Acta Crystallographica Section D
**Biological Crystallography**
ISSN 0907-4449

# Small-angle neutron scattering and contrast variation: a powerful combination for studying biological structures

**William T. Heller**

Center for Structural Molecular Biology and Chemical Sciences Division, Oak Ridge National Laboratory, PO Box 2008, MS-6393, Oak Ridge, TN 37831, USA

Correspondence e-mail: hellerwt@ornl.gov

Received 16 April 2010
Accepted 13 May 2010

The use of small-angle scattering (SAS) in the biological sciences continues to increase, driven as much by the need to study increasingly complex systems that are often resistant to crystallization or are too large for NMR as by the availability of user facilities and advancements in the modelling of biological structures from SAS data. SAS, whether with neutrons (SANS) or X-rays (SAXS), is a structural probe of length scales ranging from 10 to 10 000 Å. When applied to biological complexes in dilute solution, it provides size and shape information that can be used to produce structural models that can provide insight into function. SANS enables the use of contrast-variation methods through the unique interaction of neutrons with hydrogen and its isotope deuterium. SANS with contrast variation enables the visualization of components within multisubunit complexes, making it a powerful tool for probing protein–protein and protein–nucleic acid complexes, as well as the interaction of proteins with lipids and detergents.

## 1. Introduction

In the post-genomic era, the next great challenge is to understand the structures and interactions that exist between an expressed protein in an organism and its complex environment consisting of other proteins, nucleic acids and the macromolecular assemblies that comprise a living cell. Crystallography, NMR and electron microscopy (EM) are the primary tools of structural biology. Crystallography and NMR provide atomic resolution structures of proteins and their complexes. The resolution of EM continues to improve, producing increasingly detailed three-dimensional structures of biological macromolecules. While very powerful, these techniques are not always successful at determining the structures of biological macromolecules. Crystallography and EM are not readily applicable to highly flexible and dynamic systems as a result of the state of the sample during the measurement, while NMR cannot be readily applied to large systems.

Small-angle scattering (SAS) is a complementary method for probing biological structures that can be applied to systems ranging from isolated proteins and their complexes to complex hierarchical assemblies such as bone, tissue or biomass, across length scales ranging from 10 to 10 000 Å. SAS methods do not provide atomic resolution structural information such as can be obtained using crystallography and NMR, but they can be applied to systems of arbitrary size. Similarly, highly dynamic systems can be readily probed using SAS. When applied to structural biology, SAS using X-rays (SAXS) or

**Table 1**
Coherent nuclear scattering lengths of selected elements and isotopes (Koester *et al.*, 1991).

With the exception of hydrogen and deuterium, the value presented is the isotopic abundance-weighted average.

| Atom or isotope | $b_{coh}$ ($10^{-15}$ m) |
|---|---|
| $^{1}$H | −3.74 |
| D ($^{2}$H) | 6.67 |
| C | 6.65 |
| N | 9.36 |
| O | 5.80 |
| S | 2.84 |
| P | 5.13 |
| Ti | −3.37 |
| Fe | 9.45 |
| Au | 7.90 |

neutrons (SANS) is employed because the wavelengths are well matched to the length scales being probed. X-rays interact with the electrons in a sample, while neutrons interact with atomic nuclei. It is the interaction of neutrons with the atomic nucleus that enables contrast-variation methods for probing the internal structure of complex assemblies using SANS. In this article, a brief introduction to these methods is presented, with examples from the literature that illustrate the power of the technique for structural biology.

