Structure and dynamics of biomacromolecules in solution: recent developments and future perspectives in SANS/SAXS and neutron spectroscopy

Frank Gabel

Habilitation à Diriger des Recherches (DHDR)

Université Joseph Fourier, Grenoble

UFR de Physique

Structure and dynamics of biomacromolecules in solution: recent developments and future perspectives in SANS/SAXS and neutron spectroscopy

Dr. Frank Gabel

(Institut de Biologie Structurale, Grenoble, France)

Members of the habilitation jury:

Prof. Otto Glatter	(rapporteur)
Prof. José Teixeira	(rapporteur)
Prof. Patrice Vachette	(rapporteur)
Dr. Martin Blackledge	(examinateur)
Dr. Roland May	(examinateur)
Dr. Giuseppe Zaccai	(examinateur)
Prof. Heinrich Stuhrmann	(examinateur)

Defended on May 31st 2010 at IBS Grenoble, France

Habilitation à Diriger des Recherches (DHDR)

Université Joseph Fourier, Grenoble

UFR de Physique

Structure and dynamics of biomacromolecules in solution: recent developments and future perspectives in SANS/SAXS and neutron spectroscopy

Dr. Frank Gabel

(Institut de Biologie Structurale, Grenoble, France)

Members of the habilitation jury:

Prof. Otto Glatter	(rapporteur)
Prof. José Teixeira	(rapporteur)
Prof. Patrice Vachette	(rapporteur)
Dr. Martin Blackledge	(examinateur)
Dr. Roland May	(examinateur)
Dr. Giuseppe Zaccai	(examinateur)
Prof. Heinrich Stuhrmann	(examinateur)

Defended on May 31st 2010 at IBS Grenoble, France

Foreword

This French University Habilitation ("mémoire" to obtain the "Diplôme d'Habilitation à Diriger des Recherches", DHDR) is divided into two parts: the first one deals with the results that I have obtained after my PhD thesis in 2003, the second one discusses open questions related to these results as well as mid- and long-term perspectives.

Three different topics are presented in the result chapters: 1) a combination of small angle scattering (SAS) and nuclear magnetic resonance (NMR) for rigid-body modeling of biomacromolecular complexes, 2) the combined use of small angle X-ray (SAXS) and neutron (SANS) scattering for the study of unfolded proteins and 3) the study of biomacromolecular and solvent dynamics by neutron spectroscopy combining several instruments.

Points 1) and 3) are discussed in great detail, both in the results and perspective sections, since they represent the most advanced projects of my research. Point 2) is dealt with more briefly in the results section, since few results are available so far. However, perspectives are discussed. A special perspective chapter deals with applications on membrane proteins. Key publications for the different chapters are:

Chapter 1: Gabel *et al.* (2006) A target function for quaternary structural refinement from small angle scattering and NMR orientational restraints. *Eur. Biophys. J.* **35**(4), 313-327.
Gabel *et al.* (2008) A structure refinement protocol combining NMR residual dipolar couplings and small angle scattering restraints. *J. Biomol. NMR* **41**(4), 199-208.

Chapter 2: Gabel *et al.* (2009) Quantitative Modelfree Analysis of Urea Binding to Unfolded Ubiquitin Using a Combination of Small Angle X-ray and Neutron Scattering. *J. Am. Chem. Soc.* **131**(25), 8769-8771.

Chapter 3: Gabel (2005) Protein dynamics in solution and powder measured by incoherent elastic neutron scattering: the influence of Q-range and energy resolution. *Eur. Biophys. J.* **31**(1), 1-12.
Gabel & Bellissent-Funel (2007) C-Phycocyanin Hydration Water Dynamics in the Presence of Trehalose: An Incoherent Elastic Neutron Scattering Study at Different Energy Resolutions. *Biophys. J.* **92**(11), 4054-4063.

The habilitation thesis is focused on methodological aspects and developments of small angle scattering and neutron spectroscopy. It is obvious that the approaches discussed here and their sophisticated levels of data analysis rely fundamentally on the quality of the sample, and in particular on *monodispersity* (for SAS) and amount of material for spectroscopy. The paramount

importance of **good biochemistry** and the use of **complementary techniques** for the characterization of samples **can hardly be overestimated.** They include, amongst others, gel filtration, analytical ultracentrifugation, static and dynamic light scattering, NMR, etc. They are quite simply **indispensable** for doing **good and accurate science with SAS**, in particular in more complex systems (macromolecular complexes, membrane proteins …). If they are not presented in more detail in this thesis, it is not out of ignorance of this fact but due to the lack of space.

Acknowledgements

I would like to thank all the friends and colleagues that have supported me throughout my years in research and in particular those who have contributed to the work presented in this habilitation thesis:

- The **jury members** for taking the time to read and evaluate my work and agreeing to participate in the defense
- Prof. Dieter **Engelhardt** († 14th December 2007) for bringing me to Grenoble through the French-German exchange program between Karlsruhe and Grenoble Universities
- Dr. Giuseppe **Zaccai** (ILL Grenoble) for accepting me as a student at IBS 10 years ago and for his uninterrupted inspiration and (scientific and moral) support since that time
- Dr. Michael **Sattler** (TU Munich) for accepting me as a Postdoc in his lab at EMBL Heidelberg and for believing in the combination of SAS and the "NMR world"
- Dr. Bernd **Simon** (EMBL Heidelberg) for his patience in explaining NMR practical and theoretical aspects to me
- Dr. Marie-Claire **Bellissent-Funel** (LLB Saclay) for her collaboration in the CPC project
- Dr. Martin **Blackledge** (IBS Grenoble) and Dr. Malene **Ringkjøbing Jensen** (IBS Grenoble) for their continuing enthusiasm in the collaboration on unfolded proteins
- Dr. Bruno **Franzetti** (IBS Grenoble) for the collaborations on large molecular assemblies, for teaching me biology and for accepting me as a member of his group
- Dr. Christine **Ebel** (IBS Grenoble) for introducing me to the world of lipids, detergents and membrane proteins
- All the other **members of the LBM** at IBS (in particular Dr. Martin **Weik**) who have supported me throughout the years
- Last but not least I would like to thank my wife **Catia** and my sons **Enrico** and **Riccardo** for their patience throughout all the years and in particular during the last months while writing this thesis. Without their understanding for my weekends, public holidays and nights spent at ILL and ESRF measuring data, this work would not have been possible

Für Enrico Leonardo und Riccardo Angelo, in Liebe

Abbreviations

Ambiguous Restraints for Iterative Assignment – Crystallography & NMR Systems (ARIA-CNS)
Analytical Ultracentrifugation (AUC)
Biological Small Angle Scattering (BioSAS)
Critical Micelle Concentration (CMC)
Contrast Match point (CMP)
C-Phycocyanin (CPC)
Diplôme d'Habilitation à Diriger des Recherches (DHDR)
Deoxyribonucleic Acid (DNA)
Elastic Intensity (EI)
Elastic Incoherent Structure Factor (EISF)
European Molecular Biology Laboratory (EMBL)
Elastic Neutron Scattering (ENS)
European Synchrotron Radiation Facility (ESRF)
Half Width at Half Maximum (HWHM)
Institut de Biologie Structurale (IBS)
Institut Laue-Langevin (ILL)
Inelastic Neutron Scattering (INS)
Laboratoire de Biophysique Moléculaire (LBM)
Laboratoire Léon-Brillouin (LLB)
Molecular Dynamics (MD)
Membrane Protein (MP)
Mean Square Displacement (MSD)
Nuclear Magnetic Resonance (NMR)
Nuclear Overhauser Effect (NOE)
Neutron Scattering (NS)
Protein Data Bank (PDB)
Protein-Detergent Complex (PDC)
Paramagnetic Relaxation enhancement (PRE)
Quasielastic Neutron Scattering (QENS)
Residual Dipolar Coupling (RDC)
Root Mean Square Displacement (RMSD)
Ribonucleic Acid (RNA)
Small Angle Scattering (SAS)
Small Angle Neutron Scattering (SANS)
Small Angle X-ray Scattering (SAXS)

Contents

Postdoctoral results

1. Accuracy in biomacromolecular rigid-body modeling: the combination of small angle scattering and nuclear magnetic resonance

1.1 Introduction

Biological small angle scattering (BioSAS) has been used for several decades to study the structural properties of biomacromolecules in solution. A complete historical review is beyond the scope of this thesis. Excellent introductions to the theory, applications and history of small angle scattering (SAS) can be found in the textbooks of Guinier and Fournet (1955), Glatter and Kratky (1982) or Feigin and Svergun (1987). Pioneering results in BioSAS have been reviewed by Kratky (1963) and Kratky and Pilz (1972). Biological small angle neutron scattering (SANS) has been specifically discussed by Stuhrmann (1974), Engelman and Moore (1975), Jacrot (1976) and Zaccaï and Jacrot (1983).

The past decade (2000-2010) has witnessed a vibrant renaissance of BioSAS, reflected in an exponentially growing number of publications (*e.g.* citations referring to "small angle scattering" in the ISI Web of Knowledge[SM] database). This is in great part due to the advent of a new generation of data modeling programs and increased computing power (Svergun and Koch 2002; Koch *et al.* 2003; Petoukhov and Svergun 2007; Putnam *et al.* 2007; Lipfert and Doniach 2007). A number of different approaches have emerged and can be distinguished conceptually (Figure 1.1):

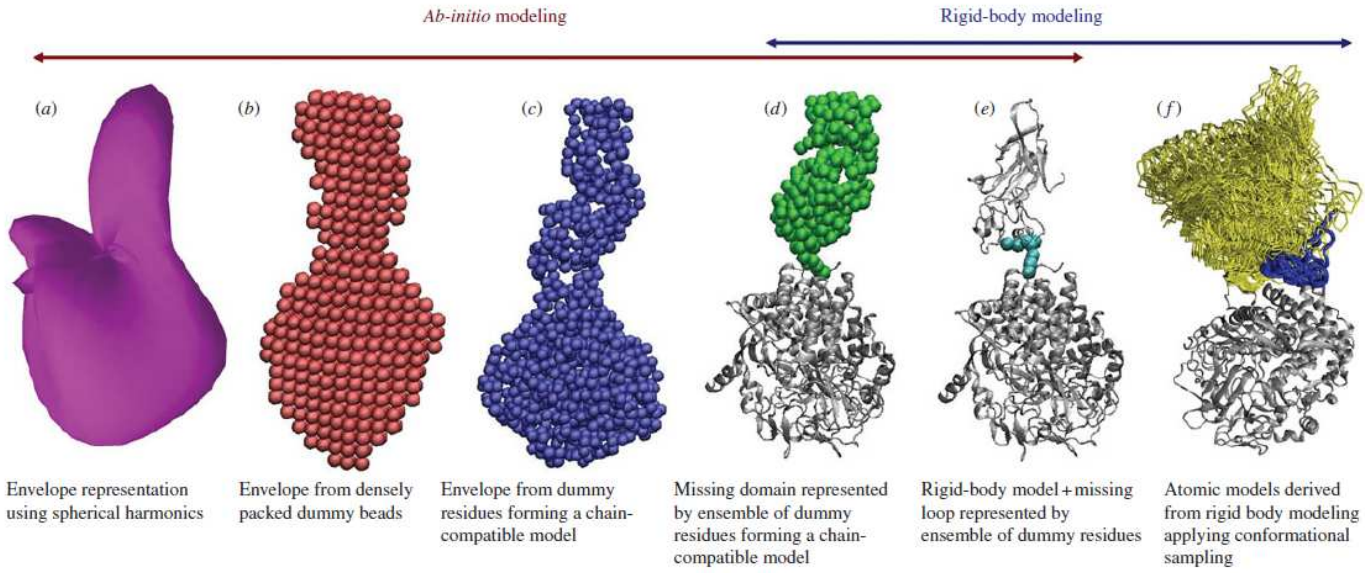

Figure 1.1: Overview of different SAS modeling approaches applied to cellulase (from Putnam *et al.* 2007).

1) ***Ab initio* modeling** can be used in cases where no *a priori* atomic resolution information of the system (or parts of it) is available. Since the early 1970s, this problem has been addressed by multipole expansions using spherical harmonics (Harrison 1969; Stuhrmann 1970a and b) to model low-resolution shapes and envelopes. It is conceptually similar to a Fourier-analysis and adapted to three-dimensional cases described by spherical coordinates. It consists in finding the coefficients of the multipole expansion for a particle shape from the experimental scattering curve. Based on experimental data regularization with *Indirect Fourier Transformation* (Glatter 1977; Svergun *et al.* 1988; Svergun 1992) a so-called "dummy atom"-approach has been developed more recently (Chacón *et al.* 1998; Svergun 1999): starting from a connected ensemble of small spherical scattering units (dummy atoms) grouped together in a volume of defined size (*e.g.* a sphere of maximum diameter D_{max}) the scattering curve can be back-calculated easily (*e.g.* Glatter 1980a) and be compared to the experimental data. Dummy atoms are removed or added in a simulated annealing process until the back-calculated scattering curve matches the experimental one best. An important requirement for the success of this method is the **connectivity** of the ensemble of dummy atoms.
2) **Rigid-body modeling** is used for macromolecular complexes of unknown overall structure made up of subunits of known high-resolution structure (by NMR or crystallography). Originally, it had been used with simple geometric bodies (spheres, ellipsoids, cylinders …) and an optimization of their respective position and orientation by an educated "trial-and-error" process (Kratky 1963; Kratky and Pilz 1972). More recently, bodies with arbitrary structures (*e.g.* protein structures from PDB) have been refined in a **grid search approach** by shifting and orienting them incrementally in a simulated annealing process until the back-calculated scattering curve matches the experimental SAS data from the complex they form (*e.g.* programs "MASSHA" and "SASREF"; Konarev *et al.* 2001; Petoukhov and Svergun 2005).
3) **"Hybrid" approaches** are used when high-resolution structures of (a) part(s) of the complex are available (programs "CREDO/CHADD" and "BUNCH"; Petoukhov *et al.* 2002; Petoukhov and Svergun 2005). The remainder of the complex (*e.g.* linkers, missing domains) is modeled by *ab initio* techniques similar to those mentioned under point 1).
4) **Ensemble averaging.** For systems built up entirely or partly of flexible regions (*e.g.* unfolded proteins), no single structure (either *ab initio* or rigid-body) can represent a structural solution in a satisfactory manner. In these cases, one can use an approach that back-calculates and scores an averaged scattering curve from an ensemble of structures against the experimental

data, ideally by taking complementary information into account, *e.g.* from NMR (Bernadó *et al.* 2005; Bernadó *et al.* 2007).

The remaining sections of this chapter focus on the rigid-body modeling approach (number 2) of the above list), and in particular present and discuss the benefits of supplementing SAS restraints by NMR restraints to increase the accuracy of the refined models.

1.2 Limitations of traditional rigid-body modeling approaches

Rigid-body modeling programs like "MASSHA" and "SASREF" optimize the respective position and orientation of two bodies in a grid-search by incremental translations and rotations in a simulated annealing process until their back-calculated scattering curve matches the experimental data best (least χ^2-fit). "GLOBSYMM" (Petoukhov and Svergun 2005) can generate an ensemble of symmetric oligomers by a brute force exhaustive search but cannot address the general case in the same way. While always producing a structural model as a result, there are a number of inherent fundamental limitations to these approaches, in particular to the grid-search:

1) **The refined structure of the complex is in general biased by the starting points** (*i.e.* the initial orientations and positions) of the two bodies. For different initial setups, different structural solutions may be proposed by the refinement process. In particular, once the system is trapped in an energetic side-minimum (a so-called "false positive") it is often not possible to retrieve the "correct" structural solution.
2) **Solutions with an equally good χ^2 value can be different**. In particular, equal matches with the experimental data (in terms of χ^2) can be obtained by deviations of the back-calculated scattering curve that lie either below or above the experimental data in several Q-ranges.
3) Another unsatisfactory point pertains to the **uniqueness and stability** of the refined structure: in other words, **is the proposed structure the only (and the best) one possible**? Equally important, what are possible **residual degrees of freedom** for the positions and orientations of the rigid bodies that are consistent with the SAS curve? That is, can one of the bodies be translated or re-oriented and still reproduce a satisfactory fit of the experimental scattering curve? And finally, how do these residual degrees of freedom (if they exist) depend on the geometry, position and orientation of the bodies, on the experimental Q-range and on error bars?

In the light of these limitations, a more satisfying way to represent structural solutions would be to display a number of structures that are consistent with the SAS curve. An example is given below (Fig. 1.2). Unfortunately, this is not consequently done.

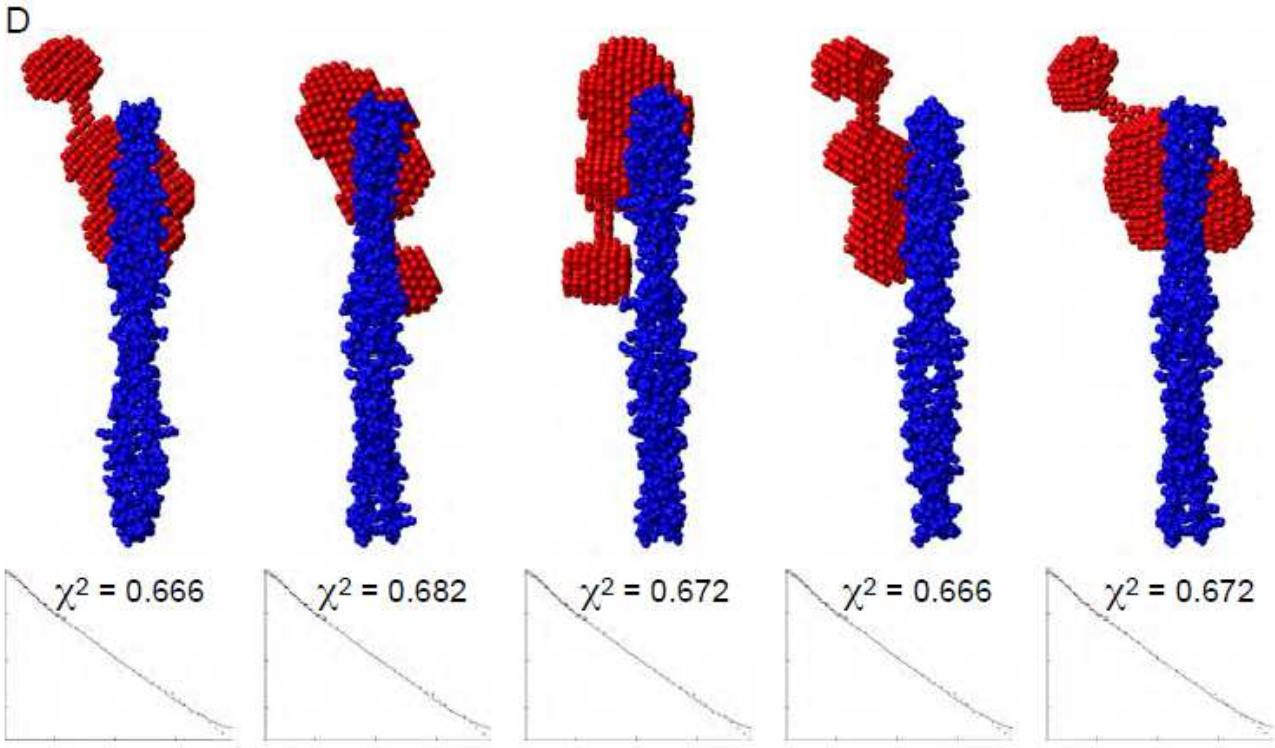

Figure 1.2: Structural models obtained for a protein-protein complex involved in the bacterial cell division of *Streptococcus pneumoniae* and the corresponding χ^2-fit values of the experimental (SANS) data (from Masson *et al.* 2009). The models visualize the residual degrees of freedom of the complex that can not be further refined by the experimental data available.

