# DENFERT v. 2 : Extension of ab initio structural modeling of hydrated biomolecules in the case of Small Angle Neutron Scattering data.

ALEXANDROS KOUTSIOUBAS,$^{a}$* SEBASTIAN JAKSCH$^{a}$ AND JAVIER PÉREZ $^{b}$*

$^{a}$Jülich Centre for Neutron Science (JCNS), Forschungszentrum Jülich GmbH,

Outstation at MLZ, Lichtenbergstrasse 1, 73 D-85747 Garching Germany, and

$^{b}$Synchrotron Soleil, Beamline SWING, Saint Aubin BP48, F-91192 Gif sur Yvette

Cedex France. E-mail: a.koutsioumpas@fz-juelich.de,

javier.perez@synchrotron-soleil.fr

## Abstract

Following the introduction of the program DENFERT [Koutsioubas & Pérez, J. Appl. Cryst. (2013). 46, 1884] that takes into account the hydration layer around solvated biological molecules during ab initio restorations of low-resolution molecular envelopes from Small Angle X-ray Scattering data, in the present work we introduce the second version of the program that provides the ability to treat Neutron Scattering data sets. By considering a fully interconnected and hydrated model during the entire minimization process we are able to simplify user input and reach more objective shape reconstructions. Additionally a new method is implemented for the subtraction of the contribution of internal inhomogeneities of bio-molecules to the measured scattering. Validation of the overall approach is performed by successfully recovering the shape of various protein molecules from experimental neutron and x-ray solution scattering data.

## 1. Introduction

Small Angle Scattering of Neutrons (SANS) and x-rays (SAXS) is gaining increasing attention from the structural biology community due to the valuable complementary information that these methods can provide concerning the low resolution structure of biological molecules and their complexes in solution (Putnam *et al.*, 2007; Koch *et al.*, 2003). Except from the relative ease of sample preparation and the ability to work in nearly physiological conditions, the popularity of this class of experimental techniques has been boosted in late years by the availability of dedicated instruments on reactor, spallation and synchrotron facilities, and by the development of data analysis tools that maximize the amount of information that can be obtained. Among these tools, the so called ab-initio bead models pioneered by Chacón (Chacon *et al.*, 1998) and Svergun (Svergun, 1999; Franke & Svergun, 2009), treat proteins and nucleic acids as particles of uniform contrast relative to the solvent and represent them as an interconnected and compact collection of dummy-atoms (beads). By the use of proper minimisation algorithms, bead arrangements that better reproduce the experimentally acquired small-angle scattering curves are found, and serve as low-resolution approximations of biomacromolecular shape.

However consideration of bio-molecules as particles of constant average contrast does not take into account the inherent hydration layer around the molecule, which typically consists of a thin layer (around $3\text{Å}$) of slightly increased contrast ($\approx 10\%$) relative to the bulk solvent, due to the special packing of water molecules near the outer surface of a protein or nucleic acid (Svergun *et al.*, 1998). In a recent publication (Koutsioubas & Pérez, 2013) we have proposed a solution to this problem by introducing a two-phase bead model where in the final solution, beads that represent the bio-molecule are covered by a second type of beads that represent the hydration layer.

We have shown that the treatment of SAXS protein data with the developed ab ini-

tio algorithm (implemented in the program DENFERT (Koutsioubas & Pérez, 2013)) leads to objective shape reconstructions having the correct volume of the molecular envelope, since the method accounts for the additional scattering contribution from the hydration layer. This extra contribution to the scattering for x-rays is a direct consequence of the positive contrast of biomolecules in aqueous buffers. On the other hand for heavy water ($D_2O$) buffers commonly used in SANS solution studies, biological molecules are characterized by a negative contrast relative to the solvent, so that the presence of a denser border layer leads to an apparent shrinking of the molecules as seen by SANS.