## 2. Theory

Here, the discussion will be restricted to what is often the simplest conceptual case for SAS, which is also the most applicable for problems in structural biology, namely a dilute solution of monodisperse particles. For a more detailed introduction to SAS, the reader is referred to the comprehensive reviews covering SAXS and SANS that are available (Jacrot, 1976; Svergun & Koch, 2003; Putnam *et al.*, 2007). The SAS intensity profile, $I(Q)$, measured from a dilute solution of particles is given by (Debye & Bueche, 1949; Guinier & Fournet, 1955)

$$I(Q) = n\left\langle \left| \int_V [\rho(\mathbf{R}) - \rho_s] \exp(-i\mathbf{Q}\cdot\mathbf{R})\, \mathrm{d}^3R \right|^2 \right\rangle. \qquad (1)$$

The quantity $n$ is the number of particles per unit volume, $\rho(\mathbf{R})$ is the scattering-length density of the particle at position $\mathbf{R}$ inside the particle and $\rho_s$ is the scattering-length density of the solvent. The entire volume of the particle, $V$, is used for the integration. The integral is also taken over all orientations of the particle with respect to the incident beam. The measured $I(Q)$ is also averaged over the entire ensemble of structures present in the sample during the period of the measurement. In (1), the quantity $\rho(\mathbf{R}) - \rho_s$, being the scattering-length density within the particle relative to the surrounding solvent, is called the contrast.

The scattering-length density of an object is a function of its elemental composition, and the associated atomic scattering lengths (specifically the coherent scattering lengths), denoted $b_{coh}$, are a measure of the strength of the interaction of an X-ray or neutron with an atom. For X-rays, the scattering length is a linear function of the atomic number. As a result, light elements scatter poorly relative to those with higher atomic number, with hydrogen, a critical element in biology, often not being resolved in all but the highest resolution protein crystal structures. The nuclear scattering lengths of all atoms are comparable in magnitude (Koester *et al.*, 1991), making light atoms relatively simple to visualize by neutron scattering methods. Examples of several nuclear scattering lengths are provided in Table 1. Interestingly, the nuclear scattering length is isotope-dependent and can be either positive or negative, although there are far fewer negative scattering lengths than positive ones. Of particular interest to structural biology is the fact that hydrogen and its isotope deuterium have such dramatically different scattering lengths. It is this difference that enables contrast variation for structural biology through the use of mixtures of $H_2O$ and $D_2O$ to provide a continuous spectrum of values for $\rho_s$.

The true utility of being able to change the scattering-length density of the solvent relative to that of the scattering particle becomes evident when working with structures composed of materials having different scattering-length densities. Such complexes are very common in biology, where proteins, lipids and nucleic acids all have inherently different average scattering lengths. Furthermore, it is often possible to produce deuterium-enriched biological materials, making it possible to reconstitute multi-component structures with selectively deuterated subunits. Example scattering-length densities of various biological macromolecules, including examples involving deuterium-labeled material, are shown in Fig. 1 as a function of the $H_2O/D_2O$ mixture in the background solution. The calculated scattering-length densities include the impact of H/D exchange within the structures. One interesting feature of the curves in the figure is the set of points where the scat-

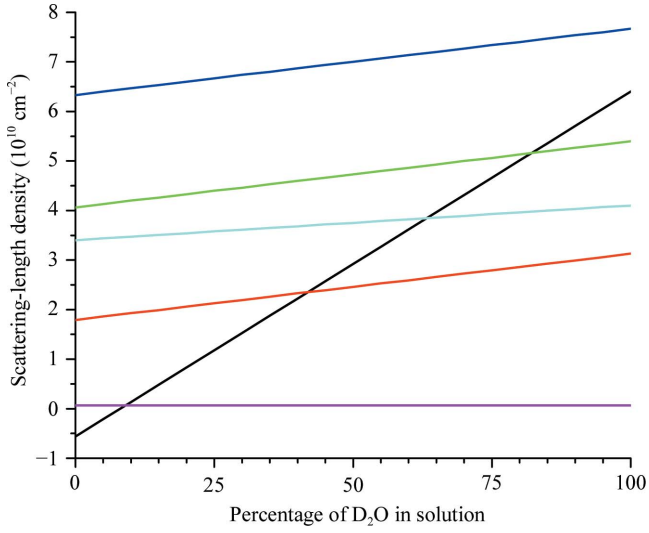


**Figure 1**
Scattering-length densities as a function of $D_2O$ for water (black), hydrogenated protein (red), protein with 50% of the protons that do not exchange substituted by deuterium (green), protein with 100% of the protons that do not exchange substituted by deuterium (blue), DNA with an equal distribution of base pairs (cyan) and the lipid dimyristoyl phosphatidylcholine (violet).