1.3 Complementary structural restraints and their incorporation into rigid-body modeling approaches

It is obvious that any kind of additional structural restraints (complementary to SAS) can help to refine a structural model and improve its accuracy. Such restraints are provided by a number of techniques, including: i) **mutagenesis** can help to identify the (surface) residues that serve as potential contact points at the interface of two particles and can be used as docking restraints (Dominguez *et al.* 2003), ii) mapping of NMR **chemical shift perturbations** (Zuiderweg 2002) can define, in a way similar to mutagenesis, patches on the surfaces of interacting partner molecules (Dominguez *et al.* 2003) and are either obtained in titration experiments of separate partners or by the depletion of a part of the sequence in a linker-connected molecule, iii) **chemical cross-linking experiments** (Back *et al.* 2003) can yield distance restraints between specific residues/atoms of macromolecules in a complex, iv) in a similar way, distance restraints can be obtained from NMR **paramagnetic relaxation enhancement** (PREs) (Battiste and Wagner 2000), v) restraints on subunit orientations in a multi-body complex are provided by NMR **heteronuclear relaxation** (Brüschweiler *et al.* 1995; Tjandra *et al.* 1997; Dosset *et al.* 2000), **chemical shift anisotropy** (Prestegard 1998) or **residual dipolar couplings** (RDCs) (Tjandra and Bax 1997; Prestegard 1998; Dosset *et al.* 2001; Bax 2003), Fig. 1.3.

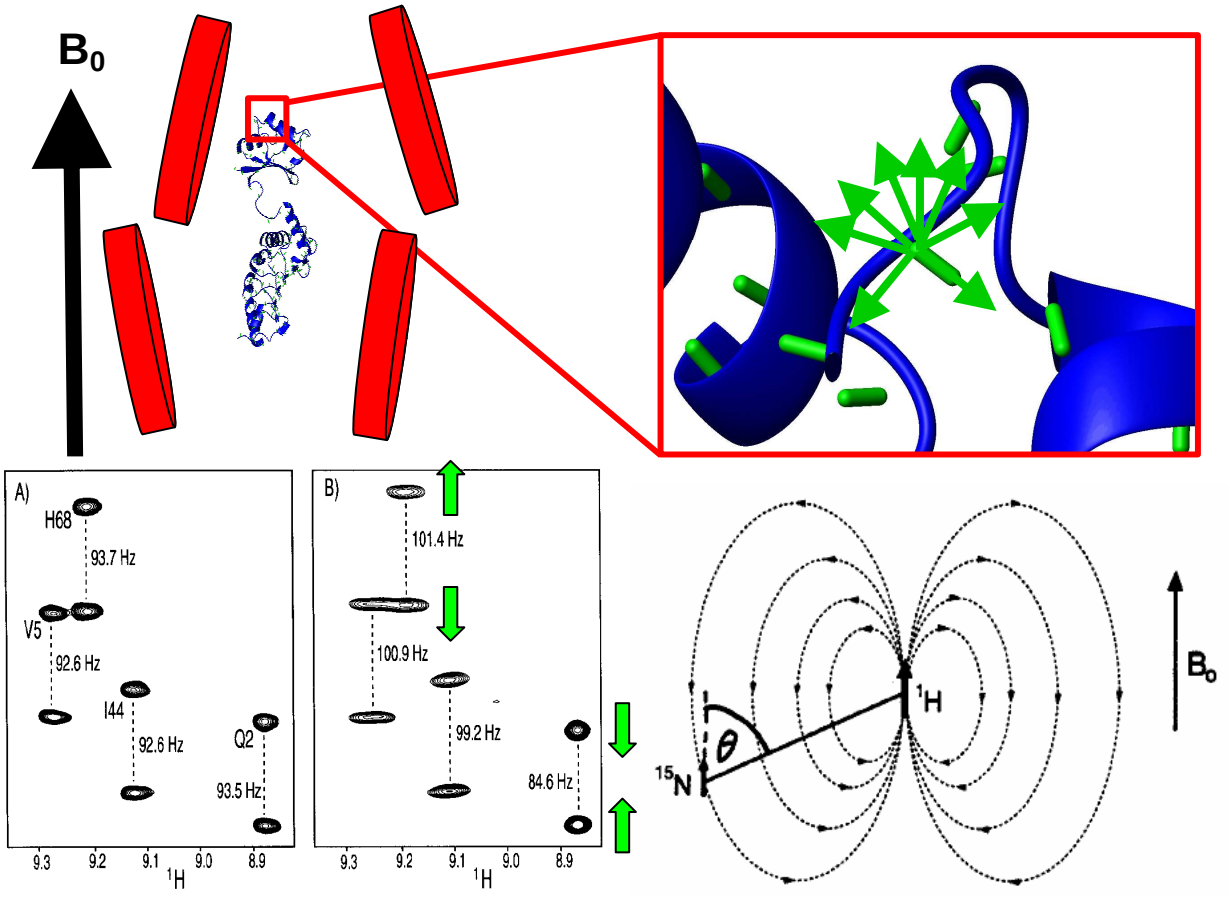

Fig. 1.3: Orientational information obtained by NMR residual dipolar couplings (RDCs): a slight anisotropy is created in liquid state NMR by introducing, *e.g.* phages, stretched gels, bicelles or diluted liquid crystals to the solution (Bax 2003, and references therein). Weak steric restraints or electrostatic interactions lead to an anisotropic tumbling of the dissolved biomacromolecules, including the nuclear bond vectors, *e.g.* ^{15}N-^{1}H (upper half of figure). The magnetic dipolar interaction (lower right-hand side) is no longer averaged to zero (as in isotropic environments) and increases or diminishes the measured *J*-couplings (lower left-hand side), yielding orientational restraints for each bond vector (lower part of figure adapted from Bax 2003).

NMR restraints (in particular RDCs) have often been used in an ***"ad hoc"* manner** to complement SAS restraints, *e.g.* by generating structures that are scored *a posteriori* against the experimental SAS curve (Yuzawa *et al.* 2004; Marino *et al.* 2006) or by using a **brute force grid search** to determine positions of RDC-oriented domains (Mattinen *et al.* 2002). Only a very limited number of protocols have been published recently that incorporate complementary NMR restraints **simultaneously** to SAS data for structural refinement. The strengths and shortcomings of two of them (Grishaev *et al.* 2005; Mareuil *et al.* 2007) are briefly discussed in the following section before the protocol developed by Gabel *et al.* 2006 is presented in detail in chapter 1.5.

1.4 Two recent examples of combined SAS-NMR refinement protocols

Grishaev *et al.* (2005) have developed a combined NMR and SAXS refinement protocol that generates a **driving force** (gradient) for each atom from a least χ^2-fit against the experimental SAXS data:

$$\nabla_{r_j}\left[\chi^2\right] \approx \sum_{k=1}^{N_Q} \frac{I_{\exp}(Q_k) - c_k I_{cal}(Q_k)}{\sigma_k^2} \sum_{i=1}^{N}\sum_{j\neq i}^{N} f_i\,(Q_k) f_j\,(Q_k) \left[\cos(Q_k r_{ij}) - \frac{\sin(Q_k r_{ij})}{Q_k r_{ij}}\right] \frac{\mathbf{r}_{ij}}{r_{ij}^2} \qquad \text{(Eq. 1.1)}$$

Q_k are the experimental wave vector transfers, N_Q their number, r_i the atomic coordinates, r_{ij} inter-atomic distances, I_{exp} the experimental scattering intensity, σ_k the standard deviations of their errors, I_{calc} the back-calculated intensity. c_k is a correction factor for the "globbic approximation" (see below), $f(Q_k)$ are atomic form factors and N the number of atoms.

The authors have applied Eq. 1.1 to refine the two-domain γS crystallin using experimental SAXS data in complement with a number of NMR restraints (NOEs, dihedral angles and RDCs). While this approach has the advantage to allow for intra-domain refinement (*e.g.* both domains are not considered as rigid bodies) and the authors were able to improve atomic RMSDs of γS crystallin with respect to a homologous γB crystallin crystal structure by applying SAXS (in addition to NMR) restraints, the protocol has a number of limitations: 1) it uses a "globbic" approximation (*i.e.* combining N = 3 to 9 heavy atoms, plus hydrogens, into a single scattering unit), 2) it drops 7 out of 8 experimental scattering points N_Q, 3) the accuracy is checked in terms of an atomic RMSD against a homologous crystal structure, 4) the starting structure is already NMR-refined, notably containing 15 inter-domain NOEs.

While points 1) and 2) are mainly required for reducing calculation time (~ N^2N_Q) by about a factor of 80 and can be justified, the globbic and surface solvent layer is used as an additional fit parameter. More severely, once the inter-domain restraints, in particular the inter-domain backbone linker, are severed, an improved structural refinement is no longer successful with SAXS (see Supporting Information in Grishaev *et al.* 2005), which is the case in "docking" two separate bodies (no connecting linker). The most fundamental problem, however, of using Eq. 1.1 for structural refinement is probably the fact that it generates individual driving forces for each single atom (or "dummy atom" in a "globbic" approximation). They are completely uncorrelated, the only criterion being that moving a given atom in a given direction is improving the χ^2-fit of the so-generated new structure. This may lead to situations where the effect (on the back-calculated scattering curve) of moving an atom A into one direction may be compensated by moving another atom B into another direction. Considering the degrees of freedom of atoms in a typical protein it is clear that this procedure cannot converge towards a unique structure in the general case without additional restraints on the atomic positions. Whether the restraints usually present in a protein structure (backbone bonds, steric and NOE restraints etc.) are sufficient or not to warrant the use of Eq. 1.1 for accurate structural refinement and to retrieve a unique structural solution is not discussed in the paper.

In another approach, Mareuil *et al.* (2007) propose a **genetic algorithm** to modify the structures of 25 "parents" (homology models or crystal structures) of γS cristallin or the S1KH bi-domain of NusA by changes in one or three backbone (ψ,ϕ)-angles that generate 125 "children". The population of children is varied by an exchange of backbone fragments. This procedure is repeated several times. 25 children are finally selected based on a fitness function, χ^2 (SAXS) + wχ^2 (RDCs), as well as RMSDs between families of structures (w is a weighing factor).

The genetic algorithm allows for intra-domain refinement and improves both the fitness and the RMSDs of the refined models. However, several questions pertaining to the efficiency and accuracy of the protocol remain open, as partly mentioned by the authors: 1) how is the respective weight of SAXS and RDC fitness and the overall weight of fitness with respect to the RMSDs to be distributed? (*e.g.* in the case of a cluster of refined structures having better fitness but a larger RMSDs than another), 2) how efficient is the protocol for severed domains (no connecting linker), 3) the authors were not able to distinguish between structures with different orientations, 4) neither the fitness nor the RMSDs of the target structures were obtained, suggesting that multiple minima exist and 5) how is an exhaustive conformational sampling achieved?

While point 3) is an inherent problem when incomplete RDC data are used and different domain orientations cannot be distinguished (see also next section), the other points indicate shortcomings of the protocol in its present state. In particular, the problem of multiple minima and the non-exhaustive conformational sampling remain to be addressed in a satisfactory manner.

1.5 A "generalized Guinier approximation" for rigid-body modeling using SAS and NMR RDC restraints (Gabel *et al.* 2006; Gabel *et al.* 2008)

Both the traditional approaches of rigid-body modeling using SAS data alone (or scoring of structures refined by other methods) as well as the two protocols presented in section 1.4 have a number of shortcomings that are unsatisfactory from the viewpoint of **uniqueness**, **accuracy** and **exhaustive sampling** of the structural models. In other words: 1) how can one be sure that a refined structure is the correct (accurate) one and not in a side-minimum of the target function? 2) **Does an ensemble (a family) of structures exist that complies with all restraints and how can it be represented?** 3) Another (more technical) problem is the χ^2-fit of the overall scattering curve: a back-calculated SAS curve can in general lie above or below the experimental data and equally well (mis)match the experimental data.

In order to address these shortcomings I have developed a protocol that refines a rigid-body system (of two bodies) simultaneously against rotational (NMR RDCs) and translational (SAS) restraints. Its particularity with respect to the approaches in the previous sections is that it is not a "bottom-up" (*i.e.* scoring a given conformation/structure against the overall SAS curve) but a **"top-down" approach** (*i.e.* refining at lowest resolution first by matching the radius of gyration, and then progressively refining against information from higher *Q*-values). One of its fundamental advantages is that it **generates and visualizes an exhaustive family of structures that comply with both SAS and NMR restraints simultaneously**.

The basic idea is to use a refinement of the positions of two rigid bodies (of known respective orientation from NMR RDC data) against an experimental SAS curve, going progressively from smaller angles to larger ones in an approach that could be baptized a **"generalized Guinier approximation"**: in a first step, the *parallel axes theorem* (*e.g.* Goldstein 1977) is exploited to position the two bodies at a distance R_1+R_2 (independent of the respective orientations, Fig. 1.4, right) that is related to the experimental radius of gyration of the complex, R_g, and the radii of gyration of the two bodies ($R_{g,1}$ and $R_{g,2}$). This approach has been exploited, *e.g.*, in the "triangulation" of the ribosome (Engelman and Moore 1972; Hoppe *et al.* 1975; Capel *et al.* 1987) or the RNA-polymerase (Stöckel *et al.* 1979). It can be expressed as follows:

$$R_g^2 = w_1 R_{g,1}^2 + w_2 R_{g,2}^2 + w_1 w_2 (R_1 + R_2)^2 \quad ; \quad w_i = \frac{\Delta\rho_i V_i}{\Delta\rho_1 V_1 + \Delta\rho_2 V_2}; i = 1,2 \qquad \text{(Eq. 1.2)}$$

w_i are the relative scattering "masses" of both bodies of scattering contrast $\Delta\rho_i$ and solvent-excluded volumes V_i.

In the following steps of the protocol, both bodies (with fixed relative orientations) are shifted on the surface of a great circle (with diameter R_1+R_2) until their positions are in agreement with the experimental SAS data in an intermediate angular range *beyond* the Guinier range (Fig. 1.4 right).

1.5.1 A generalized Guinier approximation

The mathematical starting point of the approach is the expansion of the scattering intensity into a Taylor series using Debye's equation (Debye 1915):

$$\begin{aligned} I(Q) &= \sum_i \sum_j f(\mathbf{r}_i) f(\mathbf{r}_j) \frac{\sin(Qr_{ij})}{Qr_{ij}} \\ &= \sum_i \sum_j f(\mathbf{r}_i) f(\mathbf{r}_j) \left(1 - \frac{Q^2 {r_{ij}}^2}{6} + \frac{Q^4 {r_{ij}}^4}{120} - \frac{Q^6 {r_{ij}}^6}{5040} + \frac{Q^8 {r_{ij}}^8}{362880} - + \ldots \right) \end{aligned} \qquad \text{(Eq. 1.3)}$$

The $\mathbf{r}_i$ are the (Cartesian) coordinates of the individual scattering centers (atoms or nuclei) and r_{ij} = $|\mathbf{r}_i - \mathbf{r}_i|$ the distance between two of them (Fig. 1.4, right). The f_i are the scattering lengths of the scattering centers in appropriate units. In a first approximation (for Q-values not too far beyond the Guinier range and for chemically not too heterogeneous bodies) they can be assumed as identical and constant, f_i=f_j=const. Out of mathematical convenience, the origin of the Cartesian coordinate frame can be chosen as the center of mass of both bodies (or, more accurately, the *center of scattering length density*). The coefficients of the Taylor series, $A := \sum_i \sum_j r_{ij}^{\,2}$, $B := \sum_i \sum_j r_{ij}^{\,4}$ etc, can be obtained by a (truncated) polynomial fit of the normalized experimental scattering data in a Q-range that exceeds the Guinier zone by a factor of 2-3 (Fig. 1.4, left):

$$I(Q) \approx 1 - \frac{1}{6} A_{\mathrm{exp}} Q^2 + \frac{1}{120} B_{\mathrm{exp}} Q^4 - \frac{1}{5040} C_{\mathrm{exp}} Q^6 + \frac{1}{362880} D_{\mathrm{exp}} Q^8 \qquad \text{(Eq. 1.3')}$$

The fact that a **limited** Q-range is sufficient for structural refinement is an important advantage of the approach with respect to the previous ones that require experimental data from an extended Q-range for the χ^2-fit.

