (*b*) Model built by *DAMMIN*. Several independent models were averaged and filtered in order to obtain the most probable envelope. Length and width are indicated. (*c*) Superposition of the crystal structure of the pfRad50 ATPase domain with the *ab initio* model obtained for the RecN head domain in solution.

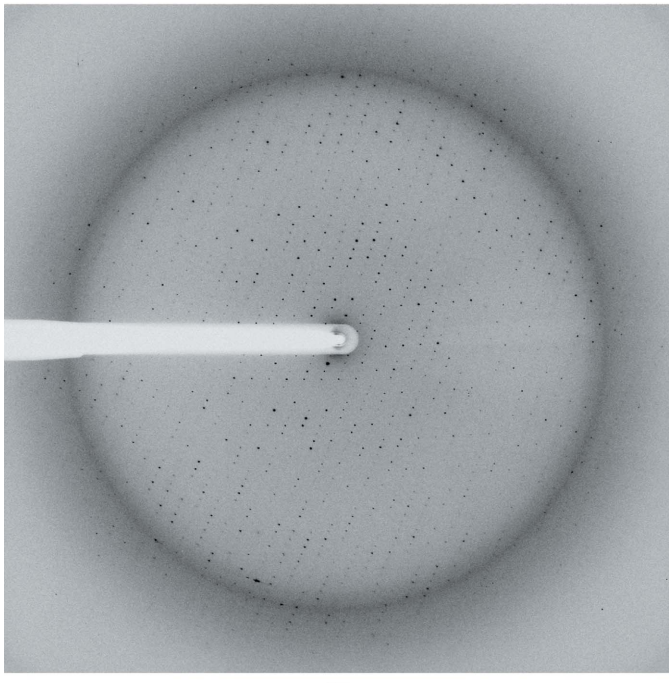
were collected while the sample was flowing in the capillary tube so as to continuously expose fresh sample to the X-ray beam. Data for the buffer were also recorded before and after each sample to be used for subtraction from the scattering curve of the protein sample.

Experimental data were processed with *Primus* (Konarev *et al.*, 2003) and the R_g (radius of gyration) was evaluated using the Guinier approximation. The subtracted curves were used as input files for *GNOM* (Svergun, 1992), from which the pair distribution function $P(r)$ and D_{max} were calculated. An estimation of the molecular weight was extrapolated directly from the Porod volume calculation using the program tool *AUTOPOROD* (Petoukhov *et al.*, 2007).

Ab initio models were generated using *DAMMIN* (Svergun, 1999), assuming *P*1 symmetry for each run. 20 independent reconstructions were aligned, averaged and filtered using the *DAMAVER* program package (Volkov & Svergun, 2003). Superposition of the resulting *ab initio* model with the crystal structure of the ATPase domain of pfRad50 (Fig. 1*c*) was performed using *SUPCOMB* (Kozin & Svergun, 2001).



(*a*)



(*b*)

Figure 2
(*a*) Crystals of SeMet-derivatized RecN head domain. The crystals grew after 1 d of vapour-diffusion equilibration. (*b*) A typical 0.5° oscillation diffraction image of the RecN head domain crystals collected on an ADSC Q315R detector.

2.4. Crystallization

Initial crystallization screenings were performed at 293 K using Greiner CrystalQuick sitting-drop vapour-diffusion plates. A Cartesian PixSys 4200 crystallization robot (High Throughput Crystallization Laboratory at EMBL Grenoble) was used to test 576 different crystallization conditions (using the method described in Dimasi *et al.*, 2007). The following commercial screens from Hampton Research were set up: Crystal Screen, Crystal Screen 2, Crystal Screen Lite, PEG/Ion, MembFac, Natrix, QuickScreen, Grid Screens (Ammonium Sulfate, Sodium Malonate, Sodium Formate, PEG 6K, PEG/LiCl and MPD) and Index. Crystals were obtained in several conditions and the most promising ones, which were from the Index screen, were used as a starting point for optimization using the hanging-drop method at 293 K. Single protein crystals appeared in condition No. 44 from the Index screen. Drops were subsequently set up screening different values of pH (6–9) and PEG concentration (15–30%). The crystals (Fig. 2) were grown in droplets consisting of 1 µl head domain protein solution (10.4 mg ml^{-1}) and 1 µl reservoir solution (0.1 *M* Tris pH 7.5, 25% PEG 3350) after 1–2 d of equilibration against 500 µl reservoir in the well.

2.5. X-ray diffraction analysis and structure determination

Crystals of the RecN head domain were tested on the ID23-2 beamline (Flot *et al.*, 2010) at the ESRF, Grenoble, France. Crystals were cryoprotected in 0.1 *M* Tris pH 7.5, 25%(*v*/*v*) PEG 3350 and 25% glycerol and were flash-cooled in liquid N_2. Diffraction data were collected at 100 K using an X-ray wavelength of 0.8726 Å and an ADSC Q315R detector. Data were collected to 3 Å resolution, processed with *iMOSFLM* (Battye *et al.*, 2011) and scaled with *SCALA* (Evans, 2006). Estimation of the Matthews coefficient was performed using the *MATTHEW_COEFF* tool (Kantardjieff & Rupp, 2003; Matthews, 1968); the solvent content was estimated to be 59%.