The aim of the present work is to generalize the two-phase bead method for both kinds of radiation used in small angle experiments by properly taking into account hydration effects and also to address some shortcomings of the previous algorithmic implementation. In the following we briefly revise the main aspects of the model and also describe the new developments. Then we show how these can be used for ab-initio shape reconstructions of proteins from SANS and SAXS experimental data. Finally a thorough comparison with the widely used ab initio programs DAMMIN and DAMMIF is performed.

## 2. Materials and methods

### 2.1. Model and algorithm

As in the original implementation of the algorithm (Koutsioubas & Pérez, 2013), here the model consists of two types of beads (biomolecule "M" and hydration "H" dummy atoms) of electron/scattering length density contrast $\delta\rho_i$, placed on a face-centered cubic grid of lattice constant $L$, so that the packing radius $r$ of the beads is equal to $L\sqrt{2}/4$. The scattering curve of each bead configuration is calculated using the Debye equation with the form factor $f_i(q)$ for each bead given by the relation

$f_i(q) = V_i \delta \rho_i$ where $V_i = L^3/4$.

A simulated annealing cooling scheme is used in order to minimize the discrepancy between the scattering curve corresponding to the bead arrangement and the experimentally acquired data. During the annealing procedure random beads are selected and trial modifications are attempted as long as the connectivity of the model is not violated. Acceptance of trial modifications depends on the variation of the score function $f(X)$ and also to the value of the annealing temperature T. Moves that reduce the score function are always accepted while moves that increase the score function are accepted with a probability $e^{-\Delta f/T}$. This score function contains only two contributions a) from the $q^2$-weighted discrepancy ($R(I, X)$ between the scaled scattering intensity of the bead configuration (X) and the experimental scattering intensity ($I$) and b) from the compactness term (Svergun, 1999; Koutsioubas & Pérez, 2013)that ensures the dense packing in the final solution.

Instead of introducing hydration beads at a later stage of the annealing procedure (Koutsioubas & Pérez, 2013), the initial configuration of $M$-beads is covered by $H$-beads so that no empty lattice sites that represent the solvent are neighbouring with $M$-beads. Then at each bead modification trial a local search is performed in order to check if $H$-beads should be added or removed from the model and the associated change in the score function is calculated. In this way during the whole annealing procedure complete coverage of the $M$-beads by a monolayer of $H$-beads is ensured. The contrast of the hydration layer beads is fixed depending on the packing radius $r$, so that the product $2r\delta\rho_{hydration}$ is always equal to the product of thickness and contrast of the actual layer.

We note that the choice to have an interconnected $M$-bead model that is always covered by $H$-beads, requires the definition of a simple score function that contains only one parameter while at the same time the correct hydration of the model is

taken into account during the whole annealing procedure at the expense of relatively increased computational cost, due to the higher overall number of beads present in the model and the need to vary multiple beads during a single trial move.

As in every dummy-atom ab initio approach, our model considers the biomolecule as a particle with no internal structure (i.e. uniform electron/scattering length density since all $M$-beads are identical). Such a homogeneous particle presents an asymptotic trend of its scattering intensity at high-$q$ given by the Porod-law ($I(q) \approx c_4 q^{-4}$) . However in reality, compact biological molecules are characterised by internal electron/scattering length density inhomogeneities of statistical nature. By inspecting a large number of small to average size globular protein crystallographic structures we have found that the characteristic values of size and amplitude of these inhomogeneities are given by $\delta d = 6 \mathring{A}$ and $\delta \rho_f = \pm 0.5 \rho_M$ respectively. The contribution of these fluctuations affects the scattering curve at higher $q$ values ($q > 0.25 \mathring{A}^{-1}$) leading to deviations from pure "shape" scattering that manifest themselves as non-Porod behavior of the scattering curve at high-$q$ (see example for lysozyme protein in fig. 1). As long as the size of the particle is larger than the characteristic size of the inhomogeneities, the high-$q$ scattering behaviour is expected to be approximated by the equation (Luzzati *et al.*, 1961; Feigin & Svergun, 1987)

$$I(q) \approx c_0 + c_4 q^{-4} \tag{1}$$

where $c_0$ is a constant representing the inhomogeneity contribution to the scattering.