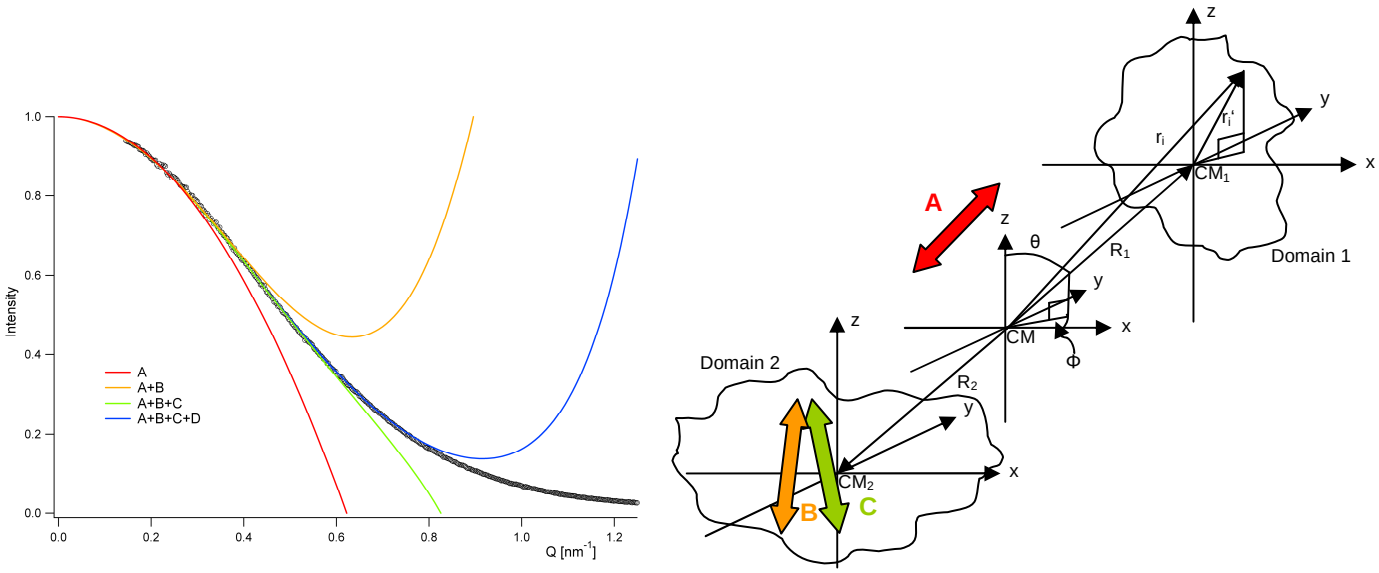

Figure 1.4: *Left*: illustration of a fit of experimental data with a truncated Taylor expansion (Eq. 1.3') up to the order of Q^2, Q^4, Q^6 and Q^8, respectively. *Right*: influence of the fit parameters A_{exp}, B_{exp}, and C_{exp} on the positions of the centers of mass in a two rigid-body system with fixed relative orientations.

The first term of the expansion is nothing else but the *Guinier approximation* (Guinier 1939) with A_{exp}=$2R_g^2$ and defines the distance between the centers (of scattering length density) of both bodies accurately, independently of their respective orientation (*parallel axes theorem*). In a coordinate frame where the centers of both bodies are expressed in polar coordinates, $0 \le \theta \le \pi$ and

$0 \leq \Phi < 2\pi$ (Fig. 1.4, right), the coefficients B_{exp}, C_{exp} etc. generate geometric restraints (target values) for the domain centers on the sphere *via*

$B=B_{exp}$ and $C=C_{exp}$ (Eq. 1.4)

N.B.: *B* and *C* depend on the geometry (anisotropy), distance, and respective orientation of both bodies. They can be calculated explicitly for any given conformation from the atomic coordinates which yields rather lengthy expressions (see appendix 2). All restraints can be combined into the following target function that describes the deviations of a given conformation (described by *A*, *B*, *C*) with respect to the target conformation (A_{exp}, B_{exp}, C_{exp}) determined from the experimental SAS curve (the λ_i are weighing factors):

$$\Psi(\theta,\phi,R_1+R_2)=\frac{\lambda_a}{6}\left|A-A_{\exp}\right|^2+\frac{\lambda_b}{120}\left|B-B_{\exp}\right|^2+\frac{\lambda_c}{5040}\left|C-C_{\exp}\right|^2 \quad \text{(Eq. 1.5)}$$

1.5.2 A family of structures in agreement with the SAS curve

As mentioned above, Eqs. 1.4 can be written down in terms of the atomic coordinates of both bodies, which yields rather lengthy expressions (appendix 2). Importantly however, their solutions can be expressed (for a given domain geometry and distance) in terms of the polar coordinates θ and Φ of the domain centers only, *e.g.* $B = B_{const} + B(\theta,\Phi)$. They represent one or more one-dimensional subspaces in the (θ,Φ)-plane defined by the condition $A = A_{exp}$ (*i.e.* great circles on the sphere with diameter R_1+R_2). In other words, **the degrees of translational freedom of both bodies (at a given orientation) are confined to one-dimensional lines on the surface of a sphere determined by their inter-domain distance** (Fig. 1.5).

If a solution to the scattering problem exists, Eqs. 1.4 define two lines that intersect at one or more discrete points (or entire lines) in the (θ,Φ)-plane (Fig. 1.5) that define **uniquely** the structures that are in agreement with the SAS curve up to order of Q^6. If the solutions of Eqs. 1.4 consist in one or more discrete points (non-degenerative case), they represent the **only** possible structures and scattering data at higher angles **contain only redundant information**. In the case of solutions consisting in a single one-dimensional line (degenerative case, *e.g.* for certain rotational symmetries in the bodies), they represent a family of structures that are all in agreement with the scattering curve).

In the ideal case (accurate and precise extraction of A_{exp}, B_{exp} and C_{exp}) **the "top-down" approach therefore fulfils all requirements mentioned at the beginning of this section** (uniqueness, accuracy and complete conformational sampling) and **identifies exhaustively all possible structural solutions** (including side-minima). It therefore remedies a fundamental shortcoming

of traditional rigid-body refinement where only one structure is produced and identified as "the" solution to the problem.

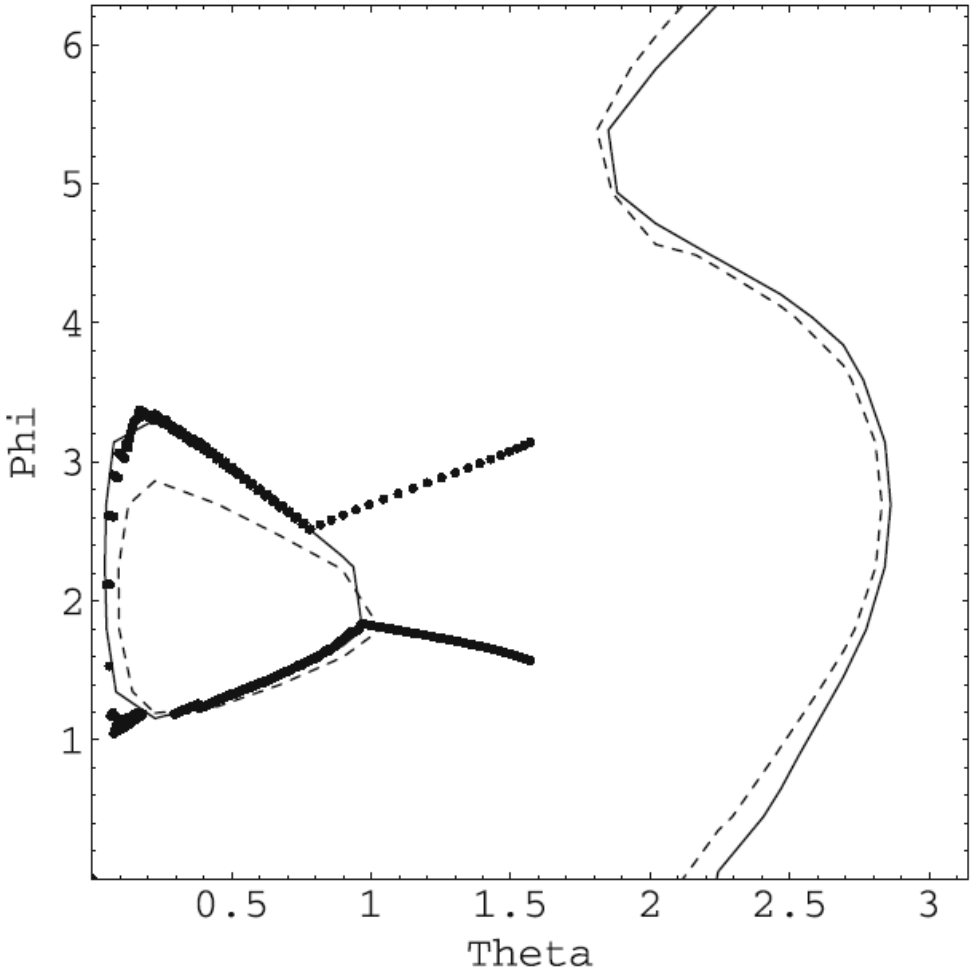

Figure 1.5: Equipotential lines B=B_{exp} (continuous) and C=C_{exp} (dashed) for the Barnase-Barstar two-body system (from Gabel *et al.* 2006). All possible positions of the body centers (for a given, fixed relative orientation) are defined (in the ideal case) by the intersection(s) of these lines in the (θ,Φ)-plane. In practice, acceptable structures are found in the vicinity of both lines. (The filled black circles are successive steps in the refinement process of a domain center from two different starting points).

Another advantage of the protocol is that it can be used to **quantify the uniqueness and variability** of a structural model *a posteriori* to its determination by indicating the remaining translational degrees of freedom of the bodies that will not (or only slightly) alter its match with the experimental SAS curve. These degrees of freedom are defined by vectors that point along the directions defined by the B=B_{exp} line in (θ,Φ)-space, orthogonally to the inter-body distance vector.

1.5.3 Efficiency of the protocol as a function of domain anisotropy and distance

A corollary of the expressions derived for the B and C terms (see appendix 2) is the possibility to quantify the refinement capacity and efficiency of the protocol as a function of domain

anisotropy and distance. In the simplest case of both rigid bodies having spherical symmetry, one retrieves the fundamental result that B and C terms no longer depend on θ or Φ and a refinement beyond fixing the inter-domain distance is not possible. A more elaborate result is that the B=B_{exp} and C=C_{exp} conditions are the more distinct the larger the expressions $r_{ij}/(R_1+R_2)$ and the larger the coefficients j_C though s_C (appendix 2): in other words, **the more anisotropic the domains and the more extended they are with respect to their mutual distance, the better the efficiency of the protocol** to determine unique structural solutions. These considerations can be elaborated further (see "Perspectives" section).

1.5.4 Uniqueness and accuracy in the "real" case

In the practical case a number of error sources are present: i) experimental SAS errors, ii) approximate extraction of A_{exp}, B_{exp} and C_{exp} due to the truncated polynomial fit, iii) presence of a hydration shell, iv) experimental RDC errors. Most of them are discussed in detail in Gabel *et al.* (2006). They are summarized briefly here:

i) **Experimental SAS errors** will limit the accuracy of the extracted A_{exp}, B_{exp} and C_{exp}. In practice, the solutions of Eqs. 1.4 are hard to distinguish and impose similar geometric restraints on the domain positions (Fig. 1.5). Therefore, even in the non-degenerative case, the structural solutions will generally consist in a family of structures aligned in the vicinity of the one-dimensional great circle B=B_{exp}. Their spread can be reduced by scoring them against the entire SAS curve or against complementary structural information from NMR, *e.g.* PREs (Battiste and Wagner 2000). Usually this scoring will select a number of structures stretched out along a finite stripe close to the crossing point defined by B_{exp}=B and C_{exp}=C (Fig. 1.6). In the examples studied so far (Gabel *et al.* 2006; Gabel *et al.* 2008) the D-term yielded no improvement on the refined structures. It rather represents a potential source of systematic errors if not determined accurately.

ii) A **truncated Taylor series** introduces systematic errors in the extracted A_{exp}, B_{exp} and C_{exp} values. While a working procedure has been explained in Gabel *et al.* (2006) to extract them in an efficient way (first A_{exp} in the Guinier range, and then progressing to higher angles), several questions pertaining to their independence and potential effects of systematic errors remain open. They are discussed in the "Perspectives" section of the thesis.

iii) The **presence of a hydration shell** of different density than the bulk solvent (Svergun *et al.* 1998) is not taken into account in the present form of the protocol and can lead to

systematic errors in the comparison between experimental and back-calculated target values. Strategies to solve this problem are discussed in detail in the "Perspectives" section.

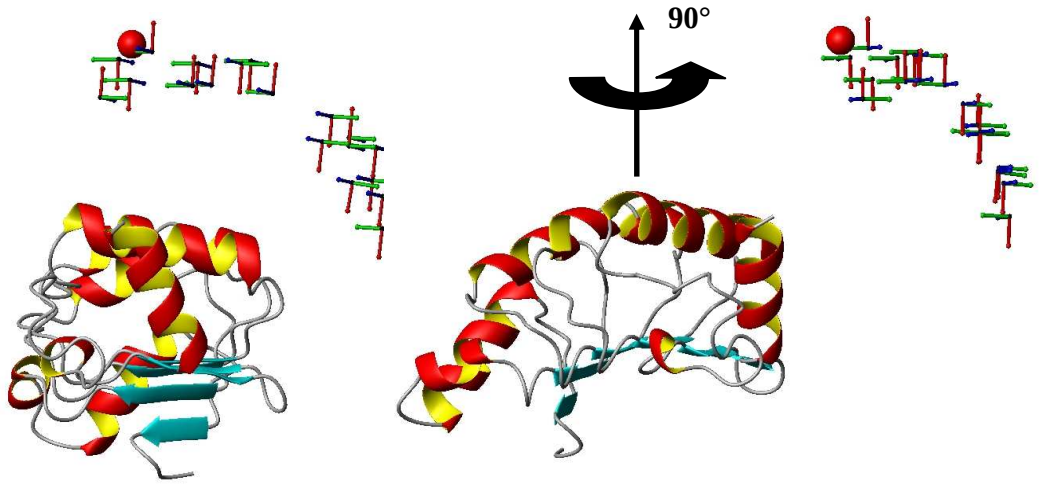

Figure 1.6: Family of structures refined against A_{exp}, B_{exp} and C_{exp}, with additional scoring against the complete SAS curve. Only the centers of mass of the second domain are shown (TAP protein from Gabel *et al.* 2008). The small red sphere denotes the position of the domain center of the target structure from crystallography. All structures represented are indistinguishable in terms of a χ^2-fit against the entire SAXS curve (with errors).

iv) **Experimental RDC errors** and the subsequent errors in the domain orientations are discussed extensively in Gabel *et al.* (2006 and 2008). In summary, if the domain orientations are not uniquely determined, *e.g.* for a single RDC set from a single alignment medium (Bax 2003), the accurate structural solution cannot be found and "false positives" can, in general, yield an equally good match of the experimental data as the "correct" structure. It is therefore imperative for the accurate functioning of the protocol that the correct domain orientations be known (*e.g.* by using several alignment media). However, once the correct orientation is known, the protocol is stable against uncertainties of the domain orientations of up to about 5 degrees in the cases studies.

1.5.5 Incorporation into ARIA-CNS *(Gabel* et al. *2008)*

The fundamental ideas developed in the preceding sub-sections have been incorporated into the program suite *ARIA-CNS* that is broadly used for structural refinement of biomacromolecules in NMR and crystallography (Brünger *et al.* 1998). It was applied and tested on a single simulated RDC dataset and SAS curve from the 31 kDa nuclear export factor TAP (Liker *et al.* 2000). Several of the above-mentioned sources of errors were tested: i) A_{exp}, B_{exp} and C_{exp} were endowed with

error marges of 3, 5 and 5%, respectively, ii) number and noise of active RDCs, iii) efficiency of the protocol with only RDC (or only SAS) restraints.

The technical details of the program are not explained here, they can be found in Gabel *et al.* (2008). We limit our discussion to the major results:

- The fourfold rotational degeneracy introduced by a single RDC dataset can be (partially) lifted when the SAS restraints are activated
- When A+B+C terms are active, the possible structures (with correct orientation) are lined up in a stripe in space defined by the conditions $A=A_{exp}$, $B=B_{exp}$ and $C=C_{exp}$
- A further χ^2-scoring against the overall SAS curve reduced this family of structures to a stripe in space of finite length (Figur 1.6)

As shown in the publication, the strengths of the protocol are in particular

i) the **economic computational time**

ii) the choice of **activating SAS restraints on different levels** (*e.g.* only A_{exp}, A_{exp}+B_{exp} etc.) and

iii) a **limited angular range** (maximum of about two to three times the Guinier range) is sufficient to determine the structural solutions

Points ii) and iii) are particularly advantageous in practical situations when no (or only noisy) data are available at higher angles. Finally, the protocol was checked against another rigid-body refinement program ("SASREF"; Petoukhov and Svergun 2005) which could not provide a better refinement within the SAS errors in the study. Limitations of the protocol as well as strategies for improvement are discussed in detail in the "Perspectives" section.

2. Combination of small angle neutron and X-ray scattering for the structural study of unfolded proteins

2.1 Introduction

Both small angle neutron and X-ray scattering have been used extensively for several decades to study the structural properties of polymers in soft condensed matter (*e.g.* Schurtenberger 2002 and references therein). They are particularly suited to provide information on several lengthscales (going from radii of gyration over persistence lengths to cross-sectional analysis). An extensive review and discussion of the results is beyond the scope of this thesis. Structural properties of unfolded proteins have also been studied by SAXS and SANS (Kirste *et al.* 1969; Calmettes *et al.* 1994; Petrescu *et al.* 1997; Petrescu *et al.* 1998; Pérez *et al.* 2001; Doniach 2001; Millett *et al.* 2002; Kohn *et al.* 2004), in part by using the concepts developed in the polymer sciences. Two very interesting and promising recent topics are the presence of **residual native structures** (Shortle and Ackerman 2001; Plaxco and Gross 2001; Mittag and Forman-Kay 2007) and the description of the unfolded state in terms of explicit **conformational ensembles** (Bernadó *et al.* 2005; Bernadó *et al.* 2007). Both findings were stimulated by **complementary information from NMR** (in particular RDCs), a fruitful combination with very high potential for future studies of unfolded proteins (Putnam *et al.* 2007; Jensen *et al.* 2009).

It has been recognized that the study of polymer structure can benefit from contrast variation and specific deuterium labeling using SANS, in particular regarding the cross-sectional analysis (Rawiso *et al.* 1987; Schurtenberger 2002). It would be desirable to systematically combine SAXS and SANS in a similar way to study structural properties of unfolded proteins. Particular questions that can be specifically (and sometimes uniquely) addressed by this combination include the following: i) how are the **interactions of protein and solute** organized in solution?, ii) **solvent properties** in the vicinity of the unfolded protein in contrast to the folded protein, iii) properties of **side-chain flexibility** and its effect on the scattering curve at higher angles, iv) deviations from the random-coil model, *e.g.* **excluded volume effects**.