tering-length density of the solvent matches that of the various materials. At these points, the contrast, and therefore the measured intensity, of that material is zero. An $I(Q)$ measured from a particle that is a complex of that material and a second material having a different scattering-length density would only result from the part of the sample having a nonzero contrast. The result is structural information on one component of the complex within the intact complex. This particular kind of experiment is known as a contrast-matching experiment.

An extension of a contrast-matching experiment is the contrast-variation experiment, in which the SANS intensity profile from a complex is measured in a series of $H_2O/D_2O$ mixtures. In the case of a binary complex consisting of subunits having different average scattering-length densities that will be denoted $\rho_1$ and $\rho_2$, the measured data sets can be described, according to Ibel & Stuhrmann (1975), as

$$I(Q, \rho_1, \rho_2) = (\rho_1 - \rho_s)^2 I_1(Q) + (\rho_2 - \rho_s)^2 I_2(Q) + (\rho_1 - \rho_s)(\rho_2 - \rho_s) I_{12}(Q). \quad (2)$$

The functions $I_1(Q)$ and $I_2(Q)$ are known as the basic scattering functions and correspond to the individual subunits within the structure of the complex. $I_{12}(Q)$ is commonly known as the cross-term and results from the scattering between the subunits in the complex. The cross-term provides information about the relative disposition of the two subunits and is information that cannot be obtained directly from a contrast-matching experiment. A measured contrast-variation series defines a set of linear equations in the three basic scattering functions that can be extracted computationally; for example, with the method implemented in the program *MULCH* (Whitten, Cai *et al.*, 2008). An example contrast-variation series calculated using the program *ORNL_SAS* (Tjioe & Heller, 2007) from a dimer of human serum albumin found in the crystal structure (PDB code 1ao6; Sugio *et al.*, 1999) in which one subunit was hydrogenated and the other completely deuterated is shown in Fig. 2. The interplay of the constructive and destructive interference that results from the opposite signs of the scattering lengths of hydrogen and deuterium can be readily seen in the profiles calculated at high $D_2O$ content (>70% $D_2O$), resulting in a depression of the forward scattering.

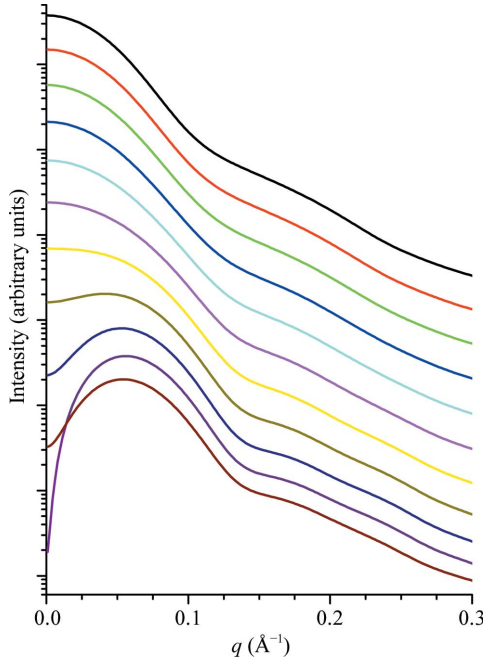


**Figure 2**
Simulated contrast-variation series calculated using *ORNL_SAS* (Tjioe & Heller, 2007) from the crystal structure of the human serum albumin dimer (PDB code 1ao6; Sugio *et al.*, 1999). One polypeptide chain was hydrogenated and the other was perdeuterated. The curves are 0% $D_2O$ (black), 10% $D_2O$ (red), 20% $D_2O$ (green), 30% $D_2O$ (blue), 40% $D_2O$ (cyan), 50% $D_2O$ (violet), 60% $D_2O$ (yellow), 70% $D_2O$ (gold), 80% $D_2O$ (dark blue), 90% $D_2O$ (purple) and 100% $D_2O$ (brown). The curves have been offset for clarity.