Following the methodology firstly introduced in in the program DAMMIN (Svergun, 1999), in our previous work (Koutsioubas & Pérez, 2013) we took this effect into account by subtracting an appropriate constant from the scattering data so that Porod behavior is recovered at high-$q$ before performing ab-initio reconstructions. This constant was estimated by fitting the high-$q$ region of the data with equation 1. However

this procedure depends on the range of the acquired experimental data. In the case where no high-$q$ data are available, the constant estimation might become troublesome.

In the present version of the algorithm we have applied a more objective method for the estimation of the internal inhomogeneity contribution. By assuming that the inhomogeneity contribution to the overall scattering is additive to the shape scattering term (Svergun, 1994), we firstly make a very rough estimation of the molecular shape through a fast simulated annealing procedure (or by fitting the experimental curve by assuming a triaxial ellipsoidal molecular envelope in the case of large molecules) and then we calculate the scattering contribution of statistical inhomogeneities of characteristic size $\delta d$ and amplitude $\delta \rho_f$ within the found molecular shape. The described calculation is performed using the Debye equation on a model where the found molecular envelope is randomly filled with dummy atoms of $6\mathring{A}$ diameter and $\delta \rho_f = +0.5 \rho_M$ or $\delta \rho_f = -0.5 \rho_M$. Note that this calculation leads always to a pseudo-plateau behaviour at high-q (fig. 1). The value corresponding to this plateau represents the constant [1] that is subtracted from the experimental curve in order to obtain an internal inhomogeneity-free "shape"-scattering curve that is further used for recovering the molecular shape.

*2.2. Protein solution preparation*

Proteins hen egg Lysozyme and bovine $\beta$ - lactoglubulin were purchased in lyophilised powder form from Sigma Aldrich. Protein powders were dissolved and dialysed two times against appropriate $D2O$ buffers ($40 mM$ acetic acid, $50 nM$ $NaCl$, $pD = 4.0$ for Lysozyme and $50 mM$ Tris, 100mM $NaCl$, $pD = 7.5$ for beta-lactoglobulin) for removing eventual conservative products and for performing $H/D$ exchange of labile protons.

---

[1] The value of this constant was found to be close to the one estimated by programs DAMMIN and DAMMIF, when reliable high-$q$ experimental data are provided as input to DAMMIN and DAMMIF.

Each dialysis was performed in $250mL$ buffer, using $3mL$ volume dialysis cassettes with $3kDa$ MWCO membrane pores, purchased from Pierce Company. The solute concentration after dialysis was determined by the absorption of protein solutions at 280nm using a Nanodrop spectrophotometer, while the extinction coefficients of the two proteins were estimated using ProtParam online server (Gasteiger *et al.*, 2005).

### 2.3. Small-angle neutron scattering experiments

Solution Small Angle Neutron Scattering (SANS) data from Lysozyme and $\beta$-Lactoglobulin proteins were collected at the KWS-1 beamline ,JCNS outstation at MLZ (Garching, Germany) (Feoktystov *et al.*, 2015). Neutron wavelength $\lambda$ was set equal to $4.5\mathring{A}$ (with $\Delta\lambda/\lambda \sim 10\%$) and two different sample-detector distances were used ($L = 1.5\&8m$), covering a range of scattering vector magnitude $0.01 < q < 0.45\mathring{A}^{-1}$ ($q = 4\pi sin\theta/\lambda$), where $2\theta$ is the scattering angle). All solutions were placed in $1mm$ thick quartz cuvettes (Helma). Scattering of the dialysis buffers was also recorded and subtracted from the corresponding protein solution scattering curves. Additional test SANS datasets of proteins dissolved in appropriate buffers were collected at the MARIA beamline at JCNS outstation at MLZ (Garching, Germany).