In contrast to polymers, such a combined SAXS/SANS data analysis on biological systems is difficult in many cases due to their fragility: both experiments ideally have to be done on the same sample within a minimum time delay. However, it is particularly promising in places where this combined use is recognized and promoted (*e.g.* through a joint SAXS/SANS proposal system set up recently between ILL and ESRF in Grenoble, France). The following section describes the benefits of such a combined SAXS/SANS analysis to study the denaturing effects of

urea on ubiquitin in a model-free way. Future prospects of this work are presented in the "Perspectives".

2.2 Urea-denaturation of ubiquitin, a recent example illustrating the benefit of combined SAXS/SANS data analysis for unfolded proteins (Gabel *et al.* 2009)

Despite the abundance of biophysical and biochemical data from unfolded proteins in the presence of denaturant, the molecular origin of solvent-induced protein denaturation remains unclear. Two models have emerged to explain the unfolding effect of urea: the first invokes the disruptive effects on water structure, so that urea acts as a **better solvent** for hydrophobic groups which are mostly buried in water-solvated proteins (Tanford 1964). Another model proposes that urea **binds directly** to multiple sites on the protein backbone, thereby destabilizing the native fold relative to the unfolded state (Schellman 1958).

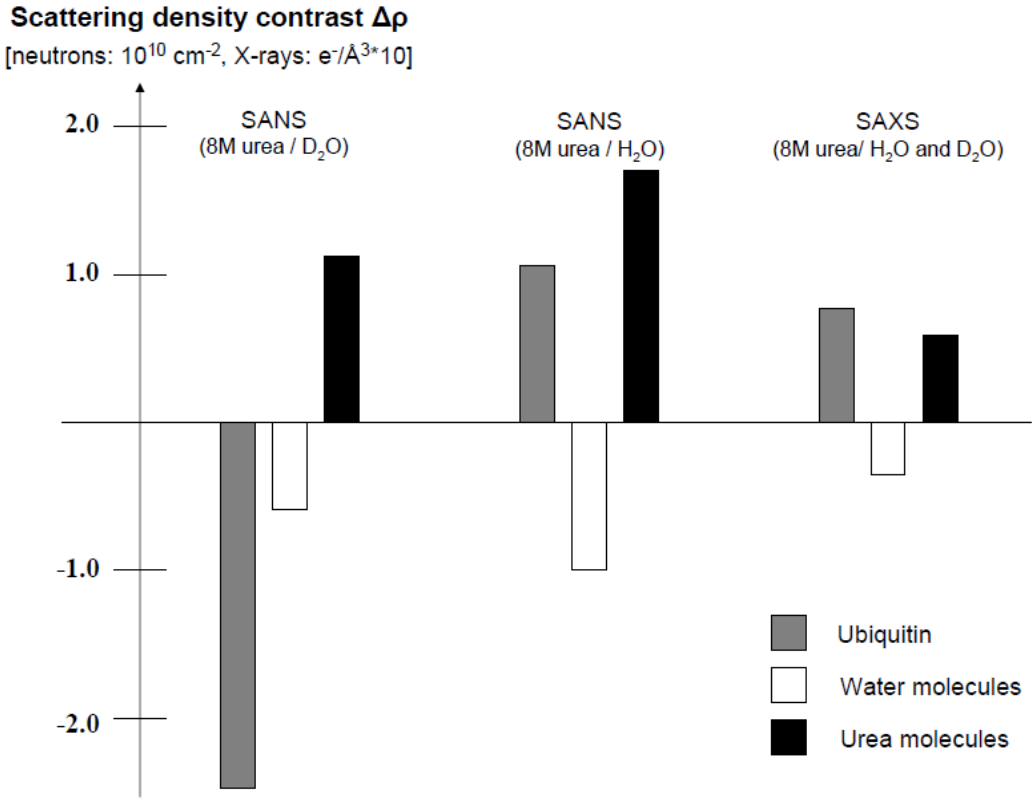

Fig. 2.1: Scattering contrast for ubiquitin, water and urea molecules for neutrons in 8M urea/(D_2O and H_2O) and for X-rays in 8M urea (identical values in H_2O and D_2O). Opposite signs in contrast of two components associated in solution induce an apparent diminution of the molecular weight while identical signs lead to an increase.

In order to probe the denaturing effect of urea on a protein we conducted a combined SANS/SAXS analysis on ubiquitin (MW=8.6 kDa) in 8M urea solution (in H_2O and D_2O at pH 2.5 and 6.5). The fact that ubiquitin, urea, H_2O and D_2O molecules have different, specific scattering length densities for neutrons and X-rays (Fig. 2.1) was exploited and the interaction between urea

and ubiquitin was probed *via* the scattered intensity in the forward direction, $I(0)$. In the context of the *invariant particle hypothesis* (*e.g.* Tardieu *et al.* 1981) it is proportional to the square of the integrated sum of the scattering length densities of the particle components times the number of particles, N:

$$I(0) \propto N\left(\sum_i (\rho_i - \rho_S)V_i\right)^2 \qquad \text{(Eq. 2.1)}$$

(ρ_i, ρ_S: scattering length densities of the particle components and of the solvent, respectively; V_i: solvent-excluded volume of the particle components).

A comparison of the overall scattering curves from ubiquitin in H_2O and D_2O urea solutions by SANS (D22, ILL, Grenoble) and SAXS (ID02, ESRF, Grenoble) revealed that at pH 6.5 the experimental data match the one predicted for folded ubiquitin over the entire Q-range, indicating that the protein is in its **native conformation** even when dissolved in high concentrations of denaturant. Lowering the pH to 2.5 resulted in a change in overall shape of the scattering curve along with a significant increase of the radius of gyration, suggesting that under these conditions the protein unfolds into a disordered conformation (Fig. 2.2).

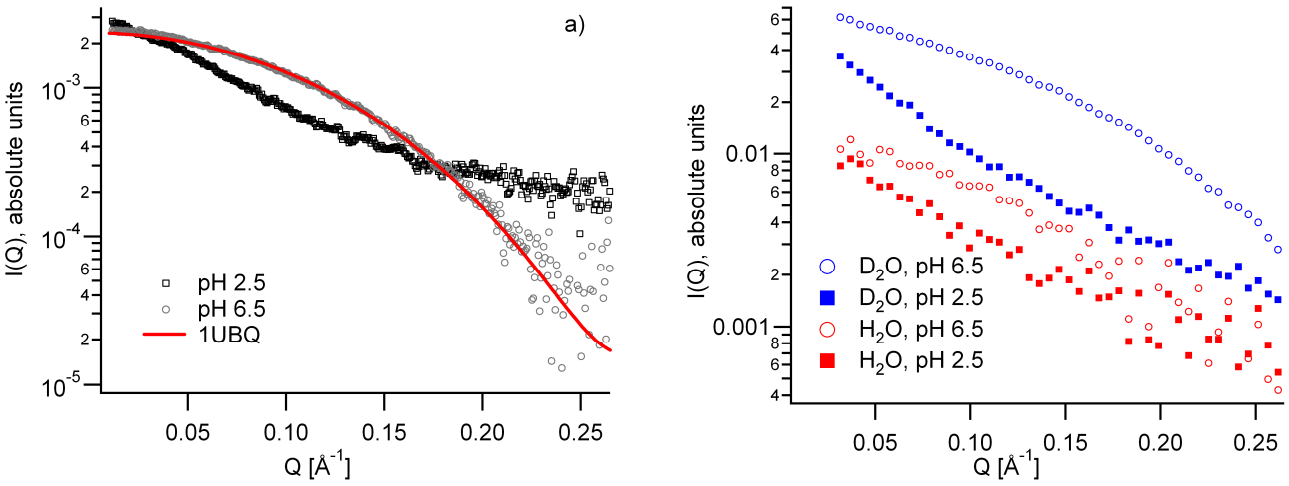

Fig. 2.2: SAXS (left) and SANS (right) curves of ubiquitin in urea at various solvent conditions.

A remarkable feature, observed by both SAXS and SANS is the change in $I(0)$ due to lowering the pH: an **increase** is observed by SAXS upon denaturation, while SANS in D_2O shows a significant **decrease**. The situation for SANS in H_2O is ambiguous due to large errors (Fig. 2.2). Both the extent and the sign of the intensity change can be explained quantitatively by assuming that a certain number ΔN_{urea} of urea molecules are recruited by each ubiquitin molecule upon unfolding:

$$\sqrt{\frac{I_{pH\,2.5}(0)}{I_{pH\,6.5}(0)}} = \frac{(\rho_{prot} - \rho_S)V_{prot} + (N_{urea}^{pH\,6.5} + \Delta N_{urea})(\rho_{urea} - \rho_S)V_{urea}}{(\rho_{prot} - \rho_S)V_{prot} + N_{urea}^{pH\,6.5}(\rho_{prot} - \rho_S)V_{prot}} = 1 + \frac{\Delta N_{urea}}{N_{urea}^{pH\,6.5} + F} \qquad \text{(Eq. 2.2)}$$

The indices "*prot*", "*urea*" and "*S*" refer to ubiquitin molecules, urea molecules and the bulk solvent (8M urea/water mixtures), respectively. $N_{urea}^{pH6.5}$ is the number of urea molecules that are potentially already bound at pH 6.5. F is a number that depends on the contrast of urea and ubiquitin in the respective experimental conditions (F=163, 82, -304 for SAXS, SANS in 8M urea/H_2O and 8M urea/D_2O, respectively). The $I(0)$ intensities were extracted by the Guinier approximation (Guinier 1939) for the folded samples (pH 6.5) and by the Debye function (Debye 1947) for the unfolded samples (pH 2.5).

Eq. 2.2 yields **three independent equations** for SAXS, SANS in 8M urea/H_2O and SANS in 8M urea/D_2O. It can be solved graphically (Fig. 2.3) and yields as a result that about **ΔN_{urea}=20 urea molecules are being recruited per ubiquitin molecule upon unfolding**. The results are not very sensitive to the number of urea molecules already bound at pH 6.5. However, an absolute measurement of the molecular weight using water as a calibration of the SANS data (Jacrot and Zaccai 1981) suggested that not more than about 10 urea molecules are bound in the native state.

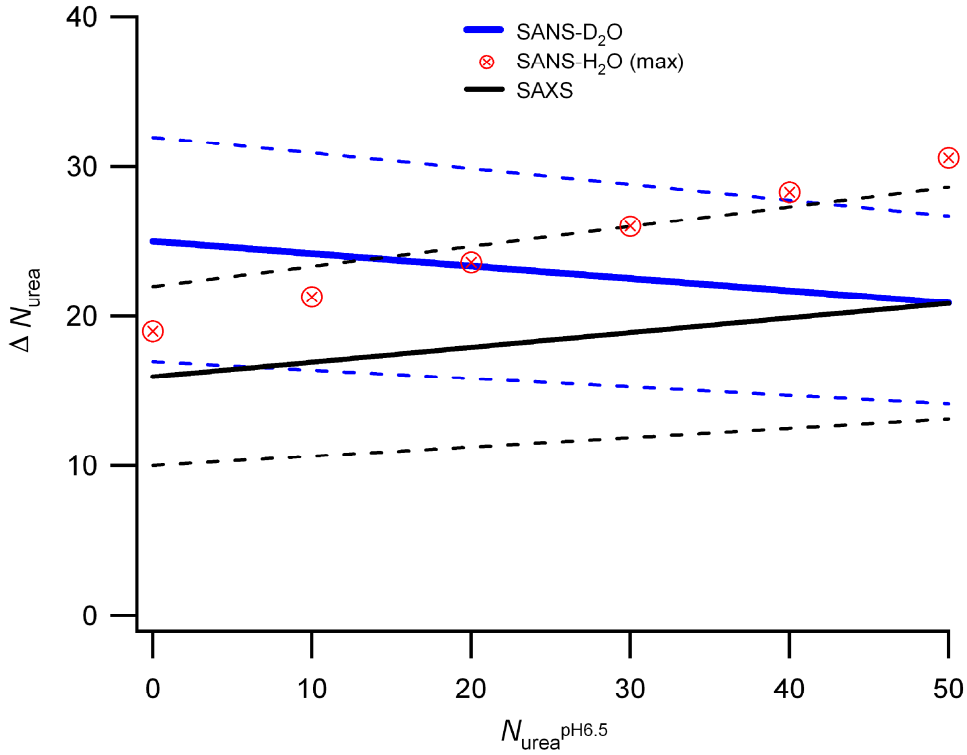

Fig. 2.3: Number of urea molecules recruited by ubiquitin upon unfolding is reported as a function of the number of urea molecules potentially already bound at pH 6.5. Thick lines are obtained from the average values of the $I(0)$. The thinner, broken lines are the extreme values using the maximum errors from the intensities. For SANS-H_2O, only the maximum values are shown.

The results rely strongly on the **complementarity** of neutrons and X-rays and could not have been unambiguously obtained by either technique alone. For example, the SAXS data alone could have been interpreted in terms of a hydration shell of different density than the bulk solvent (Svergun *et al.* 1998) that increases its volume upon unfolding (by about a factor of 3 for a

3Å thick shell). However, this interpretation is inconsistent with the SANS results. The combined data analysis rather suggested that the hydration shell of unfolded ubiquitin does not have the same thickness as the one of the folded protein and/or that it must have a different density.

In conclusion, the complementary scattering properties of H_2O, D_2O, ubiquitin and urea for SANS and SAXS permitted a **quantitative and model-free** analysis of the interaction of urea with the protein ubiquitin under denaturing conditions. The results contribute to a more detailed understanding of the molecular mechanisms of protein denaturation, by supporting the model of direct binding of urea to the protein.

2.3 Urea-denaturation of ubiquitin: structural interpretation beyond the interaction analysis

The model-free analysis of the previous section describes the interaction between urea and ubiquitin molecules from the intensities scattered at zero angle, $I(0)$. It is clear that the scattering curves at higher angles contain additional, complementary structural information, in particular on the unfolded state. *E.g.*, a Kratky representation (Fig. 2.4) of the SAXS data suggested that the unfolded state cannot be adequately described by a random-walk model but that **excluded volume effects** etc. need to be taken into account for an accurate interpretation. Further points like **cross-sectional analysis** should confirm and complement the findings of the $I(0)$ analysis. These and other points are discussed in the "Perspectives" section.

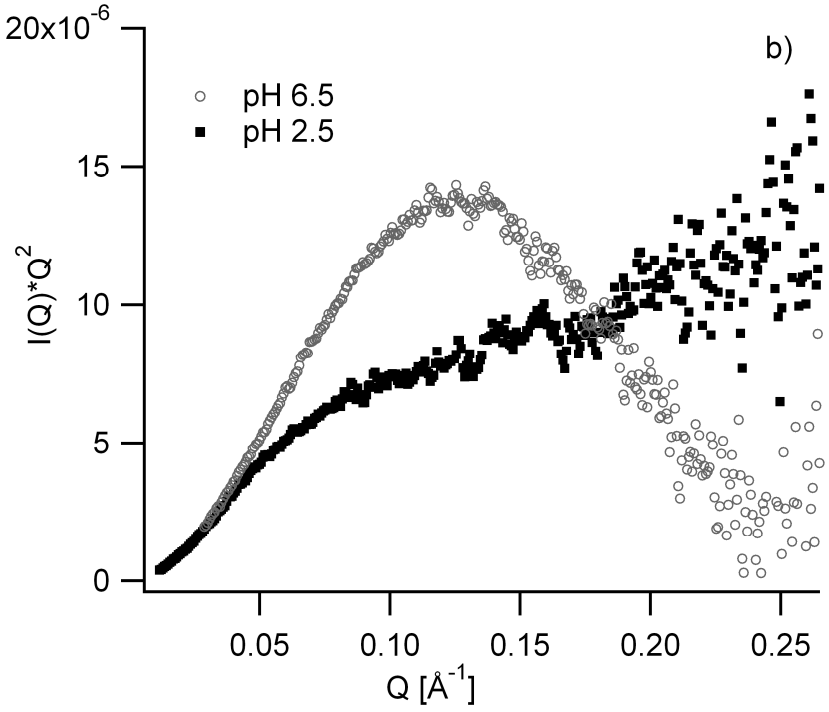

Fig. 2.4: Kratky-plot of the SAXS data in urea/H_2O (from Gabel *et al.* 2009). The data at pH 6.5 show the typical profile of a compact particle in solution, the pH 2.5 data are representative for an unfolded protein. However, the latter can not be explained by a simple random-walk model (Gaussian chain) since a clear plateau is missing in the plot.

3. Biomacromolecular and solvent dynamics by elastic neutron spectroscopy: the advantages of combined data analysis from several instruments and analogies to small angle scattering

3.1 Introduction

Neutron spectroscopy (NS) has been employed widely over the last 30 years to study the molecular dynamics of biomacromolecules, either in amorphous powder (dry or hydrated) or in solution state. Pioneering work includes studies on the vibrational spectra of hexokinase (Jacrot *et al.* 1982) and lysozyme (Middendorf *et al.* 1984) the quasielastic spectrum of phycocyanin (Bellissent-Funel *et al.* 1989) and the discovery of the dynamical transition involving all atoms in myoglobin by elastic spectroscopy (Doster *et al.* 1989). A review of work up to the year 2000 has been published by Gabel *et al.* (2002).

Systems studied include a variety of soluble native and denatured (Receveur *et al.* 1997; Paciaroni *et al.* 1999; Pérez *et al.* 1999; Fitter 1999; Bu *et al.* 2001; Wood *et al.* 2008a) and membrane proteins (Ferrand *et al.* 1993; Fitter *et al.* 1996) *in vitro* as well as macromolecules *in vivo* (Tehei *et al.* 2004; Jasnin *et al.* 2008a). Various sample conditions have been exploited, including hydration degree (Diehl *et al.* 1997; Lehnert *et al.* 1998; Pérez *et al.* 1999 ; Stadler *et al.* 2009) and type of solvent (Cordone *et al.* 1999; Réat *et al.* 2000; Tsai *et al.* 2000; Tehei *et al.* 2001; Paciaroni *et al.* 2002; Gabel *et al.* 2004). The dynamic properties of aqueous solvent were equally studied as a function of temperature (Teixeira *et al.* 1985), the presence of ions (Bellissent-Funel *et al.* 1984; Salmon 1987) or other solutes (Russo *et al.* 2004; Magazù *et al.* 2004), both in the vicinity of biomacromolecules (Randall *et al.* 1978; Bellissent-Funel *et al.* 1996; Settles and Doster 1996; Bon *et al.* 2002; Wood *et al.* 2008a) as well as in entire biological cells (Trantham *et al.* 1984; Ford *et al.* 2004; Tehei *et al.* 2007; Jasnin *et al.* 2008b; Stadler *et al.* 2008). A recent review of water dynamics in several biological systems is given by Frölich *et al.* (2009).