## 3. Examples from the literature

SANS with contrast variation has been extensively applied to problems in structural biology. Protein–protein and protein–nucleic acid complexes are the most appropriate systems in the context of this review, but other examples are also presented. In the case of protein–protein complexes, the ability to apply contrast variation requires selective deuteration of one subunit, while SANS with contrast variation can be applied to protein–nucleic acid complexes without deuterium labeling. Here, several examples from both classes of problems are presented.

Perhaps the classical set of examples of the application of SANS with contrast variation involved the study of various ribosomes, the macromolecular machines that assemble proteins from amino acids. The earliest studies probed the internal structure of the 30S ribosome (Langer *et al.*, 1978; Ramakrishnan *et al.*, 1981, 1984). Contrast-variation methods were used to determine the relative distances between subunits in the multisubunit complex. Ultimately, this research led to a structural model of the disposition of the subunits in space (Capel *et al.*, 1987).

These early studies were followed by subsequent studies of the larger 50S and 70S ribosomes. A map of the distribution of protein and RNA within the 50S ribosome from *Escherichia coli* was generated using SANS with contrast variation and shape restoration by spherical harmonics (Svergun, Koch, Pedersen *et al.*, 1994; Svergun, Koch & Serdyuk, 1994; Svergun, Pedersen *et al.*, 1994). The *E. coli* 70S ribosome was also studied by SANS with contrast variation and selective deuteration. A total of 42 data sets were collected and again modeled using spherical harmonics (Svergun *et al.*, 1997*a*,*b*). A later model was generated from the same series of data using the program *DAMMIN* (Svergun & Nierhaus, 2000).

Another protein complex extensively characterized by SANS with contrast variation is the troponin complex, which serves to regulate muscle contraction in response to a calcium signal. Both skeletal and cardiac variants of the three-subunit (TnC, TnI and TnT) complex have been studied. The binary TnC–TnI complex was the first troponin complex to be studied, providing important insight into the interaction

between the calcium-binding TnC subunit and the inhibitor TnI subunit (Olah *et al.*, 1994; Olah & Trewhella, 1994; Stone *et al.*, 1998). Later work expanded to the study of the ternary TnC–TnI–TnT complex and shed light on the organization of the three subunits using low-resolution modeling methods (Heller *et al.*, 2002, 2003) and ultimately modeling using high-resolution structural information that provided important insight into the function of the complex (King *et al.*, 2005).

These are hardly the only applications of SANS with contrast variation to protein–protein complexes. The calmodulin–myosin light-chain kinase (MLCK) system was studied in a series of experiments that shed light onto the activation of MLCK that starts with the receipt of a calcium signal by calmodulin (Krueger *et al.*, 1997, 1998). Similarly, SANS with contrast variation and selective labeling made it possible to understand the organization and activation of protein kinase A (Zhao *et al.*, 1998; Heller *et al.*, 2004). The structure of the chaperonin complex GroEL–GroES was also elucidated using this approach (Krueger *et al.*, 2003). Similarly, insight into the interaction of neurolignin with $\beta$-neurexin was obtained with the help of the technique (Comoletti *et al.*, 2007), as were the structures of two microtubule motors (Fujiwara *et al.*, 1995) and the structure of a complex involved in histidine kinase inhibition (Whitten *et al.*, 2007). Other examples of the application of SANS with contrast variation to the study of multisubunit protein complexes exist in the literature (Krueger *et al.*, 2000; Whitten, Jeffries *et al.*, 2008; Li *et al.*, 2009).