Initial structure-determination trials involved molecular replacement (MR) using *Phaser* (McCoy *et al.*, 2007). The ATPase domain of pfRad50, which displays 47% sequence similarity to the RecN head domain, and the head domain of tmSMC were used as search models. However, molecular replacement failed to provide a clear solution.

Subsequently, SeMet-substituted head domain was produced using the method described by Doublié (1997) and purified to homogeneity using the protocol established for the native protein. Crystals were obtained under the same conditions as used for the native RecN head domain and were tested for diffraction analysis. An energy scan performed on the ID14-4 beamline (McCarthy *et al.*, 2009) gave a clear absorption spectrum, from which it was possible to determine the energy corresponding to the peak of the selenium absorption edge. A complete single-wavelength anomalous dispersion (SAD) data set was collected at this peak wavelength for experimental phasing.

Data were processed with *iMOSFLM* and the averaged intensities were obtained with *SCALA*. The calculated structure factors were then used for substructure determination using the *SHELX* program package (Sheldrick, 2010). The positions of the Se atoms were determined using *SHELXD*. The substructure coordinate file was then submitted to *Phaser* (McCoy *et al.*, 2007), which is integrated in *AutoSol* in *PHENIX* (Adams *et al.*, 2010; Terwilliger *et al.*, 2009), for refinement of heavy-atom coordinates and occupancies. *Phaser* found

Table 1
Data-collection statistics.

Values in parentheses are for the highest resolution shell.

	Native	Se peak
Beamline	ID23-2, ESRF	ID14-4, ESRF
Space group	$P2_1$	$P2_1$
Unit-cell parameters (Å, °)	a = 130.0, b = 62.4, c = 134.8, $\alpha = \gamma$ = 90.00, β = 102.2	a = 129.9, b = 62.0, c = 133.8, $\alpha = \gamma$ = 90.00, β = 102.7
Resolution range (Å)	64.12–2.98 (3.14–2.98)	50.00–3.00 (3.16–3.00)
Oscillation range (°)	0.5	0.5
Wavelength (Å)	0.873	0.9795
Mosaicity (°)	0.63	0.94
Total No. of reflections	161489 (22164)	216096 (31862)
No. of unique reflections	43217 (6016)	41994 (6098)
Completeness (%)	99.4 (95.7)	99.6 (100)
$R_{\rm merge}$† (%)	12.8 (74.9)	12.4 (41.9)
$R_{\rm p.i.m.}$‡ (%)	7.7 (45.5)	6.7 (22.9)
Mean $I/\sigma(I)$	9.0 (1.9)	9.2 (3.3)
Multiplicity	3.7 (3.7)	5.1 (5.2)
Anomalous multiplicity		2.7 (2.7)
Molecules per asymmetric unit	4	4
$V_{\rm M}$ (Å^3 Da^{-1})	3.14	3.13
$CC_{\rm anom}$§		5.3
SHELXD ($CC_{\rm all}/CC_{\rm weak}$)		31.81/19.16
FOM (HA refinement)¶ (%)		27.8
FOM (DM)†† (%)		70

† $R_{\rm merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I(hkl)$ is the integrated intensity for a given reflection. ‡ $R_{\rm p.i.m.} = \sum_{hkl}\{1/[N(hkl)-1]\}^{1/2}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$. § Correlation coefficient from *SHELXC*. ¶ Figure of merit from *Phaser*. †† Figure of merit from *RESOLVE*.

16 Se sites, which corresponded to the four molecules in the asymmetric unit. Data statistics are given in Table 1.

The partial model built by *AutoSol* was then submitted to *AutoBuild* (Terwilliger *et al.*, 2008) for more extensive model building while improving the electron-density map. Manual completion of model building and refinement are under way.

3. Results and discussion

Expression of the RecN head domain resulted in a typical yield of about 100 mg protein per litre of bacterial cell culture. The protein was purified by standard protocols as described above and gel-filtration chromatography yielded purified head domain (the purity was checked by SDS–PAGE) with a corresponding molecular weight of ∼43 kDa.

Dynamic light-scattering measurements performed on the purified protein provided an estimation of the hydrodynamic radius, which was found to be 3.3 nm. Polydispersity was measured and found to be reasonable (PdI of 21%, which is typical for a monodisperse sample).

Initial crystallization screens resulted in crystals of different morphologies, mostly plates and needle clusters, in a broad variety of conditions. Optimized crystals for diffraction analysis grew within 1 d at 293 K. Droplets were set up by mixing 1 µl protein solution (at 10.4 mg ml^{-1}) with 1 µl precipitant solution consisting of 0.1 *M* Tris pH 8 and 25% PEG 3350.

The best crystal diffracted to a limiting resolution of 3 Å. Structure determination by molecular replacement failed, so a SAD experiment using selenomethionine-incorporated protein crystals was performed instead. Structure determination and model building are ongoing.