In contrast to SANS (only momentum exchange), NS measures **both momentum and energy exchange between neutrons and a sample as a function of *Q***. While coherent scattering is analyzed in crystals (Niimura 1999) and SANS, disordered biological samples (amorphous powders or solutions) scatter thermal neutrons (~Å) mainly incoherently. It can be described in terms of a space-time Fourier transform of the atomic auto-correlation (or "self-diffusion distribution") function (van Hove 1954; Vineyard 1958; Rahman *et al.* 1962). In general, the proton incoherent scattering cross section dominates the (coherent and incoherent) ones of most other atoms (including deuterium) by more than an order of magnitude (Köster *et al.* 1991). Thus thermal neutrons are an **ideal probe for the study of proton "single particle" dynamics in**

(disordered) biological samples. In a manner of speaking, the (sometimes disturbing) incoherent background signal in a SANS experiment is the signal containing information in an NS experiment and is analyzed as a function of energy exchange and *Q*.

The timescales (pico- to nanoseconds) and lengthscales (~Å) probed by NS are similar to those accessible by NMR relaxation (Palmer 2001; Brüschweiler 2003) and molecular dynamics (MD) simulations. They can be straightforwardly compared to the latter (Smith 1991). Since scattering from a disordered biological sample is dominated by the incoherent cross section of hydrogens, adapted labeling schemes (exchanging hydrogen against deuterium, so-called "inverse" labeling in contrast to NMR) have been used to focus specifically on the hydrogenated parts of a sample, *e.g.* specific amino acids in proteins (Réat *et al.* 1998; Wood *et al.* 2008b).

3.2 Elastic *vs.* inelastic spectroscopy

Three major types of experiments can be distinguished conceptually in neutron spectroscopy: **inelastic** (INS), **quasielastic** (QENS) and **elastic neutron scattering** (ENS) (Fig. 3.1). In INS, vibrational atomic motions exchanging a discrete amount of energy with the neutrons are probed while QENS measures continuous exchanges of energy close to zero energy transfer that are associated with, *e.g.*, diffusive motions or overdamped vibrations (often symmetrically distributed around ω=0). ENS, finally, measures the elastically scattered neutrons (within the instrumental energy resolution). It can contain both contributions from (truly elastic) as well as quasielastic processes.

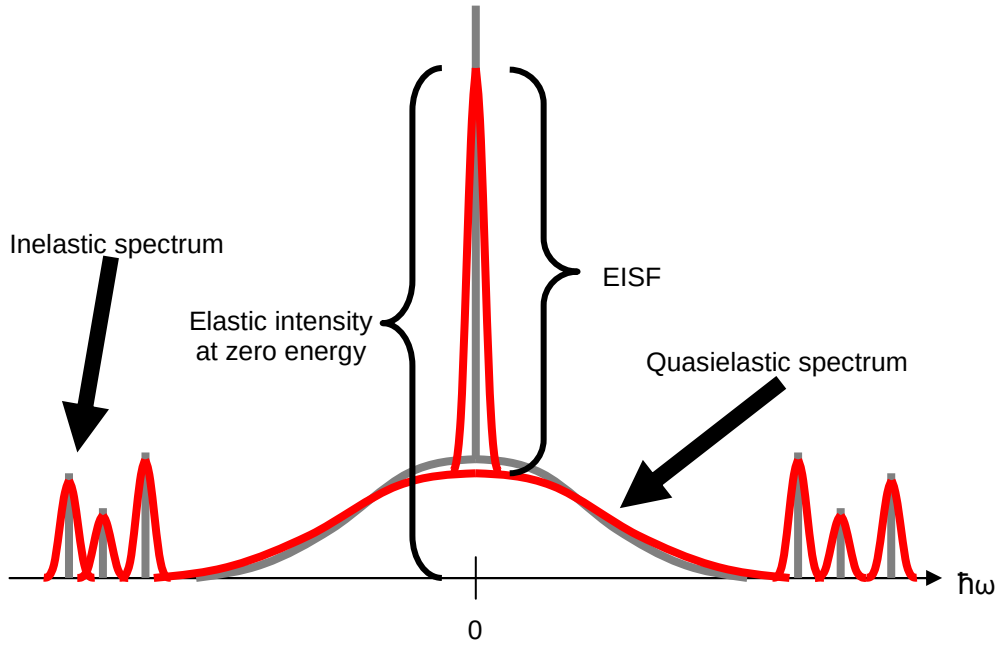

Figure 3.1: Schematic representation of a neutron spectrum from a disordered biological sample at a given *Q*-value. The elastic, quasielastic and inelastic parts of the theoretical scattering law $S(Q,\omega)$ are in gray, a finite energy resolution $R(\omega)$ is reflected by a broadening of the lines (in red).

While all three types of experiments are complementary, they usually employ different parameters to interpret molecular dynamic processes and a direct comparison of, *e.g.*, QENS and ENS data of the same system is not always straightforward (Gabel *et al.* 2002). Particularly frustrating is the **comparison and interpretation of ENS data from different instruments** (energy resolution, *Q*-range), even of the same biological system. For example, atomic mean square displacements (MSDs) or force constants (Zaccai 2000) at higher temperatures (> 200 K; when quasielastic contributions are present) cannot be compared when extracted in different *Q*-ranges and/or at different instrumental energy resolutions (Gabel 2005; Gabel and Bellissent-Funel 2007).

While ENS data usually contain less information than QENS data, the former are often preferred to the latter for practical reasons when **instrumental flux is low** and/or if the **sample is weakly scattering**. In addition, data analysis in terms of MSDs is technically easier and model-free. However, this approach poses the problem (even more than in QENS) of attributing and interpreting the measured MSDs from a complex sample (*e.g.* protein in solution, entire biological cells) as a contribution from the component of interest (*e.g.* intracellular proteins). In addition it requires that the user be certain that contributions from undesired components (*e.g.* hydration water; see section 3.5) may be neglected.

The difficulties in comparing ENS with QENS data and ENS data from different instruments as well as the problems of assigning MSDs from ENS to specific components in a complex system have motivated the proposal of a **convolution formalism for the interpretation of ENS data** (Gabel 2005; Gabel and Bellissent-Funel 2007). The purpose of the following sections is to elucidate this approach in more detail. Its limitations and possible future developments are discussed in the "Perspectives" section at the end of the thesis.

3.3 Interpretation of ENS data in terms of convoluted scattering laws

The approach of measuring only neutrons scattered elastically and interpreting the signal in terms of a convolution product of the theoretical scattering law with the instrumental energy resolution function $R(\omega)$, evaluated at $\omega=0$ (Eq. 3.1) has been used as an economic, time-saving approach since the 1970s in the case of small chemical molecules. A pioneering example of this approach (also termed "*fixed window spectroscopy*" or "*elastic scan*" in the case of zero-energy transfer) is the study of methyl group rotations in tetramethyl-ammonium-manganese chloride $N(CH_3)_4MnCl_3$ (TMMC) as a function of temperature and *Q* (reviewed in Springer 1977).

$$EI := S(Q,0) = \int_{-\infty}^{+\infty} S(Q,\omega)R(\omega)d\omega \qquad \text{(Eq. 3.1)}$$

While being a well-known phenomenon in QENS (Lechner 2001 and references therein) the convolution formalism had been hardly used by the biological community until the late 1990s. It has regained considerable interest over the last 10 years, both from a theoretical point of view (Kneller and Calandrini 2007; Zorn 2009) and a practical one, *e.g.* for explaining instrument-resolution effects on the protein dynamical transition (Daniel *et al.* 1999; Hayward and Smith 2002; Becker and Smith 2003; Becker *et al.* 2004). It is particularly important and useful when different types of motion (diffusive, vibrational), possibly from different components in the sample (macromolecules, water), are present (Gabel 2005): in these cases, it is possible by using this approach to focus on, *e.g.*, macromolecular motions and minimize the "contamination" of the elastic signal from the surrounding hydration water or solvent, **even in H_2O**.

3.4 Elastic intensity *vs.* Elastic Incoherent Structure Factor

The measured elastic intensity (*EI*) (Eq. 3.1) needs to be distinguished conceptually from the elastic incoherent structure factor (*EISF*) (Fig. 3.1). Speaking in a general manner, for finite energy resolutions the *EI* may contain a "contamination" from QENS intensity scattered closely around ω=0. For example, a simple composite scattering law

$$S_{theo}(Q,\omega) = A_0(Q)\delta(\omega) + A_1(Q,\omega) \qquad \text{(Eq. 3.2)}$$

(with $EISF=A_0$) yields after convolution with the instrumental energy resolution $R(\omega)$:

$$S_{measured}(Q,\omega) = A_0(Q)R(\omega) + \int_{-\infty}^{+\infty} A_1(Q,\omega')R(\omega-\omega')d\omega' \qquad \text{(Eq. 3.3)}$$

and therefore the elastic intensity

$$EI = S_{measured}(Q,0) = A_0(Q)R(0) + \int_{-\infty}^{+\infty} A_1(Q,\omega)R(\omega)d\omega \qquad \text{(Eq. 3.4)}$$

The *EI* and *EISF* can in principle only be de-convoluted strictly with the help of a quasielastic scan (Bée 1988). If both the instrumental energy resolution and the quasielastic scattering law are described by Lorentzian functions with half widths at half maximum (HWHM) Γ and Γ_{Lor}, respectively, Eq. 3.4 can be calculated analytically:

$$EI = S_{measured}(Q,0) = \frac{A_0(Q)}{\pi\Gamma} + \frac{A_1(Q)}{\pi(\Gamma+\Gamma_{Lor})} \qquad \text{(Eq. 3.4')}$$

For Q-values where $\Gamma \ll \Gamma_{Lor}$, the *EI* is dominated by truly elastically scattered neutrons. The *EISF* can then be obtained either by integrating the elastic peak in a quasielastic scan or by

normalizing the *EI* to a very low temperature (where no quasielastic contributions are present and A_0 is a constant as a function of *Q*, *i.e.* the atoms are immobile).

3.5 The choice of instrumental energy resolution and *Q*-range in ENS and the possibility to focus on biomacromolecular dynamics in H_2O solutions (Gabel 2005)

In the case of complex biological samples (proteins in hydrated powder, solution or embedded in membranes, entire biological cells) one is often confronted with the problem of accurately interpreting and attributing the origin of the (incoherent) elastic signal as a function of *Q* and temperature. A major problem is the **de-convolution of the signal stemming from protons in the biomacromolecules from those in the aqueous environment.** Usually D_2O was the solvent of choice to minimize incoherent signal from water in order to focus on biomacromolecular dynamics, rather than the more "natural" solvent H_2O. Another problem is the **interpretation of the elastic intensity in terms of intramolecular dynamics (*e.g.* atomic MSDs) of biomacromolecules in solution** since these motions are convoluted by the global (translational and rotational) macromolecular motions.

In order to address these issues quantitatively, I have applied the convolution formalism (Eq. 3.1) to several "model" scattering laws, representing proteins in D_2O and H_2O solutions (Gabel 2005). In these (simplified) scattering laws, the intramolecular mean square displacements $\langle u^2 \rangle$ of the proteins are described by the Gaussian approximation, $S_{Gauss}(Q,\omega) = \exp\left(-\frac{1}{6}\langle u^2\rangle Q^2\right)\delta(\omega)$, while their global motions are described by translational diffusion with diffusion coefficient D_1. The water solvent motions are described by translational diffusion (diffusion coefficient D_2):

$$S_{theo}(Q,\omega) = p_1 \exp\left(-\frac{1}{6}\langle u^2\rangle Q^2\right)\frac{1}{\pi}\frac{D_1 Q^2}{\omega^2 + (D_1 Q^2)^2} + p_2\frac{1}{\pi}\frac{D_2 Q^2}{\omega^2 + (D_2 Q^2)^2} + B(Q,\omega) \qquad \text{(Eq. 3.5)}$$

$B(Q,\omega)$ is the inelastic background and p_1 and p_2 stand for the proton populations in the proteins and the H_2O solvent, respectively. Using Eq. 3.1 with a Lorentzian resolution function, $R(\omega) = \frac{1}{\pi}\frac{\Gamma}{\Gamma^2 + \omega^2}$, equation 3.5 yields the measured elastic intensity

$$EI = S_{meas}(Q,0) = p_1 \exp\left(-\frac{1}{6}\langle u^2\rangle Q^2\right)\left(1 + y_1 Q^2\right)^{-1} + p_2\left(1 + y_2 Q^2\right)^{-1}, \quad y_i = \frac{D_i}{\Gamma}; i = 1,2 \qquad \text{(Eq. 3.6)}$$

The factors y_i are **very useful parameters** to evaluate the contributions of the global macromolecular and solvent diffusive motions to the *EI* as a function of *Q*. If they are in the range $0.1\text{Å}^2 < y < 10\text{Å}^2$ (instrument energy resolution of the order of the quasielastic broadening at $Q = 1\text{Å}^{-1}$), the elastic intensity will be sensitive to them, *i.e.* it will vary as a function of *Q*. If ***y*** **<**

0.01Å² or y > 100Å², the instrument will **not be able to detect the diffusive motions**: in the first case, the atoms will appear to **stand still**, and in the latter, the scattered signal will appear as a **flat background**.

Figure 3.2 shows the logarithms of the measured normalized *EI* for the model system ''proteins in H_2O'' (Eq. 3.6). The model parameters are $\langle u^2 \rangle$=0.5 Å², p_1=0.15, p_2=0.85, D_1=5·10^{-11}, 5·10^{-12} and 5·10^{-13} m² s^{-1}, D_2=2.5·10^{-9} m² s^{-1} and Γ=0.5, 5 and 50 μeV. The curves are grouped according to their y_2-value: 3.3Å² (crosses), 33Å² (plusses) and 330Å² (dashes). Each group is subdivided into three branches, representing different y_1-values (cf. Table 3.1), in increasing order from top to bottom in each group. The single uppermost straight thick line represents the pure intramolecular motions $\langle u^2 \rangle$=0.5 Å² without diffusive contributions (y_1=y_2=0, p_2=0). The thick lines represent linear fits in the range 2 Å^{-2} < Q^2 < 4 Å^{-2} to extract MSDs.

Γ/y_2	0.5 μeV/330			5 μeV/33			50 μeV/3.3		
y_1	6.6	0.66	0.066	0.66	0.066	0.0066	0.066	0.0066	0.00066
$\langle u^2 \rangle$ (1–2 Å^{-2})	4.20 ± 0.07	2.55 ± 0.02	0.91 ± 0.00	2.82 ± 0.04	1.26 ± 0.02	0.95 ± 0.02	2.25 ± 0.04	2.04 ± 0.04	2.00 ± 0.00
$\langle u^2 \rangle$ (2–4 Å^{-2})	2.40 ± 0.03	1.85 ± 0.01	0.85 ± 0.00	1.88 ± 0.02	0.93 ± 0.00	**0.65 ± 0.00**	1.29 ± 0.01	1.08 ± 0.01	1.06 ± 0.00

Table 3.1 : Extracted intramolecular MSDs for the model system « proteins in H_2O solution », Eq. 3.6, for different values y_1 (proteins) and y_2 (water), see text. The fat entry defines the best experimental conditions to focus on intramolecular dynamics.

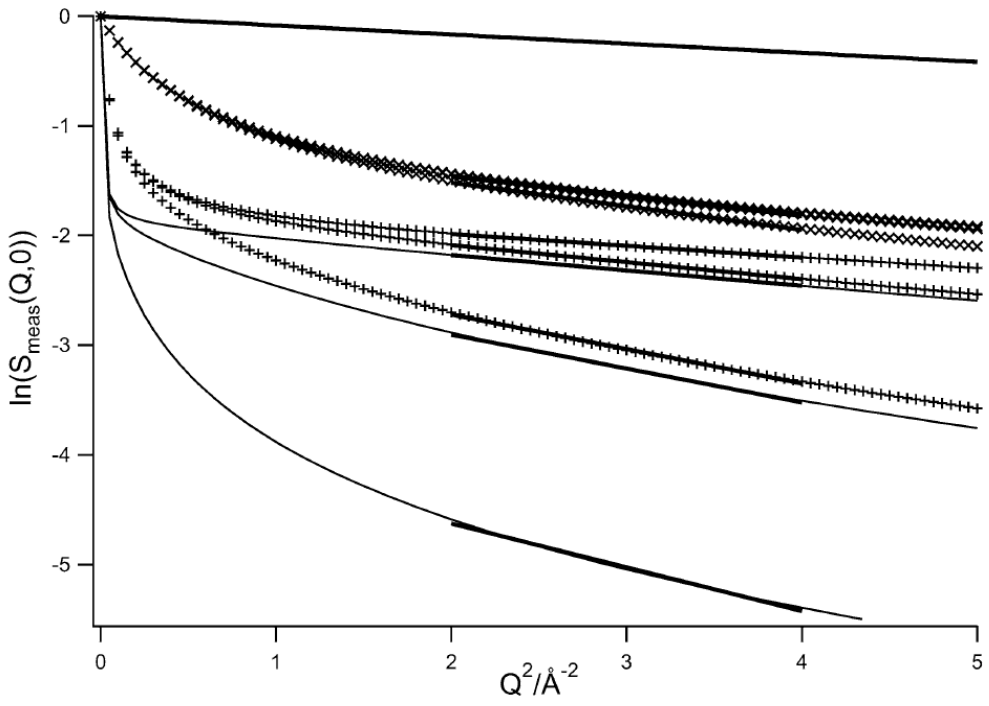

Figure 3.2: Logarithms of the measured normalized EI for the model system ''proteins in H_2O'' (Eq. 3.6). The isolated thick line in the top is the elastic signal of a pure Gaussian scattering law.

The thick straight lines in Fig. 3.2 illustrate the deviations between the elastic signal of a pure Gaussian law for protein intramolecular motions, and the same law in the presence of

translational diffusive motions, both from the protein and the surrounding solvent. The differences are two-fold and both depend on p_1, p_2, and y_1 and y_2:

1) the **slope** of the linear fit (= measured MSDs)
2) the **absolute intensities**

A major quantitative result of this work are the **optimal experimental conditions (energy resolution and *Q*-range) to focus on intramolecular motions of biomacromolecules in solution**: y_1=0.0066Å^2 and y_2=33Å^2, corresponding to a protein diffusion coefficient D_1=5·10^{-13} m^2 s^{-1} and a water diffusion coefficient D_2=2.5·10^{-9} m^2 s^{-1}, respectively, on an instrument with an energy resolution half width at half maximum (HWHM) of 5 μeV yield the best results (fat entry in Table 3.1). However, even in that case, and for the **favorable *Q*-range** 2 < Q^2[Å^{-2}] < 4, there is a residual "contamination" from the diffusive motions which overestimates the "real" intramolecular atomic motions by about 30 %.