As the studies of the ribosomes show, SANS with contrast variation is a powerful tool for probing protein–nucleic acid interactions. The complex of Taq polymerase with DNA was probed to provide insight into the structural rearrangements that take place upon complex formation (Ho *et al.*, 2004). The interaction of the double-stranded DNA-break repair protein Ku with its target DNA has also been probed (Zhao *et al.*, 1999). Differences in the assembly of the methionine repressor MetJ on its target DNA were also probed using SANS with contrast variation and suggested a physical mechanism for the DNA-length-dependent binding of the protein (Augustus *et al.*, 2006). Similarly, a study of a Holliday junction binding protein, RuvA, with its target DNA demonstrated that the DNA structure is sandwiched by multisubunit complexes of the protein (Chamberlain *et al.*, 1998). These are only a few examples in which the application of SANS with contrast variation has provided important insight into protein–nucleic acid complexes.

The technique is not only limited to simple protein and protein–nucleic acid complexes. It has also been successfully applied to the study of virus structures (Jacrot *et al.*, 1977; Chauvin *et al.*, 1978; Cusack *et al.*, 1981, 1985; Kruse *et al.*, 1982; Inoue & Timmins, 1985; Sato *et al.*, 1995). Similarly, bacteriophages are well suited to contrast-variation methods (Kuzmanovic *et al.*, 2003). The natural difference in scattering-length densities of lipids and surfactants also makes SANS with contrast variation well suited to the study of membrane-protein structures, such as the KcsA potassium channel (Zimmer *et al.*, 2006) and light-harvesting complex II (Cardoso *et al.*, 2009). Similarly, the ability to differentiate between components by contrast variation makes it possible to probe protein structures in more complex matrices, such as gels (Pozzo, 2009; Luo *et al.*, 2009). In all cases, the ability to differentiate between various components in a larger system and understand their structures proved invaluable for understanding the larger system.

## 4. Conclusions

The power of SANS is fully realised when combined with contrast-variation methods. Contrast variation enables the visualization of structures within the intact system, providing information that is unavailable to other structural techniques. In addition to leveraging the natural scattering-length density differences of many biological macromolecules, the application of selective deuteration of these materials makes it possible to clearly differentiate between components within complexes. Small-angle scattering is truly a powerful complementary tool to the traditional tools of structural biology.

This research at Oak Ridge National Laboratory's Center for Structural Molecular Biology (Project ERKP291) was supported by the Office of Biological and Environmental Research using facilities supported by the US Department of Energy, managed by UT-Battelle LLC under contract No. DE-AC05-00OR22725. The submitted manuscript has been authored by a contractor of the US Government under Contract DE-AC05-00OR22725. Accordingly, the US Government retains a nonexclusive royalty-free license to publish or reproduce the published form of this contribution, or allow others to do so, for US Government purposes.