Small-angle X-ray scattering (SAXS) was performed on the RecN head domain. Guinier analysis resulted in an estimated radius of gyration of 2.95 nm. Based on calculation of the Porod volume (Glatter & Kratky, 1982), the molecular weight of the RecN head domain in solution is estimated to be ∼46.5 kDa (monomer size: 43 kDa). The resulting model suggests a relatively elongated structure and superposition of the Rad50 head domain with our SAXS model shown in Fig. 1(*c*) suggests that the RecN head domain adopts a similar overall fold to the Rad50 ATPase structure (Hopfner *et al.*, 2001) used in the unsuccessful molecular-replacement trials.

The data for this work were collected at the ESRF on the ID14-3 beamline for the SAXS analysis and on the ID14-4 and ID23-2 beamlines for the crystallographic part. We thank the beamline staff for assistance and advice during data collection.

References

Adams, P. D. *et al.* (2010). *Acta Cryst.* D**66**, 213–221.
Ayora, S., Carrasco, B., Cárdenas, P. P., César, C. E., Cañas, C., Yadav, T., Marchisone, C. & Alonso, J. C. (2011). *FEMS Microbiol. Rev.* **35**, 1055–1081.
Battista, J. R. (1997). *Annu. Rev. Microbiol.* **51**, 203–224.
Battye, T. G. G., Kontogiannis, L., Johnson, O., Powell, H. R. & Leslie, A. G. W. (2011). *Acta Cryst.* D**67**, 271–281.
Bentchikou, E., Servant, P., Coste, G. & Sommer, S. (2010). *PLoS Genet.* **6**, e1000774.
Dimasi, N., Flot, D., Dupeux, F. & Márquez, J. A. (2007). *Acta Cryst.* F**63**, 204–208.
Doublié, S. (1997). *Methods Enzymol.* **276**, 523–530.
Evans, P. (2006). *Acta Cryst.* D**62**, 72–82.
Flot, D., Mairs, T., Giraud, T., Guijarro, M., Lesourd, M., Rey, V., van Brussel, D., Morawe, C., Borel, C., Hignette, O., Chavanne, J., Nurizzo, D., McSweeney, S. & Mitchell, E. (2010). *J. Synchrotron Rad.* **17**, 107–118.
Glatter, O. & Kratky, O. (1982). *Small-angle X-ray Scattering*. London: Academic Press.
Hirano, T. (2002). *Genes Dev.* **16**, 399–414.
Hopfner, K. P., Karcher, A., Craig, L., Woo, T. T., Carney, J. P. & Tainer, J. A. (2001). *Cell*, **105**, 473–485.
Kantardjieff, K. A. & Rupp, B. (2003). *Protein Sci.* **12**, 1865–1871.
Kidane, D., Sanchez, H., Alonso, J. C. & Graumann, P. L. (2004). *Mol. Microbiol.* **52**, 1627–1639.
Konarev, P. V., Volkov, V. V., Sokolova, A. V., Koch, M. H. J. & Svergun, D. I. (2003). *J. Appl. Cryst.* **36**, 1277–1282.
Kozin, M. B. & Svergun, D. I. (2001). *J. Appl. Cryst.* **34**, 33–41.
Löwe, J., Cordell, S. C. & van den Ent, F. (2001). *J. Mol. Biol.* **306**, 25–35.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
McCarthy, A. A., Brockhauser, S., Nurizzo, D., Theveneau, P., Mairs, T., Spruce, D., Guijarro, M., Lesourd, M., Ravelli, R. B. G. & McSweeney, S. (2009). *J. Synchrotron Rad.* **16**, 803–812.
McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
Pernot, P., Theveneau, P., Giraud, T., Fernandes, R. N., Nurizzo, D., Spruce, D., Surr, J., McSweeney, S., Round, A., Felisaz, F., Foedinger, L., Gobbo, A., Huet, J., Villard, C. & Cipriani, F. (2010). *J. Phys. Conf. Ser.* **247**, 012009.
Petoukhov, M. V., Konarev, P. V., Kikhney, A. G. & Svergun, D. I. (2007). *J. Appl. Cryst.* **40**, s223–s228.
Sheldrick, G. M. (2010). *Acta Cryst.* D**66**, 479–485.
Svergun, D. I. (1992). *J. Appl. Cryst.* **25**, 495–503.
Svergun, D. I. (1999). *Biophys. J.* **76**, 2879–2896.
Terwilliger, T. C., Adams, P. D., Read, R. J., McCoy, A. J., Moriarty, N. W., Grosse-Kunstleve, R. W., Afonine, P. V., Zwart, P. H. & Hung, L.-W. (2009). *Acta Cryst.* D**65**, 582–601.
Terwilliger, T. C., Grosse-Kunstleve, R. W., Afonine, P. V., Moriarty, N. W., Zwart, P. H., Hung, L.-W., Read, R. J. & Adams, P. D. (2008). *Acta Cryst.* D**64**, 61–69.
Volkov, V. V. & Svergun, D. I. (2003). *J. Appl. Cryst.* **36**, 860–864.
Weitzman, M. D., Lamarche, B. J. & Orazio, N. I. (2010). *FEBS Lett.* **584**, 3682–3695.