### 2.4. Computer program

The presented algorithm is implemented in the second version of the program DEN-FERT written in Fortran 90. As in the original version, the program processes the output of program GNOM (Svergun, 1992) that provides a regularised scattering curve, together with an estimation of the maximum particle diameter $D_{max}$. Instrumental resolution of the SANS instrument (wavelength spread, beam divergence) is provided as extra input to GNOM. A unified executable can treat both types of used radiation (neutrons or x-rays) and a typical run in "fast mode" requires about an hour on the

single core of modern processors.

## 3. Results and Discussion

In order to evaluate the stability and effectiveness of the generalized and updated algorithm, we have executed a series of shape reconstructions from "simulated" scattering curves that were calculated from crystallographic structures of proteins using the programs CRYSOL and CRYSON (Svergun *et al.*, 1995) and also using test SANS datasets of protein solutions measured at the MARIA instrument. These initial tests (results not shown) established the robustness of the algorithm and it's success in the accurate recovery of bio-molecular shape from the scattering data. However final validation of the algorithm was based on shape reconstructions from experimental SANS data (KWS-1 instrument) of two proteins of known crystallographic structure (Lysozyme and $\beta$-lactoglobulin). Additionally SAXS experimental data collected in our previous work (Koutsioubas & Pérez, 2013) were retreated together with a set of five "standard protein" entries found in the SASBDB database (Valentini *et al.*, 2014). A summary of the results is presented in tables 1 and 2.

Initially let us discuss the results based on SANS data. Both scattering curves for Lysozyme and $\beta$-lactoglobulin were treated by assuming a hydration layer of 10% contrast and $3\mathring{A}$ thickness, and a scattering length density of the protein [2] equal to $3.11 \times 10^{-6} \mathring{A}^{-2}$. As discussed above, when $D_2O$ buffers are used ($\rho_{D_2O} = 6.4 \times 10^{-6} \mathring{A}^{-2}$), the relative contrast of a biomolecule (protein or nucleic acid) is negative. That means that a more dense hydration layer will produce an "artificial shrinking" of the particle as seen by SANS, which is the opposite effect than what is observed in a SAXS experiment. For example from Guinier fits (in the range $qR_g < 1.25$) of the

---

[2] Note that due to the $H/D$ exchange process occurring in solvents containing a certain ratio of $D_2O$, the scattering length density of proteins and nucleic acids is a function of $D_2O$ ratio (see (Jacrot, 1976) for details) and should be adapted accordingly.

present experimental results (fig. 2 and (Koutsioubas & Pérez, 2013)) for Lysozyme, the radius of gyration found by neutrons and x-rays is quite different $R_{g\_SAXS} = 15.0\mathring{A}$, $R_{g\_SANS} = 13.2\mathring{A}$. When the hydration layer contribution is not taken into account (as in the case of shape reconstruction with DAMMIN) this same shrinking or swelling effect is reflected on the total volume of the molecular envelope of the final model.

As seen in figure 3, if we ignore the hydration layer contribution in SANS reconstructions, the final model is characterised by an apparent "shrinking" of the molecular envelope. For the two studied proteins here the observed shrinking of the envelope volume ($V$) is of the order of $15 - 20\%$. This effect is expected to be more pronounced especially for "small" proteins (with molecular weight smaller than $50kDa$) since the hydration layer scattering contribution is comparable to the scattering from the bulk of the protein. Even though these effects do not largely alter the structural similarity of the final models (see table 1), the accurate recovery of the molecular envelope volume might prove quite valuable in cases where it is further used as a placeholder for evaluating folding prediction models or for constraining molecular simulations.

We note that the use of a scattering length density contrast equal to 10% and thickness of $3\mathring{A}$(identical to what we used previously for treating SAXS data) in the SANS ab initio reconstructions with DENFERT appears to compensate adequately for the hydration layer scattering contribution. This is also further confirmed by treating a set of "standard-protein" solution x-ray scattering data that are deposited in SAS-BDB (see table 2 and fig. 4) where with DENFERT v.2 we are, in all cases able to recover accurate values concerning the molecular envelope volume. In this respect the mentioned set of values for hydration layer contrast and thickness should always lead to reliable reconstructions at least in the case of protein structure reconstructions.