## References

Augustus, A. M., Reardon, P. N., Heller, W. T. & Spicer, L. D. (2006). *J. Biol. Chem.* **281**, 34269–34276.
Capel, M. S., Engelman, D. M., Freeborn, B. R., Kjeldgaard, M., Langer, J. A., Ramakrishnan, V., Schindler, D. G., Schneider, D. K., Schoenborn, B. P., Sillers, I. Y., Yabuki, S. & Moore, P. B. (1987). *Science*, **238**, 1403–1406.
Cardoso, M. B., Smolensky, D., Heller, W. T. & O'Neill, H. (2009). *J. Phys. Chem. B*, **113**, 16377–16383.
Chamberlain, D., Keeley, A., Aslam, M., Arenas-Licea, J., Brown, T., Tsaneva, I. R. & Perkins, S. J. (1998). *J. Mol. Biol.* **284**, 385–400.
Chauvin, C., Witz, J. & Jacrot, B. (1978). *J. Mol. Biol.* **124**, 641–651.
Comoletti, D., Grishaev, A., Whitten, A. E., Tsigelny, I., Taylor, P. & Trewhella, J. (2007). *Structure*, **15**, 693–705.
Cusack, S., Miller, A., Krijgsman, P. C. J. & Mellema, J. E. (1981). *J. Mol. Biol.* **145**, 525–543.
Cusack, S., Ruigrok, R. W. H., Krygsman, P. C. J. & Mellema, J. E. (1985). *J. Mol. Biol.* **186**, 565–582.
Debye, P. & Bueche, A. (1949). *J. Appl. Phys.* **20**, 518–525.
Fujiwara, S., Kull, F. J., Sablin, E. P., Stone, D. B. & Mendelson, R. A. (1995). *Biophys. J.* **69**, 1563–1568.
Guinier, A. & Fournet, G. (1955). *Small-Angle Scattering of X-rays.* New York: Wiley.
Heller, W. T., Abusamhadneh, E., Finley, N., Rosevear, P. R. & Trewhella, J. (2002). *Biochemistry*, **41**, 15654–15663.
Heller, W. T., Finley, N. L., Dong, W.-J., Timmins, P., Cheung, H. C., Rosevear, P. R. & Trewhella, J. (2003). *Biochemistry*, **42**, 7790–7800.