In conclusion, it is possible to choose the instrumental energy resolution **and** *Q*-range in a manner to **reduce contributions from bulk solvent water (even H$_2$O)** to the elastically scattered signal and to focus on macromolecular dynamics. A typical instrument which is able to do this is the backscattering spectrometer IN13 at the Institut Laue-Langevin (ILL), Grenoble. Its characteristics (HWHM of 4 μeV, *Q*-range up to 5 Å^{-1}) make it perfectly suited to focus on protein dynamics, both in aqueous solution as well as in living biological cells, if the intracellular water dynamics in the latter are not slowed down too much with respect to free bulk water.

More problematic is the de-convolution of protein intramolecular motions (*e.g.* MSDs) from their global diffusive motions (D_1): they require that the parameter y_1 be inferior to about 0.01. This is the case for a protein translational diffusion coefficient D_1 < ~ 10^{-12} m^2 s^{-1}. This is, however, not obvious, in particular for small and medium-size (< 100 kDa) proteins in dilute solutions that have diffusion coefficients of ~10^{-10} m^2 s^{-1} (Serdyuk *et al.* 2007). It is up to the user to make sure that the **condition y_1 < 0.01Å^2** is fulfilled in the biological system studied when the variations of the *EI* are to be interpreted as **intra**molecular macromolecular motions. Working with **large proteins** (> 100 kDa), in **crowded protein solutions**, in **D$_2$O** (increased viscosity) and/or in **entire biological cells** may help to reach that condition by slowing down protein global translational diffusion.

The effect of protein global **rotational** motions and **deviations from the Gaussian approximation** of the intramolecular dynamics are discussed in Gabel (2005). They are not elaborated further here.

3.6 A combined ENS data analysis from two backscattering spectrometers with a tenfold difference in energy resolution: a study of D-CPC hydration water on IN13 and IN16 (Gabel and Bellissent-Funel 2007)

Apart from enabling to focus on specific types of motions in complex biological systems (see preceding section), the convolution formalism allows **extraction of quasielastic parameters** such as characteristic times, *EISF*s etc. and **interconnects both types of experiments** more intimately than a simple interpretation of MSDs from the Gaussian approximation. Moreover, MSDs depend in general (at higher temperatures where quasielastic contributions are present) on the instrumental energy resolution.

In order to test and apply the convolution formalism, a lyophilized powder of deuterated C-phycocyanin (CPC), hydrated in a D-trehalose-H_2O mixture (130 mg dry protein, 55 mg dry trehalose, 76 mg H_2O) was studied by elastic scans on the ILL backscattering instruments IN13 and IN16 in the temperature range from 20 to 320 K. The aim of the project was to find a scattering law that describes the elastically scattered signal (dominated by the hydration water) **simultaneously well** on both instruments and to relate it to QENS data from literature.

In a first attempt the data were fitted by the traditional Gaussian approximation in the Q-range $0.9 < Q^2[Å^{-2}] < 3.3$. The resulting apparent MSDs are plotted in Fig. 3.3.

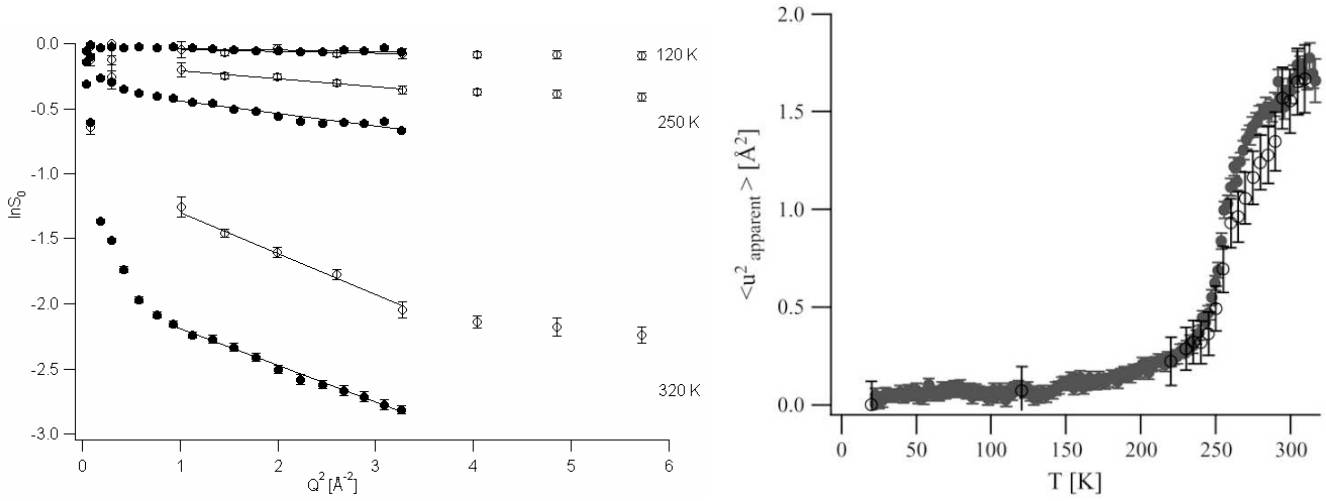

Figure 3.3: *Left*: Logarithm of the elastic intensity as a function of Q, normalized to the lowest temperature (20 K) at 120, 250 and 320 K (from top to bottom). Open circles represent IN13 data, filled circles IN16 data. The lines represent fits to extract atomic MSDs by the Gaussian approximation. *Right*: Apparent atomic MSDs from D-CPC hydrated water. IN13 (open black circles) and IN16 (filled gray circles), respectively.

Several conclusions can be drawn from Figure 3.3:

- D-CPC hydration water scatters mainly **vibrationally below 200 K**
- above about 220 K, the **Gaussian approximation breaks down** since the MSDs are different from both instruments

- the **Guinier-plot contains more information** than simply MSDs (slope in a given *Q*-range)

In order to improve the unsatisfactory results from the Gaussian approximation, the data were fitted using the convolution formalism (Eq. 3.1). A theoretical scattering law $S_{theo}(Q,\omega)$, describing the hydration water dynamics, was sought from the elastic data from both instruments. An essentially important requirement was that a single scattering law fitted the **datasets from both instruments simultaneously well at each temperature**. The scattering law capable to fulfill these conditions was a combination of a population *p* of atoms diffusing in a sphere (simplified after Volino and Dianoux 1980) and a population 1-*p* diffusing translationally freely:

$$S_{theo}(Q,\omega) = (1-p)\left[\left(\frac{3j_1(qa)}{Qa}\right)^2\delta(\omega) + \left(1-\left(\frac{3j_1(qa)}{Qa}\right)^2\right)\frac{1}{\pi}\frac{\Gamma}{\Gamma^2+\omega^2}\right] + p\frac{1}{\pi}\frac{DQ^2}{\left(DQ^2\right)^2+\omega^2} \quad \text{(Eq. 3.7)}$$

D is the translational diffusion coefficient, Γ is the relaxation parameter describing the confined diffusion in a sphere of radius *a*, and j_1 is the spherical Bessel function of 1st order. Convolution with the respective instrumental Lorentzian energy resolution Γ_{Lor} yielded (after normalization):

$$S_{theo}(Q,\omega) = (1-p)\left[\left(\frac{3j_1(qa)}{Qa}\right)^2 + \left(1-\left(\frac{3j_1(qa)}{Qa}\right)^2\right)\frac{1}{1+\frac{\Gamma}{\Gamma_{Lor}}}\right] + p\frac{1}{1+\frac{D}{\Gamma_{Lor}}Q^2} \quad \text{(Eq. 3.8)}$$

Figure 3.4 shows the **simultaneous fit** (by only applying the respective instrumental Γ_{Lor}) of Eq. 3.8 to the IN13 and IN16 data at 320 K. The data are described in a more satisfactory way than with the Gaussian approximation: larger *Q*-ranges, differences in the absolute intensities are captured and exploited.

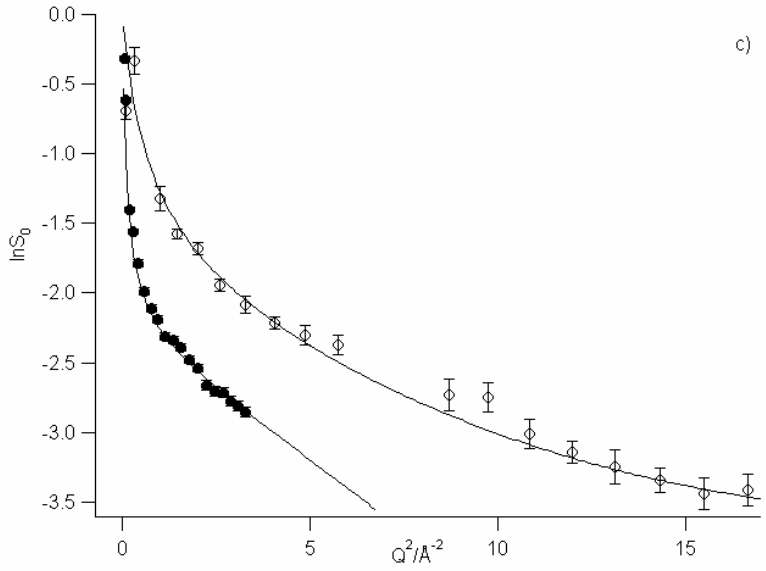

Figure 3.4: Fit of IN13 (open circles) and IN16 (filled circles) data at 320 K using Eq. 3.7, convoluted with the respective instrumental energy resolution.

A fit using Eq. 3.8 was possible to all datasets above 240 K. The temperature dependence of the parameters extracted is shown in Fig. 3.5.

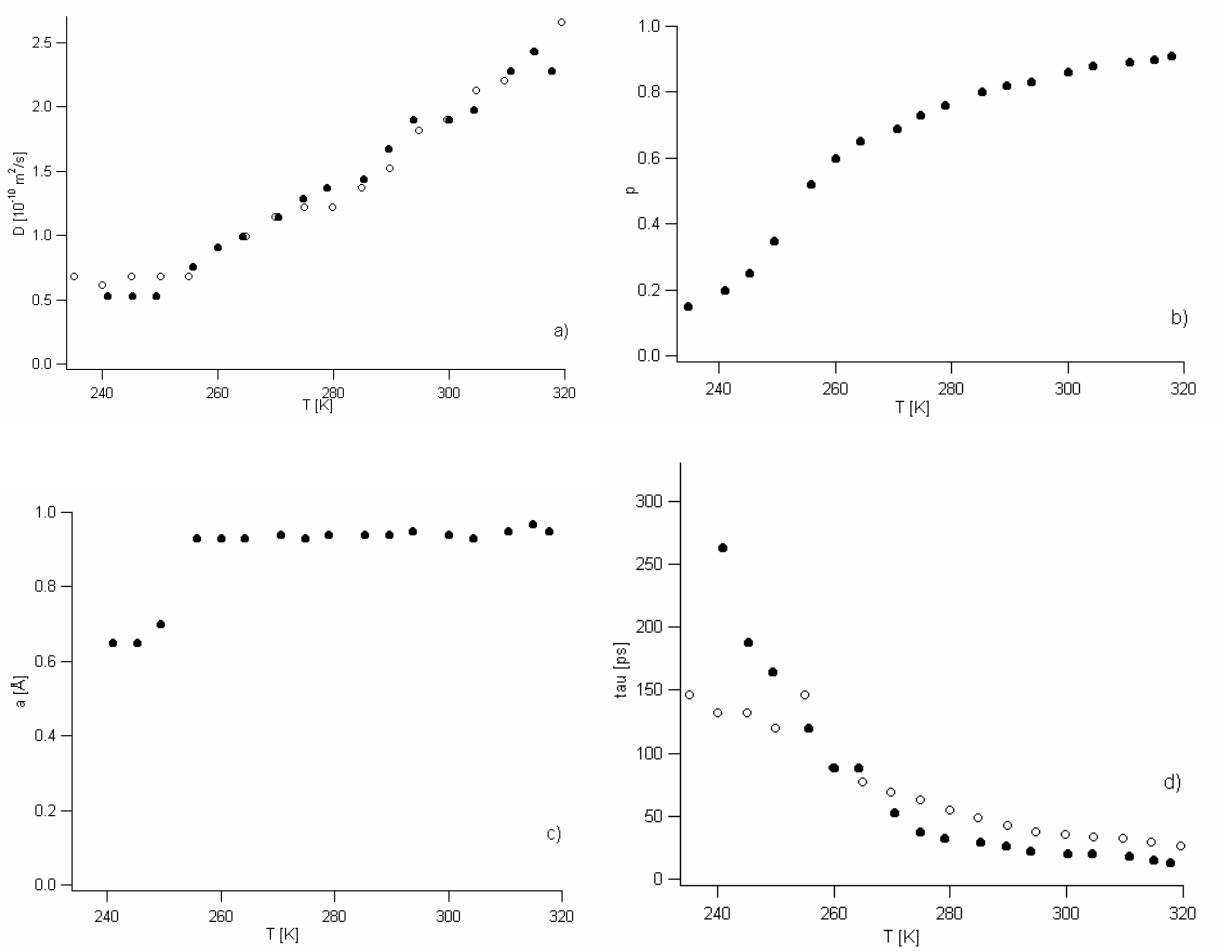

Figure 3.5: a) Diffusion coefficient *D*, b) population *p*, c) sphere radius *a*, and d) relaxation parameter $\tau=1/\Gamma$ as a function of temperature from IN13 (open circles) and IN16 (closed circles). Where only filled circles are shown, the fit parameters on IN13 and IN16 were identical.

The results of the work are discussed in detail in Gabel and Bellissent-Funel (2007). They are briefly summarized here:

- hydration water dynamics (diffusion coefficient) **are slowed down by a factor of ~15** in the presence of a protein surface and trehalose with respect to bulk water at the same temperature
- the convolution formalism (Eq. 3.1), applied to several instruments simultaneously is a **very powerful and selective tool** to determine a scattering law for the atomic motions in biological samples
- the convolution formalism can be used to **extract quasielastic parameters rapidly over a wide temperature range**

3.7 Analogies and differences between the *EISF* and Small Angle Scattering

In Gabel (2005) some mathematical analogies between the Gaussian approximation used in ENS and the Guinier approximation in SAS were pointed out and discussed. This analogy had been employed to interpret atomic MSDs extracted from ENS in different *Q*-ranges in terms of two different populations undergoing motions with different amplitudes (Réat *et al.* 1997). It allows, in principle, to interpret ENS data at higher *Q*-values in terms of mean square displacements of populations of atoms with small(er) amplitudes. While this analogy is very tempting, it needs to be applied with great care due to the following reasons:

1) **The formal analogy is strictly valid only between the *EISF* and the SAS data as a function of *Q***. It is therefore, in general, not applicable to the elastic intensity (*EI*) at higher temperatures (T>200K) that is contaminated by quasielastic contributions (see preceding sections). An accurate application of the analogy requires the demonstration that these contributions can be neglected at a given energy resolution and in a given *Q*-range. Under certain (very restrictive) experimental conditions (section 3.5) one can focus specifically on protein **intra**molecular motions, even in an H_2O solvent.
2) The separation of an ENS signal into several *Q*-ranges (linear zones in a Guinier-plot ln*I vs.* Q^2) and their consecutive **interpretation in terms of several (monodisperse) populations with different, specific, MSDs needs to be used with care**. In general, such an interpretation is difficult in SAS since two monodisperse particle populations with very distinct radii of gyration and a monodiserse population of, *e.g.*, elongated ellipsoidal particles produce very similar Guinier plots (Guinier and Fournet 1955). This holds true for ENS where the situation is in general worse since atomic motions in biomacromolecules display a **great degree of polydispersity**.
3) Another fundamental difference between the SAS and ENS formalism is the different weight of particles of different sizes and atomic motions of different amplitudes, respectively: while the contribution to the signal grows with the square of the volume for particles (of the same average chemical composition) in SAS, it is independent from the volume sampled by the atomic motions in ENS (since only single atoms scatter neutrons incoherently). In other terms, there is a **"thinning-out effect"** in ENS (Gabel 2005) that implies that the difference between atoms of small and large motions is less pronounced than it is for particles of different size in SAS.

Perspectives

Perspectives of my work concern the improvement of the developed protocols and their application to ambitious topics in structural biology. Future inhouse projects at IBS Grenoble include **large molecular assemblies** (Dr. Bruno Franzetti) **unfolded proteins** (Dr. Martin Blackledge) and **membrane proteins** (Dr. Christine Ebel). The following "Perspectives" chapters provide more details on future developments and projects. The chapter numbering corresponds to the one of the "Results" part of the thesis.

1. Limitations, possible improvements and applications of the SAS-NMR protocol

The protocol in its present form (Gabel *et al.* 2006) and implemented in *ARIA-CNS* (Gabel *et al.* 2008) requires a number of technical improvements in order to become more **user-friendly** and **generally applicable**. For example, the protocol in its present form requires that the users make modifications in several *ARIA-CNS* modules. Furthermore, potential systematic errors in the extraction of A_{exp}, B_{exp}, and C_{exp} and the presence of a hydration shell require a **grid-search** around the target values. Sections 1.1 through 1.3 are dealing with the more technical programming tasks to remedy these shortcomings. Fundamental questions, applications and possible generalizations of the approach are addressed in sections 1.4 through 1.11.

1.1 Accurate extraction of target parameters A_{exp}, B_{exp}, and C_{exp} from experimental SAS curve

The target parameters are extracted by a least χ^2-fit of a polynomial function in a restricted Q-range from the experimental SAS curve, starting with A_{exp} in the Guinier-range and moving on to higher Q-values including progressively the higher parameters. In its present form it is done in an interactive mode by the user. The fundamental problem is the accurate extraction of the parameters and the optimization of the Q-range to be used for the least χ^2-fit. This is basically the following mathematical problem: **how well does a fit of the data by a truncated polynomial in a limited Q-range extract the coefficients of the underlying infinite Taylor series that describes the scattering curve completely?**

1.2 Accurate calculation of the b_i for heterogeneous bodies

At the moment, the protocol is limited to homogeneous bodies of the same scattering length densities, *i.e.* the b_i of all atoms are identical. This is a good approximation for *A*, *B*, and *C* in the *Q*-ranges used ($QR_g \leq$ 2-3). However, in some cases, in particular for biomacromolecules showing a heterogeneous scattering length density on a ~ 10-30 Å lengthscale (phosphate groups *vs.* base pairs in DNA and RNA, hydrophobic core *vs.* hydrophilic surface residues in proteins), the present approximation might not be good enough. The quality of the homogeneous approximation needs to be checked quantitatively by comparing the back-calculated scattering curves from several systems using either b_i=const or the accurate scattering lengths. If necessary, the more accurate approach needs to be incorporated into the protocol at the expense of computational time. Using the accurate scattering lengths will allow **refinement of heterogeneous system** (*e.g.* protein-DNA/RNA complex). It will affect in particular the position of the pivotal point between both rigid bodies (Eq. 1.2) and the positions of the $B=B_{exp}$ and $C=C_{exp}$ lines. In this case, the user needs to provide information on the type of radiation and solvent scattering length density (H_2O, D_2O, presence of solutes etc.)