Comparing to the previous version of the program, in the new implementation we

begin with a compact arrangement of beads that is fully covered by a single layer of hydration beads. In every single iteration of the simulated annealing algorithm the interconnectivity and full coverage by hydration beads is respected. That means that no externally defined parameters related to connectivity and hydration are needed and only a single parameter related to the overall compactness of the final model needs to be defined. By performing extensive tests we have related the value of the compactness parameter with the packing radius ($r$) of the beads and the estimated size of the final model, and this value is proposed by default to the user.

By performing ab initio reconstructions of SAXS data with the new algorithm, we generally observed lower NSD values between the recovered protein envelopes and the crystallographic structures (comparing to DAMMIN and DAMMIF programs, the improvement on the normalised spatial discrepancy (NSD) values may go up to 10%, see table 2) and also a better stability of the final solution (smaller variation of the NSD values). This improvement over the previous version has mainly to do with the smoother approach towards the final solution, since in the current version hydration beads are not introduced artificially at a later stage of the minimization procedure. Of course this approach involves a much higher number of dummy atoms during the entire annealing procedure, thus making the algorithm more computationally demanding. This fact makes program execution relatively slower than other programs like DALAI_GA (Chacon *et al.*, 1998), DAMMIN (Svergun, 1999) and especially DAMMIF (Franke & Svergun, 2009), a price that has to be payed for obtaining the maximum amount of information from small-angle scattering curves with a dummy-atom model. The characteristic time for a DENFERT reconstruction on a single core of a modern processor is expected to be in the range of an hour.

In general we expect that most usage cases will involve treatment of x-ray data sets, since SAXS instruments are characterised by better resolution, small counting times

and smaller required sample volumes, and thus are more popular in the structural biology community, while the main advantage of SANS is related to the ability to vary the contrast in multi-component systems like biomolecular complexes. In cases where one of the components is contrast matched, the presented algorithm may provide reliable low-resolution information about biomolecular structure from a neutron scattering data set.

## 4. Conclusions

In summary, we have presented the improved second version of the ab initio low-resolution structure recovery program DENFERT that includes the additional ability to treat SANS data sets. We have shown that the hydration layer contribution can be taken into account in an identical way as in the SAXS case. User input has been simplified while algorithmic and methodological improvements over the previous version lead to more reliable final models that are representative of protein or nucleic acid volume and shape in solution. The executable code of the program DENFERT v.2 is available for all major platforms (Mac, Linux, Windows) from http://www.synchrotron-soleil.fr/Recherche/LignesLumiere/SWING/denfert

### References

Brownlow, S., Cabral, J. H. M., Cooper, R., Flower, D. R., Yewdall, S. J., Polikarpov, I., North, A. C. & Sawyer, L. (1997). *Structure*, **5**(4), 481 – 495.

Chacon, P., Morn, F., Daz, J., Pantos, E. & Andreu, J. (1998). *Biophysical Journal*, **74**(6), 2760 – 2775.

Diamond, R. (1974). *Journal of Molecular Biology*, **82**(3), 371 – 391.

Feigin, L. A. & Svergun, D. I. (1987). *Structure Analysis by Small-Angle X-Ray and Neutron Scattering*. Springer Science+Business Media New York.

Feoktystov, A. V., Frielinghaus, H., Di, Z., Jaksch, S., Pipich, V., Appavou, M.-S., Babcock, E., Hanslik, R., Engels, R., Kemmerling, G., Kleines, H., Ioffe, A., Richter, D. & Brückel, T. (2015). *Journal of Applied Crystallography*, **48**(1), 61–70.

Franke, D. & Svergun, D. I. (2009). *Journal of Applied Crystallography*, **42**(2), 342–346.

Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M., Appel, R. & Bairoch, A. (2005). *The Proteomics Protocols Handbook*. Humana Press.

Jacrot, B. (1976). *Reports on Progress in Physics*, **39**(10), 911.

Koch, M. H. J., Vachette, P. & Svergun, D. I. (2003). *Quarterly Reviews of Biophysics*, **36**, 147–227.