Heller, W. T., Vigil, D., Brown, S., Blumenthal, D. K., Taylor, S. S. & Trewhella, J. (2004). *J. Biol. Chem.* **279**, 19084–19090.
Ho, D. L., Byrnes, W. M., Ma, W. P., Shi, Y., Callaway, D. J. E. & Bu, Z. M. (2004). *J. Biol. Chem.* **279**, 39146–39154.
Ibel, K. & Stuhrmann, H. B. (1975). *J. Mol. Biol.* **93**, 255–265.
Inoue, H. & Timmins, P. A. (1985). *Virology*, **147**, 214–216.
Jacrot, B. (1976). *Rep. Prog. Phys.* **39**, 911–953.
Jacrot, B., Chauvin, C. & Witz, J. (1977). *Nature (London)*, **266**, 417–421.
King, W. A., Stone, D. B., Timmins, P. A., Narayanan, T., von Brasch, A. A. M., Mendelson, R. A. & Curmi, P. M. G. (2005). *J. Mol. Biol.* **345**, 797–815.
Koester, L., Rauch, H. & Seymann, E. (1991). *At. Data Nucl. Data Tables*, **49**, 65–120.
Krueger, J. K., Gallagher, S. C., Wang, C.-L. A. & Trewhella, J. (2000). *Biochemistry*, **39**, 3979–3987.
Krueger, J. K., Olah, G. A., Rokop, S. E., Zhi, G., Stull, J. T. & Trewhella, J. (1997). *Biochemistry*, **36**, 6017–6023.
Krueger, J. K., Zhi, G., Stull, J. T. & Trewhella, J. (1998). *Biochemistry*, **37**, 13997–14004.
Krueger, S., Gregurick, S. K., Zondlo, J. & Eisenstein, E. (2003). *J. Struct. Biol.* **141**, 240–258.
Kruse, J., Timmins, P. A. & Witz, J. (1982). *Virology*, **119**, 42–50.
Kuzmanovic, D. A., Elashvili, I., Wick, C., O'Connell, C. & Krueger, S. (2003). *Structure*, **11**, 1339–1348.
Langer, J. A., Engelman, D. M. & Moore, P. B. (1978). *J. Mol. Biol.* **119**, 463–485.
Li, J. Q., Callaway, D. J. E. & Bu, Z. M. (2009). *J. Mol. Biol.* **392**, 166–180.
Luo, G. M., Zhang, Q., Del Castillo, A. R., Urban, V. & O'Neill, H. (2009). *ACS Appl. Mater. Interfaces*, **1**, 2262–2268.
Olah, G. A., Rokop, S. E., Wang, C.-L. A., Blechner, S. L. & Trewhella, J. (1994). *Biochemistry*, **33**, 8233–8239.
Olah, G. A. & Trewhella, J. (1994). *Biochemistry*, **33**, 12800–12806.
Pozzo, D. C. (2009). *Langmuir*, **25**, 1558–1565.
Putnam, C. D., Hammel, M., Hura, G. L. & Tainer, J. A. (2007). *Q. Rev. Biophys.* **40**, 191–285.
Ramakrishnan, V., Capel, M., Kjeldgaard, M., Engelman, D. M. & Moore, P. B. (1984). *J. Mol. Biol.* **174**, 265–284.
Ramakrishnan, V. R., Yabuki, S., Sillers, I. Y., Schindler, D. G., Engelman, D. M. & Moore, P. B. (1981). *J. Mol. Biol.* **153**, 739–760.
Sato, M., Satomiyamoto, Y., Kameyama, K., Ishikawa, N., Imai, M., Ito, Y. & Takagi, T. (1995). *J. Biochem. (Tokyo)*, **118**, 1297–1302.
Stone, D. B., Timmins, P. A., Schneider, D. K., Krylova, I., Ramos, C. H. I., Reinach, F. C. & Mendelson, R. A. (1998). *J. Mol. Biol.* **281**, 689–704.
Sugio, S., Kashima, A., Mochizuki, S., Noda, M. & Kobayashi, K. (1999). *Protein Eng.* **12**, 439–446.
Svergun, D. I., Burkhardt, N., Pedersen, J. S., Koch, M. H. J., Volkov, V. V., Kozin, M. B., Meerwink, W., Stuhrmann, H. B., Diedrich, G. & Nierhaus, K. H. (1997*a*). *J. Mol. Biol.* **271**, 588–601.
Svergun, D. I., Burkhardt, N., Pedersen, J. S., Koch, M. H. J., Volkov, V. V., Kozin, M. B., Meerwink, W., Stuhrmann, H. B., Diedrich, G. & Nierhaus, K. H. (1997*b*). *J. Mol. Biol.* **271**, 602–618.
Svergun, D. I. & Koch, M. H. J. (2003). *Rep. Prog. Phys.* **66**, 1735–1782.
Svergun, D. I., Koch, M. H. J., Pedersen, J. S. & Serdyuk, I. N. (1994). *J. Mol. Biol.* **240**, 78–86.
Svergun, D. I., Koch, M. H. J. & Serdyuk, I. N. (1994). *J. Mol. Biol.* **240**, 66–77.
Svergun, D. I. & Nierhaus, K. H. (2000). *J. Biol. Chem.* **275**, 14432–14439.
Svergun, D. I., Pedersen, J. S., Serdyuk, I. N. & Koch, M. H. J. (1994). *Proc. Natl Acad. Sci. USA*, **91**, 11826–11830.
Tjioe, E. & Heller, W. T. (2007). *J. Appl. Cryst.* **40**, 782–785.
Whitten, A. E., Cai, S. & Trewhella, J. (2008). *J. Appl. Cryst.* **41**, 222–226.
Whitten, A. E., Jacques, D. A., Hammouda, B., Hanley, T., King, G. F., Guss, J. M., Trewhella, J. & Langley, D. B. (2007). *J. Mol. Biol.* **368**, 407–420.
Whitten, A. E., Jeffries, C. M., Harris, S. P. & Trewhella, J. (2008). *Proc. Natl Acad. Sci. USA*, **105**, 18360–18365.
Zhao, J. K., Hoye, E., Boylan, S., Walsh, D. A. & Trewhella, J. (1998). *J. Biol. Chem.* **273**, 30448–30459.
Zhao, J. K., Wang, J. T., Chen, D. J., Peterson, S. R. & Trewhella, J. (1999). *Biochemistry*, **38**, 2152–2159.
Zimmer, J., Doyle, D. A. & Grossmann, J. G. (2006). *Biophys. J.* **90**, 1752–1766.