1.3 Inclusion of a hydration shell

The protocol in its present form treats the rigid bodies as "naked", *i.e.* atomic models deprived of a hydration shell. However, as has been shown by comparing SAXS and SANS data in H_2O and D_2O (Svergun *et al.* 1998), proteins in aqueous solution possess, in general, a hydration shell of different (physical and scattering length) density than the bulk solvent. It is incorporated in programs that calculate the scattering curve of an atomic structure such as CRYSOL/CRYSON (Svergun *et al.* 1995). The inclusion of a hydration shell for the calculation of *A*, *B*, and *C* will modify both the inter-body distance (Eq. 1.2) as well as the positions of the domains. In practice, a hydration shell can be added in two ways: i) using the spherical harmonics approach (Stuhrmann 1970b; Svergun *et al.* 1995) and adding a 3 Å thick shell from the outermost atoms of the PDB structure or ii) adding directly water molecules on the surface using the *ARIA-CNS* program suite. In both cases, the hydration shell needs to be properly weighted as a function of solvent and radiation used. Unfortunately, it is improbable that a single type of hydration shell applies in all cases: *e.g.*, the hydration shell of highly charged RNAs or DNAs will probably have different properties than those of proteins close to their pI.

1.4 Stability and independence of A_{exp}, B_{exp}, and C_{exp}

A crucial point of the protocol is the stability of the solutions as a function of (systematic) errors of the target values. These occur by the fit of a truncated polynomial (see above) and have been partially discussed in Gabel *et al.* (2006). In this context, it is imperative to understand if and how a systematic error in one of the values (*e.g.* B_{exp}) affects the other parameters (*e.g.* C_{exp}) and if it can be counterbalanced. In other words, can two systematic errors compensate in the sense that they would still yield an accurately refined structure? There are indications that this is true, related to the properties of the fit of a Taylor series (Gabel *et al.* 2006, Fig. 6). This study is important for the stability of the protocol and needs to be carried out systematically.

1.5 Efficiency of the protocol as a function of domain anisotropy and inter-domain distance

While it is obvious that two spherically symmetric bodies cannot be refined further than their distance by SAS, the efficiency of the protocol will improve with domain anisotropy. It would be interesting to formulate this statement in a more quantitative manner as a function of a domain **"anisotropy parameter"** (both in terms of deviations from a spherical shape and in size differences between both bodies) that can be used to assess the refinement possibilities and accuracy to be expected *prior* to the actual small angle scattering experiments. A solution to this problem would be of general interest also for two bodies of arbitrary orientation.

The capacity to refine a two-body system will not only depend on the *shape* of both partners but also on their mutual *distance*: it is immediately clear that in the case of extremely distant partners (inter-partner distance >> dimension of each partner), their mutual position cannot be easily refined beyond fixing their distance. On the other hand, for very close partners (with fixed relative orientation), the mutual position cannot be refined very efficiently, either. Therefore, there probably exists an optimal distance regarding the refinement efficiency. To determine this distance and put it into quantitative relationship with the partners' individual shapes and orientations would be very useful.

1.6 Application to macromolecular complexes composed of more than 2 bodies

An application to macromolecular complexes with more than two bodies is in principle possible. It would be a generalization of the *triangulation approach* that has been used to chart the quaternary arrangement of protein subunits in the ribosome (*e.g.* Engelman and Moore 1972; Hoppe *et al.* 1975). However, in the case of triangulation only the **inter**particle interference term

between the two subunits of interest is studied by using a "*mixed isomorphous replacement*" approach (Hoppe 1973; May 1991). The present approach uses **both** the **intra- and inter**particle interference terms of the two subunits of interest. They need to be recorded, ideally without contamination by the remainder of the structure. The latter point is difficult since proteins cannot be matched homogeneously (*e.g.* Zaccai and Jacrot 1983). RNA and DNA, on the other hand, can be matched more homogenously. In the general case, labeling schemes similar to the "*Glassy Ribosome*" approach (Nierhaus *et al.* 1983) minimizing contributions from the partners of no interest should be used. Since no data beyond 2-3 times the Guinier range is needed, there may be a chance that particle inhomogeneities at the matchpoint will not play a role.

1.7 Generalization of the developed protocol to arbitrarily oriented partners

While the protocol has been tailor-made to fit the needs of a given setup (domain orientation known from NMR RDCs plus additional SAS information), it would be interesting to know how far the developed concepts and results can be applied to a two-body system with **arbitrarily** oriented partners. In many cases, this general situation is dealt with by a grid-search approach (*e.g.* Konarev *et al.* 2001) with the drawback that the best refined structure depends on the initial conditions (positions and orientations of partners) and that secondary solutions (so-called "false positives") are possible. The main potential benefit of a generalization of the present approach, if possible, would certainly be to identify **families of structures** that are compatible with the SAS data and that can be restricted/refined further by using additional structural restraints from complementary techniques (cross-linking, mutagenesis etc.). These restraints could define **restricted two-dimensional regions** in the (θ,ϕ)-plane that **limit the positions of the two rigid bodies**.

1.8 Existence and nature of structural solutions

The ensemble of structures obtained as a function of A_{exp}, B_{exp}, and C_{exp} is given by the solution of the equations in appendix 2. It confines the centers of scattering length density of both bodies to great circles on the surface of a sphere with a diameter that corresponds to the inter-body distance. So far, the solutions of the defining equations are obtained numerically (graphically). A fundamental question that arises is: **can they be obtained analytically?** In other words, can the existence and nature of the structural solutions of a rigid two-body system be predicted and written down explicitly as a function of the geometry of its two components and the shape of the scattering curve? In its present description, this mathematical problem belongs to the field of

geometric algebra. A general mathematical theorem states that a system of polynomial equations of the same degree has always a one-dimensional complex solution. In our case, however, the **real** solutions of the equations are of interest. A possible strategy and keyword to tackle this problem could be *Bézout's theorem* (*e.g.* Brieskorn and Knörrer 1981). In any case, the general solution of this problem will require collaboration with partners with excellent mathematical skills.

1.9 Application to large molecular complexes (collaboration Dr. Bruno Franzetti)

De novo domain refinement using RDC restraints from larger macromolecular complexes (> 80 kDa) is presently out of reach for NMR even if high-resolution structures of the subunits are available (Tugarinov and Kay 2003). As a consequence, the use of the higher terms of the refinement potential (*B* and *C*) is impossible for these systems at the present state-of-the-art. However, the term defining the **inter-domain distance** (*A*) as well as a large number of other restraints such as **cross-linking, mutagenesis, PREs, NOEs** etc. are already implemented in the protocol and can be used easily (Gabel *et al.* 2008). Applied under these conditions, the protocol has the advantage (with respect to rigid-body modeling) to provide a family of structures that are compatible with all structural restraints and which can be further refined, *e.g.* by scoring against the full SAS curve.

It is planned to apply this approach to large molecular assemblies such as the family of the aminopeptidase TET (Franzetti *et al.* 2002). First results regarding the thermo-stability and thermo-activation have already been obtained (Durá *et al.* 2009). Important structural questions regarding substrate binding, processing and release as well as putative concomitant conformational changes can be addressed readily by using SAXS/SANS. The thermo-stability (fast tumbling rates) as well as the symmetry of these systems makes them particularly well-suited for structural studies combining SAS and NMR.

1.10 Taking flexible parts into account (collaboration Dr. Martin Blackledge)

By definition, the protocol assumes that both bodies be rigid, *i.e.* that no flexible parts occur in the complex to be refined. This condition is unfortunately not always fulfilled (*e.g.* flexible linker, header or tail). How to incorporate such modifications into the present protocol needs to be worked out both theoretically and practically. It is evident that these parts cannot be incorporated as rigid in a given conformation since it would alter the parameters *B* and *C* in a very biased way. One possibility to tackle this problem would be to take the flexible parts (tail

and header) into account as a **weighted ensemble of conformers** (Bernadó *et al.* 2005; Bernadó *et al.* 2007) from which new parameters *A*, *B*, and *C* can be calculated. The difficulty lies then in finding a right weighing procedure of the conformations of the flexible parts and their overall weight with respect to the rigid parts of the bodies. It is planned to address these issues both theoretically and experimentally in close collaboration with Dr. Martin Blackledge (IBS Grenoble).

1.11 Alternative descriptions: spherical coordinates, multipole expansions, principal axes systems and pair distance distribution functions

It is conceivable that the whole approach of expanding the scattering data into a power series and using Cartesian coordinates for the atomic positions to calculate the target parameters is not adapted to the geometry of the problem. **Polar coordinates** as used in other approaches (*e.g.* spherical harmonics and a multipole expansion) may be better suited and the resulting equations easier to deal with. As an alternative, if Cartesian coordinates are to be kept, a **transformation into the principal axes system** of the initial reference frame should equally yield considerable simplifications of the equations. However, it is unclear how the change of the principal axes of the two-body system as a function of the positions of the individual bodies is to be dealt with. Another alternative would be to calculate a **pair distance distribution function** from atomic coordinates for a given domain configuration during the refinement process and compare it to the one determined from the experimental data. This has been demonstrated for simple geometric bodies (Glatter 1979). However, a target function and driving force would need to be developed specifically for this approach.

2. Study of unfolded proteins by SAXS/SANS and NMR

2.1 Accurate structural information at high *Q*

A major complication with small angle scattering analysis of unfolded proteins is that they scatter to high angles (~ 1.0 $Å^{-1}$) with low intensity. The weak signal at higher angles contains precious information on structural features at shorter distances (*e.g.* surface and hydration shell properties, and cross sectional information). Inaccurate solvent subtraction of one percent or less can lead to erratic data interpretation, in particular in Kratky plots. In order to carry on our study on urea-denatured ubiquitin, data need to be recorded to high angles with good signal/noise. The approach to maximize the *Q*-range has been used in the polymer community by using specific instruments, *e.g.* D1B at ILL (Rawiso *et al.* 1987) where data on polystyrene was measured up to very high angles (Q=1.5 $Å^{-1}$). This latter approach in combination with concentrated solutions (>> 10 mg/ml) should yield accurate exploitable data in our case (if inter-particle effects can be neglected at high angles).

2.2 Complementarities of SAXS/SANS for cross-sectional analysis

While the complementarities of SAXS and SANS have been exploited routinely in the polymer sciences (Schurtenberger 2002 and references therein), its potential has been hardly applied to unfolded proteins. However, a combination of SAXS and SANS would be of particular benefit in a **cross-sectional analysis** exploiting different contrast conditions to study the following issues: i) nature and geometry of the **hydration water shell** close to the protein surface and its comparison to the folded case, ii) **side-chain flexibility** and iii) **interaction** of smaller solute molecules with the protein. All three effects are particularly pronounced in unfolded proteins (with respect to globular, folded proteins) due to the increased surface to volume ratio. Results from Gabel *et al.* (2009) suggest that some of the usually accepted concepts in globular proteins such as a 3 Ångström thick, 10% denser (than bulk solvent) (Svergun *et al.* 1998) cannot be applied without further questioning to chemically denatured proteins.

Several approaches are possible to address the raised issues and analyze future high-angle data:

1) An explicit analytical function describing more sophisticated models, including persistence lengths, cross-sectional terms etc. (*e.g.* Rawiso *et al.* 1987; Pérez *et al.* 2001)
2) An ensemble of atomic-resolution conformers, either with modeled (solvent and/or small solute) molecules at specific interaction sites or by using a hydration shell of flexible density as a phenomenological fit parameter (Svergun *et al.* 1995) in order to model side-chain flexibility and/or solutes associated with the unfolded chain.

2.3 Combination with NMR structural restraints

In the last few years the power of combining SAS with NMR restraints for unfolded proteins has been realized and is reflected in a growing number of publications (*e.g.* Bernadó *et al.* 2005; Bernadó *et al.* 2007; Jensen *et al.* 2009; Bernadó and Blackledge 2009). However, so far this approach has been mainly limited to using SAXS. Supplementing the NMR data available on denatured ubiquitin by SANS data would help to refine the structural models of the unfolded proteins in terms of side-chain flexibility and solvent properties close to the protein surface (see preceding section). This information is crucial for the accurate fit of back-calculated average scattering curves of unfolded protein ensembles in an intermediary and high *Q*-range (>0.3Å^{-1}).

3. Elastic neutron scattering analysis on several instruments; analogies to SAS and possible combination with NMR

3.1 Application of convolution formalism to more than 2 instruments

Comparison of elastic data as a function of *Q* from two instruments differing by a factor of 10 in instrumental resolution allows extracting as well as rejecting scattering laws and quasielastic parameters (Gabel 2005; Gabel and Bellissent-Funel 2007). One limitation of this approach is its insensitivity to describe the variation of quasielastic scattering components that are more than about a factor of 10 narrower or wider than the instrumental energy resolution. Therefore, the power and selectivity of this approach can be improved by **comparing data from a larger number of instruments (three or more)** that differ pair-wise by at least an order of magnitude in resolution. Since each instrument introduces, in a certain manner, an independent data set, this combination allows extraction of scattering laws with a larger number of fit parameters. It is in principle applicable to all types of sample states (protein powders, solutions, entire biological cells …).

3.2 Library of convoluted scattering laws

A **library of convoluted scattering laws** describing atomic motions in biological systems would be of enormous practical use to the neutron community using elastic scans. It should include a number of translational, rotational, and vibrational laws as well as more sophisticated models (stretched exponential, jump-diffusion, confined diffusion etc). All analytically "convolutable" scattering laws (as well as products and linear combinations of them) can be written down explicitly, taking different experimental resolution functions into account (triangular, Gaussian, Lorentzian). This would allow users with an educated guess of dynamics in their samples to choose the optimal instrument (both in terms of energy resolution and *Q*-range) in order to focus on specific types of motion. A **graphical representation of the predicted elastic intensity** as a function of *Q* with interactively tunable parameters (*e.g.* diffusion coefficients, instrument resolution etc.) would be very useful and practical.

3.3 Comparison/combination of elastic and quasielastic spectra

An elastic scan contains less information than QENS or INS spectra and is mainly used for its rapidity and model-free analysis of weakly scattering samples. A systematic combination of both types of experiments would be very advantageous. Unfortunately, up to date, this combination is not carried out routinely; in most cases QENS and ENS experiments measure different quantities

or use different models (*e.g.* diffusion coefficients and residence times *vs.* atomic MSDs). I believe that the convolution method for elastic scans is an ideal means to overcome these incompatible analyses since it can be interpreted in terms of parameters equally employed in QENS. A good experiment would consist in recording **elastic data at several temperatures over the whole experimental *Q*-range and supplementing/comparing them with QENS data** with high quality (long exposure time) **at a few selected *Q* values**. Finally, both types of datasets would need to be interpreted in terms of a single scattering law that fits both simultaneously well.

3.4 The problem of multiple scattering

Elastic scans at low *Q*-values will be particularly influenced by multiple scattering (Sears 1975; Settles and Doster 1997; Gabel 2000; Zorn 2009). While some programs exist that correct for this defect for QENS data they will need to be adapted for elastic scan data treatment. As in the case of the existing programs, this will probably require an assumption of a scattering law *a priori* that has to be convoluted and weighted for the back-calculation of the elastic signal as a function of *Q*.

3.5 Effect of coherent scattering

In the interpretation of ENS and QENS data, the coherent contributions are in general neglected. They can, however, have a differential behavior as a function of hydration state as was demonstrated in dry and D_2O-hydrated deuterated CPC (Bellissent-Funel *et al.* 1997). In the light of these results it would be worthwhile to investigate if normalization of the scattered intensity of a sample to the lowest temperature is indeed able to eliminate contributions of coherent scattering to the variations of the elastically scattered intensity as a function of temperature. An upper estimation of the coherent contribution to the observed variation of the elastic intensity should be calculated as a function of sample composition and physical state (powder, solution).

3.6 Elaboration of analogies between SAS and ENS

In Gabel (2005), the analogies and differences between the atomic mean square displacements and the radii of gyration in SAS were elaborated. In my opinion, it would be worth and interesting to push these analogies further. It is, in principle, conceivable to apply the whole arsenal used in small angle scattering data interpretation (monodisperse and polydisperse cases) on *EISF* data. However, as already pointed out (Gabel 2005), motions of larger amplitude contribute less than motions of smaller amplitude (**"thinning-out" effect**). Therefore, a

transcription of SAS results to EISF data cannot be done 1:1. It will require different weighing mechanisms for motions that sample different "volumes" and needs to be developed mathematically in a rigorous way. Since motions of atoms in a biological sample display a great degree of polydispersity, **methods developed in SAS to interpret particle-size distributions** (Glatter 1980b) should turn out to be particularly helpful, *e.g.* a Gaussian distribution of spheres of different size (Glatter 2002).

3.7 Combination of neutron spectroscopy with NMR relaxation

An ambitious perspective of my work regarding molecular dynamics would be a possible combined interpretation of neutron spectroscopic and NMR relaxation data. While both techniques access the pico- to nanosecond timescale of atomic motions on an Ångström lengthscale, different aspects contribute to the signal from each technique: neutron spectroscopy, on the one hand, measures essentially the space-time Fourier transform of the atomic (nuclear) trajectories (van Hove 1954), weighted by the instrumental energy resolution at a given *Q*-value. NMR relaxation, on the other hand, is based on spectral density of rotational motions of atomic bonds (Lipari and Szabo 1982; Palmer 2001; Brüschweiler 2003). **Both techniques are therefore complementary for the description of molecular dynamics.** Despite this fact, attempts to use a common mathematical formalism to describe molecular motions from neutron spectroscopy and NMR simultaneously are extremely rare (Fleischer and Fukara 1994; Zanotti *et al.* 1999).