Koutsioubas, A. & Pérez, J. (2013). *Journal of Applied Crystallography*, **46**(6), 1884–1888.

Kozin, M. B. & Svergun, D. I. (2001). *Journal of Applied Crystallography*, **34**(1), 33–41.

Luzzati, V., Witz, J. & Nicolaieff, A. (1961). *Journal of Molecular Biology*, **3**(4), 367 – 378.

Putnam, C. D., Hammel, M., Hura, G. L. & Tainer, J. A. (2007). *Quarterly Reviews of Biophysics*, **40**, 191–285.

Svergun, D. (1999). *Biophysical Journal*, **76**(6), 2879 – 2886.

Svergun, D., Barberato, C. & Koch, M. H. J. (1995). *Journal of Applied Crystallography*, **28**(6), 768–773.

Svergun, D. I. (1992). *Journal of Applied Crystallography*, **25**(4), 495–503.

Svergun, D. I. (1994). *Acta Crystallographica Section A*, **50**(3), 391–402.

Svergun, D. I., Richard, S., Koch, M. H. J., Sayers, Z., Kuprin, S. & Zaccai, G. (1998). *Proceedings of the National Academy of Sciences*, **95**(5), 2267–2272.

Valentini, E., Kikhney, A. G., Previtali, G., Jeffries, C. M. & Svergun, D. I. (2014). *Nucleic Acids Research*.

Table 1. *Protein structures recovered from experimental scattering data using DENFERT v.2. Results of molecular envelope volume (V) and normarized spatial discrepancy (NSD) represent the mean values from ten independent shape reconstructions, using a hydration layer of $3\text{Å}$ thickness and $10\%$ contrast relative to the solvent. Corresponding values of reconstructions using the program DAMMIN v. 5.3 (Svergun, 1999) are given in parentheses. Note that the default subtraction constants for the recovery of "shape" scattering curves are used with both programs. \* SANS data. \*\* SAXS data (Koutsioubas & Pérez, 2013). † Molecular envelope volume calculated using the program CRYSOL (Svergun* et al.*, 1995) for the corresponding crystallographic structure. ‡ NSD between the ab initio model and the crystallographic structure calculated using the program SUPCOMB20 (version 2.0 (Kozin & Svergun, 2001)).*

| Protein | $D_{max}(\text{Å})$ | $V_{en}^{\dagger}(10^3\text{Å}^3)$ | $r(\text{Å})$ | $V(10^3\text{Å}^3)$ | $NSD^{\ddagger}$ |
|---|---|---|---|---|---|
| Lysozyme* | 36 | 18.3 | 1.5 | 20.1(15.2) | 0.72(0.75) |
| $\beta$-lactoglobulin* | 65 | 50.7 | 2.0 | 52.6(40.9) | 0.84(0.84) |
| Lysozyme** | 50 | 18.3 | 1.5 | 19.0(20.7) | 0.70(0.79) |
| HSA** | 80 | 95.1 | 2.5 | 95.0(107) | 1.07(1.18) |
| Actin** | 75 | 60.1 | 2.5 | 56.5(64.8) | 0.93(0.93) |

Table 2. *Comparison of 5 different ab-initio protein structure reconstructions performed with DENFERT and DAMMIF (Franke & Svergun, 2009) for SAXS experimental data deposited in SASBDB (Valentini et al., 2014) (www.sasbdb.org). DENFERT reconstructions were performed in slow mode and by assuming a hydration layer of 3Å thickness and 10% contrast relative to the solvent. DAMMIF results (presented in parentheses) were those found in the SASBDB entries. † Molecular envelope volume calculated using the program CRYSOL (Svergun et al., 1995) for the corresponding crystallographic structure. ‡ NSD between the ab initio model and the crystallographic structure calculated using the program SUPCOMB20 (version 2.0 (Kozin & Svergun, 2001)).*

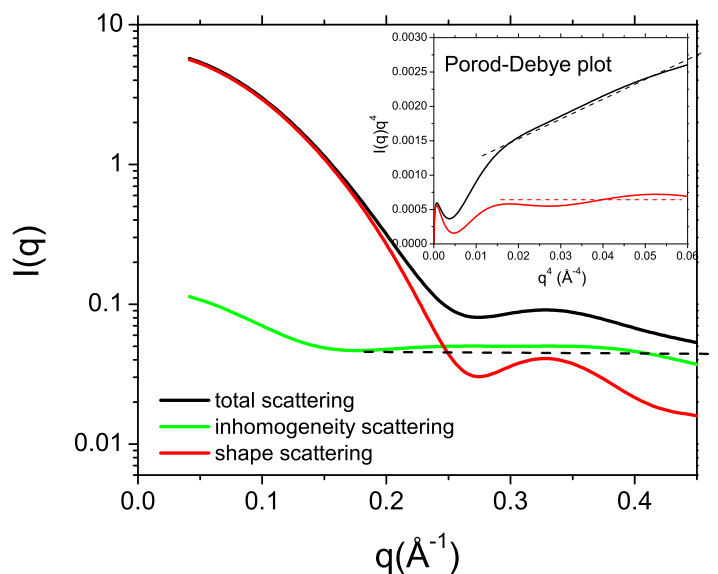| SASBDB entry | Protein | $D_{max}(\mathring{A})$ | $V_{en}^{\dagger}(10^3\mathring{A}^3)$ | $V(10^3\mathring{A}^3)$ | $NSD^{\ddagger}$ |
|---|---|---|---|---|---|
| SASDAA2 | conalbumin | 110 | 111 | 109(114) | 0.83(0.83) |
| SASDAQ2 | ubiquitin | 49 | 11.3 | 12.2(16.0) | 0.87(0.99) |
| SASDAB2 | cytochrome-c | 35 | 15.8 | 13.5(14.7) | 0.75(0.88) |
| SASDAH2 | myoglobin | 50 | 23.2 | 21.9(30.3) | 0.97(1.13) |
| SASDA32 | serum albumin | 93 | 94.7 | 91.5(130) | 0.72(0.84) |

Fig. 1. Example of internal inhomogeneity contribution (x-ray case) for protein
lysozyme. (black curve) SAXS curve lysozyme, (green curve) scattering due to inter-
nal inhomogeneities of size $\delta d = 6\mathring{A}$ and amplitude $\delta\rho_f = \pm 0.5\rho_M$, where $\rho_M$ the
average electron density of the protein, (red curve) lysozyme "shape" curve as deter-
mined by subtracting a constant (dashed line) from the the lysozyme SAXS curve.
Note that the scattering due to inhomogeneities reaches a plateau for high-$q$ and
that is also the region where it becomes a significant part of the total scattering.
In the inset the total and "shape" scattering curves are plotted in a Porod-Debye
representation $I(q)q^4$ vs $q^4$. As expected after an initial modulation, the "shape"
scattering curve fluctuates around a constant value (Porod behaviour), while the
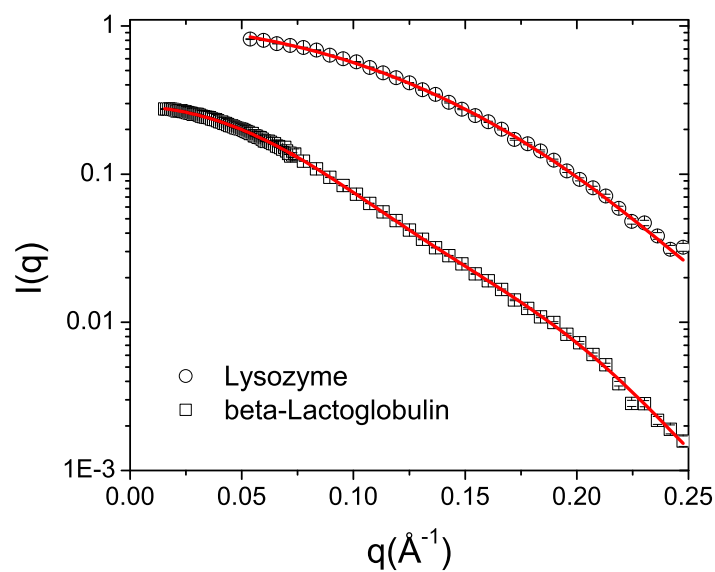total scattering curve increases linearly as expected from eq. 1.

Fig. 2. Experimental SANS data for lysozyme and $\beta$-lactoglobulin proteins together with the scattering curves (red full lines) of the restored models using DENFERT v.2 and by assuming a hydration layer of 10% contrast and $3\mathring{A}$ thickness. Note that for all shape restorations we consider the scattering curves up to $q = 0.25\mathring{A}^{-1}$.
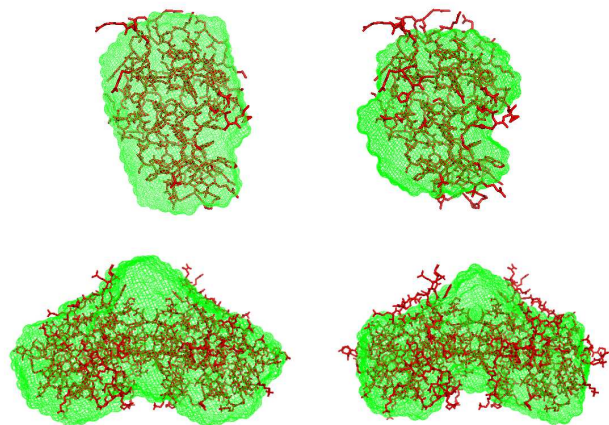
Fig. 3. Ab initio shape reconstruction of lysozyme(top) and $\beta$-lactoglobulin (bottom) from SANS data using DENFERT v.2 (left) and DAMMIN (right). The crystallographic structures of the proteins 6lyz (Diamond, 1974) and1beb (Brownlow *et al.*, 1997) are plotted in red color.
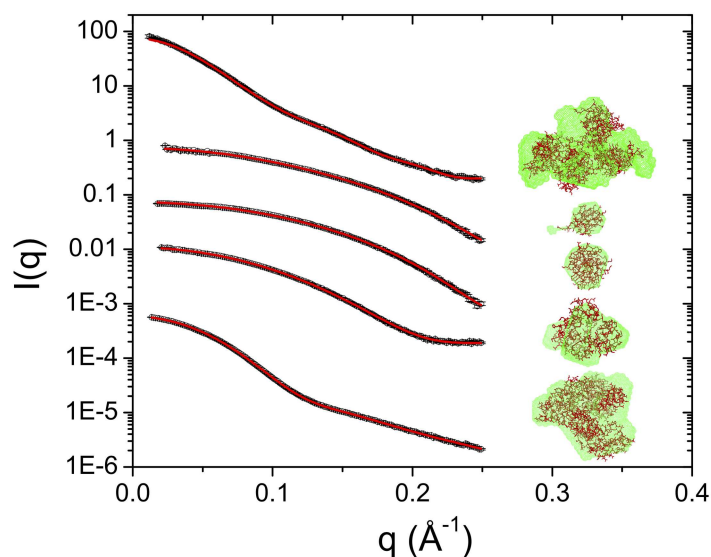
Fig. 4. Ab initio reconstructions using DENFERT v.2 (assuming a hydration layer of 10% contrast and 3Å thickness) for five SASBDB protein entries. The order from top to bottom is conalbumin, ubiquitin,cytochrome-c,myoglobin,bovine serum albumin. Crystallographic structures are plotted in red, and the found molecular envelopes in green.The obtained parameters of each reconstruction are given in table 2.

**Synopsis**

The second updated version of the program DENFERT is introduced for ab-initio bio-macromolecular shape restorations from small-angle scattering data. The two major improvements concern the ability to threat small-angle neutron scattering datasets and the proper consideration of biomolecular hydration during the entire shape restoration procedure.