4. Low-resolution modeling of membrane proteins using SANS

4.1 Introduction

Membrane proteins (MP) constitute an estimated 30-40% of all proteins in eukaryotic cells (Wallin and Heijne 1998). They fulfill a number of important biological functions such as transport (ions, small substrates…), signaling, reception (optical or chemical), cell-cell recognition and adhesion, energy production and others (White 2009 and references therein) and are a major target for drugs (Korepanova *et al.* 2005). In contrast to water-soluble proteins MPs are notoriously difficult to study by high-resolution structural techniques such as crystallography and NMR (Lacapère *et al.* 2007): in 2009, less than 1% (about 200) out of 60000 structures in the protein data base (PDB) were membrane proteins (White 2009).

Small angle neutron scattering and contrast variation can provide low-resolution models of otherwise inaccessible MP structures and architecture using the arsenal of modern methods presented in chapter 1.1. However, structural studies of protein-detergent (Guo *et al.* 1990) and membrane proteins complexes using SAXS/SANS (*e.g.* Hunt *et al.* 1997; Wang *et al.* 2003; Bu *et al.* 2003; Zimmer *et al.* 2006; Johs *et al.* 2006; Lipfert and Doniach 2007 and references therein) are rather scarce up to date. This is mainly due to two reasons: i) the **particularly difficult biochemical preparation** of biologically active and monodisperse samples (le Maire *et al.* 2000) in sufficient quantities for structural biology and ii) the data analysis due to possibly **incomplete matching** and **residual scattering signal of the detergent moieties** (free micelles and detergent monomers as well as protein-bound detergent).

4.2 A "*via regia*" to low-resolution membrane protein structures: high-resolution detergent models, multi-phase modeling or a "perfect" detergent?

Over the last two years, I have participated in several SANS projects of detergents (mainly in collaboration with Dr. Christine Ebel) with the aim to prepare the structural study of membrane proteins with the arsenal of modern modeling techniques presented in chapter 1.1. Several approaches are conceivable for these projects:

1) *The use of high-resolution models of detergent moieties*. Free detergent and micellar structures display, in general, a (scattering length) heterogeneity on their molecular length-scale. Their scattering signals therefore need to be **accurately subtracted** from the overall scattering curve (protein-lipid/detergent complexes (PDCs) plus free detergent and micelles) in order to model the remaining protein-detergent particles. To achieve this goal requires an accurate

structural model of the free micelles from detergent-only scattering curves. A simple example would be a **core-shell model** but also more sophisticated models using **different phases** (*e.g.* program "MONSA"; Svergun 1999) or **atomic models** built from detergent monomer structures are possible. In addition, the fraction of scattered signal due to micelles from the overall signal (micelles, free detergent and PDCs) needs to be known accurately. Analytical ultracentrifugation (AUC) will be **indispensable** in this approach. A bottleneck of high-resolution modeling in this approach is certainly the detergent bound to the protein which is more difficult to model than free micelles. However, neutron data from MP crystals (Pebay-Peyroula *et al.* 1995) might provide helpful clues and inspire approaches to model their contribution in MP-detergent complexes in solution accurately.

2) Once the signal of free detergent/micelles is accurately subtracted (preceding section) the protein-detergent complexes can be modeled. Only very few examples going beyond a radius of gyration / molecular weight analysis are available to date (Johs *et al.* 2006). A possible way of tackling this problem is the use of **several phases** ("MONSA") which has been applied to protein-DNA or RNA complexes (Putnam *et al.* 2007 and references therein). An important issue is the minimal number of phases to be used. From a point of view of scattering length density, a reasonable number would be **three**: one for the protein, one for the detergent heads and one for the detergent tails. Ideally, SANS data would be recorded for several contrast conditions, masking the protein, the detergent tails or heads. An additional restraint to be imposed and implemented in the data modeling programs is the **connected topology** of the individual phases.

3) From a SANS structural point of view, the perfect detergent would have a **homogeneous scattering length density**, in particular head *vs.* tail. (More important of course are its properties to dissolve membrane proteins from their lipids and keep them in a physiologically active structure!). Structural requirements for a homogeneous detergent can in principle be calculated from the chemical composition, atomic scattering length densities and partial specific volumes. A completely deuterated detergent has a scattering length density that exceeds that of pure D_2O. It can in general not be matched in any fraction of H_2O/D_2O in an aqueous solvent. Commercially available deuterated detergents often have different contrasts head *vs.* tail. To my best knowledge such a perfect detergent, regarding both scattering properties and interaction with solubilized membrane proteins does not (yet?) exist. It would have to be **tailor-made**. The next section presents a favorable candidate in a project I have been involved in.

4.3 Hemifluorinated surfactants, a promising class of detergents for membrane protein studies in solution (Breyton *et al*. 2009)

Hemifluorinated surfactants are promising detergent candidates that appear to combine the structural properties required from a SANS point of view (negligible scattering signal at the match-point) as well as the biochemical "gentleness" regarding the solubility of MPs in a functional and stable state. The results of Breyton *et al.* (2009) are briefly summarized here:

- F_6-Diglu and F_6-Triglu can be described by compact, **globular micelle structures** with radii of gyration of 18 and 20 Å, respectively.
- They have aggregation numbers of 67 and 34 molecules per micelle, respectively
- F_6-Triglu scattering curves below the CMC (at 1.7 and 0.5 mg/ml) **display no structural features** and do not differ from the buffer signal
- Their contrast match point (CMP) is about **45% D_2O**
- **No residual scattering signal** at the CMP from the detergent was visible up to C = 10 mg/ml

Future experiments on several protein systems are planned or under way (OmpX, bacteriorhodopsin). Imposed by the particular CMP, these systems will need to be deuterated.

References

- Back, J.W., de Jong, L., Muijsers, A.O. and de Koster, C. G. (2003) Chemical Cross-linking and Mass Spectrometry for Protein Structural Modeling. *J. Mol. Biol.* **331**, 303-313.
- Battiste, J.L. and Wagner, G. (2000) Utilization of site-directed spin labeling and high-resolution heteronuclear nuclear magnetic resonance for global fold determination of large proteins with limited nuclear overhauser effect data. *Biochemistry* **39**(18), 5355-5365.
- Bax A. (2003) Weak alignment offers new NMR opportunities to study protein structure and dynamics. *Protein Sci.* **12**(1), 1-16.
- Bée, M. (1988) *Quasielastic neutron scattering.* Adam Hilger, Bristol.
- Becker, T. and Smith, J.C. (2003) Energy resolution and dynamical heterogeneity effects on elastic incoherent neutron scattering from molecular systems. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* **67**(2 Pt 1), 021904.
- Becker, T., Hayward, J.A., Finney, J.L., Daniel, R.M. and Smith, J.C. (2004) Neutron frequency windows and the protein dynamical transition. *Biophys. J.* **87**(3), 1436-1444.
- Bellissent-Funel, M.-C., Kahn, R., Dianoux, A.J., Fontana, M.P., Maisano, G., Migliardo, P. and Wanderlingh, F. (1984) Incoherent quasielastic neutron scattering from H_2O and aqueous $ZnCl_2$ solutions. *Mol. Phys.* **52**(6), 1479-1486.
- Bellissent-Funel, M.-C., Teixeira, J., Chen, S.H., Dorner, B., Middendorf, H.D. and Crespi, H.L. (1989) Low-frequency collective modes in dry and hydrated proteins. *Biophys. J.* **56**(4), 713-716.
- Bellissent-Funel, M.-C., Zanotti, J.-M. and Chen, S.H. (1996) Slow dynamics of water molecules on the surface of a globular protein. *Faraday Discuss.* **103**, 281-294.
- Bellissent-Funel, M.C., Filabozzi, A. and Chen, S.H. (1997) Measurement of coherent Debye-Waller factor in in vivo deuterated C-phycocyanin by inelastic neutron scattering. *Biophys. J.* **72**(4), 1792-1799.
- Bernadó, P., Blanchard, L., Timmins, P., Marion, D., Ruigrok, R.W. and Blackledge, M. (2005) A structural model for unfolded proteins from residual dipolar couplings and small-angle x-ray scattering. *Proc. Natl. Acad. Sci. USA.* **102**(47), 17002-17007.
- Bernadó, P., Mylonas, E., Petoukhov, M.V., Blackledge, M. and Svergun, D.I. (2007) Structural characterization of flexible proteins using small-angle X-ray scattering. *J. Am. Chem. Soc.* **129**(17), 5656-5664.
- Bernadó, P. and Blackledge, M. (2009) A self-consistent description of the conformational behavior of chemically denatured proteins from NMR and small angle scattering. *Biophys. J.* **97**(10), 2839-2845.
- Bon, C., Dianoux, A.J., Ferrand, M. and Lehmann, M.S. (2002) A model for water motion in crystals of lysozyme based on an incoherent quasielastic neutron-scattering study. *Biophys. J.* **83**(3), 1578-1588.
- Breyton, C., Gabel, F., Abla, M., Pierre, Y., Lebaupain, F., Durand, G., Popot, J.-L., Ebel, C., and Pucci, B. (2009). Micellar and biochemical properties of (hemi)fluorinated surfactants are controlled by the size of the polar head. *Biophys. J.* **97**(4), 1077-1086.
- Brieskorn, E. and Knörrer, H. *Ebene algebraische Kurven.* Birkhäuser (1981).
- Brünger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, M., Pannu, N.S., Read, R.J., Rice, L.M., Simonson, T. and Warren, G.L. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. D Biol. Crystallogr.* **54**(Pt 5), 905-921.
- Brüschweiler, R., Liao, X. and Wright, P.E. (1995) Long-range motional restrictions in a multidomain zinc-finger protein from anisotropic tumbling. *Science* **268**(5212), 886-889.
- Brüschweiler, R. (2003) New approaches to the dynamic interpretation and prediction of NMR relaxation data from proteins. *Curr. Opin. Struct. Biol.* **13**(2), 175-183.
- Bu, Z., Cook, J. and Callaway, D.J. (2001) Dynamic regimes and correlated structural dynamics in native and denatured alpha-lactalbumin. *J. Mol. Biol.* **312**(4), 865-873.

- Bu, Z., Wang, L. and Kendall, D.A. (2003) Nucleotide binding induces changes in the oligomeric state and conformation of Sec A in a lipid environment: a small-angle neutron-scattering study. *J. Mol. Biol.* **332**(1), 23-30.
- Calmettes, P., Durand, D., Desmadril, M., Minard, P., Receveur, V. and Smith, J.C. (1994) How random is a highly denatured protein? *Biophys. Chem.* **53**(1-2), 105-113.
- 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. and Moore, B.P. (1987) A complete mapping of the proteins in the small ribosomal subunit of Escherichia coli. *Science* **238**(4832), 1403-1406.
- Chacón, P., Morán, F., Díaz, J.F., Pantos, E. and Andreu, J.M. (1998) Low-resolution structures of proteins in solution retrieved from X-ray scattering with a genetic algorithm. *Biophys. J.* **74**(6), 2760-2767.
- Cordone, L., Ferrand, M., Vitrano, E. and Zaccai, G. (1999) Harmonic behavior of trehalose-coated carbon-monoxy-myoglobin at high temperature. *Biophys. J.* **76**(2), 1043-1047.
- Daniel, R.M., Finney, J.L., Réat, V., Dunn, R., Ferrand, M. and Smith, J.C. (1999) Enzyme dynamics and activity: time-scale dependence of dynamical transitions in glutamate dehydrogenase solution. *Biophys. J.* **77**(4), 2184-2190.
- Debye, P. (1915) Zerstreuung von Röntgenstrahlen. *Ann. Phys.* (Leipzig) **28**, 809–823.
- Debye, P. (1947) Molecular-weight Determination by Light Scattering. *J. Phys. Coll. Chem.* **51**(1), 18-32.
- Diehl, M., Doster, W., Petry, W. and Schober, H. (1997) Water-coupled low-frequency modes of myoglobin and lysozyme observed by inelastic neutron scattering. *Biophys. J.* **73**(5), 2726-2732.
- Dominguez, C., Boelens, R. and Bonvin, A.M. (2003) HADDOCK: a protein-protein docking approach based on biochemical or biophysical information. *J. Am. Chem. Soc.* **125**(7), 1731-1737.
- Doniach S. (2001) Changes in Biomolecular Conformation Seen by Small Angle X-ray Scattering. *Chem. Rev.* **101**, 1763-1778.
- Dosset, P., Hus, J.C., Blackledge, M. and Marion, D. (2000) Efficient analysis of macromolecular rotational diffusion from heteronuclear relaxation data. *J. Biomol. NMR* **16**(1), 23-28.
- Dosset, P., Hus, J.C., Marion, D. and Blackledge, M. (2001) A novel interactive tool for rigid-body modeling of multi-domain macromolecules using residual dipolar couplings. *J. Biomol. NMR* **20**(3), 223-231.
- Doster, W., Cusack, S. and Petry, W. (1989) Dynamical transition of myoglobin revealed by inelastic neutron scattering. *Nature* **337**(6209), 754-756.
- Durá, M.A., Rosenbaum, E., Larabi, A., Gabel, F., Vellieux, F.M. and Franzetti, B. (2009) The structural and biochemical characterizations of a novel TET peptidase complex from Pyrococcus horikoshii reveal an integrated peptide degradation system in hyperthermophilic Archaea. *Mol. Microbiol.* **72**(1), 26-40.
- Engelman, D.M. and Moore, P.B. (1972) A new method for the determination of biological quarternary structure by neutron scattering. *Proc. Nat. Acad. Sci. USA* **69**, 1997-1999.
- Engelman, D.M. and Moore, P.B. (1975) Determination of quaternary structure by small angle neutron scattering. *Annu. Rev. Biophys. Bioeng.* **4**, 219-241.
- Feigin, L.A. and Svergun, D.I. (1987) Structure Analysis by Small Angle X-ray and Neutron Scattering. *Plenum Press.*
- Ferrand, M., Dianoux, A.J., Petry, W. and Zaccaï, G. (1993) Thermal motions and function of bacteriorhodopsin in purple membranes: effects of temperature and hydration studied by neutron scattering. *Proc. Natl. Acad. Sci. USA* **90**(20), 9668-9672.
- Fitter, J., Lechner, R.E., Büldt, G. and Dencher, N.A. (1996) Internal molecular motions of bacteriorhodopsin: hydration-induced flexibility studied by quasielastic incoherent neutron scattering using oriented purple membranes. *Proc. Natl. Acad. Sci. USA* **93**(15), 7600-7605.
- Fitter, J. (1999) The temperature dependence of internal molecular motions in hydrated and dry alpha-amylase: the role of hydration water in the dynamical transition of proteins. *Biophys. J.* **76**(2), 1034-1042.
- Fleischer, G. and Fujara, F. (1994) NMR as a generalized incoherent scattering experiment. In: *NMR Basic Principles and Progress* **30**, Springer Berlin, pp. 161-207.
- Ford, R.C., Ruffle, S.V., Ramirez-Cuesta, A.J., Michalarias, I., Beta, I., Miller, A. and Li, J. (2004) Inelastic incoherent neutron scattering measurements of intact cells and tissues and detection of interfacial water. *J. Am. Chem. Soc.* **126**(14), 4682-4688.

- Franzetti, B., Schoehn, G., Hernandez, J.F., Jaquinod, M., Ruigrok, R.W. and Zaccai, G. (2002) Tetrahedral aminopeptidase: a novel large protease complex from archaea. *EMBO J.* **21**(9), 2132-2138.
- Frölich, A., Gabel, F., Jasnin, M., Lehnert, U., Oesterhelt, D., Stadler, A.M., Tehei, M., Weik, M., Wood, K. and Zaccai, G. (2009) From shell to cell: neutron scattering studies of biological water dynamics and coupling to activity. *Faraday Discuss.* **141**, 117-130.
- Gabel, F. (2000) Etude des Membranes Biologiques par Diffusion de Neutrons. Rapport Stage DEA. Université Joseph Fourier, Grenoble, France.
- Gabel, F., Bicout, D., Lehnert, U., Tehei, M., Weik, M. and Zaccai, G. (2002) Protein dynamics studied by neutron scattering. *Q. Rev. Biophys.* **35**(4), 327-367.
- Gabel, F., Weik, M., Doctor, B.P., Saxena, A., Fournier, D., Brochier, L., Renault, F., Masson, P., Silman, I. and Zaccai, G. (2004) The influence of solvent composition on global dynamics of human butyrylcholinesterase powders: a neutron-scattering study. *Biophys. J.* **86**(5), 3152-2165.
- Gabel, F. (2005) Protein dynamics in solution and powder measured by incoherent elastic neutron scattering: the influence of Q-range and energy resolution. *Eur. Biophys. J.* **34**(1), 1-12.
- Gabel, F., Simon, B. and Sattler, M. (2006) A target function for quaternary structural refinement from small angle scattering and NMR orientational restraints. *Eur. Biophys. J.* **35**(4), 313-327.
- Gabel, F. and Bellissent-Funel, M.-C. (2007) C-phycocyanin hydration water dynamics in the presence of trehalose: an incoherent elastic neutron scattering study at different energy resolutions. *Biophys. J.* **92**(11), 4054-4063.
- Gabel, F., Simon, B., Nilges, M., Petoukhov, M., Svergun, D. and Sattler, M. (2008) A structure refinement protocol combining NMR residual dipolar couplings and small angle scattering restraints. *J. Biomol. NMR* **41**(4), 199-208.
- Gabel, F., Jensen, M.R., Zaccaï, G. and Blackledge, M. (2009) Quantitative modelfree analysis of urea binding to unfolded ubiquitin using a combination of small angle X-ray and neutron scattering. *J. Am. Chem. Soc.* **131**(25), 8769-8771.
- Glatter, O. (1977) A new method for the evaluation of small-angle scattering data. *J. Appl. Cryst.* **10**, 415-421.
- Glatter, O. (1979). The Interpretation of Real-Space Information from Small-Angle Scattering Experiments. *J. Appl. Cryst.* **12**, 166-172.
- Glatter, O. (1980a) Computation of Distance Distribution Functions and Scattering Functions of Models for Small-Angle Scattering Experiments. *Acta Phys. Austriaca* **52**, 243-256.
- Glatter, O. (1980b) Determination of Particle-Size Distribution Functions from Small-Angle Scattering Data by Means of the Indirect Transformation Method. *J. Appl. Cryst.* **13**, 7-11.
- Glatter, O. and Kratky, O. (1982) Small Angle X-ray Scattering. *Academic Press.*
- Glatter, O. (2002) The Inverse Scattering Problem in Small-Angle Scattering. In: *Neutrons, X-rays and Light.* Eds. Lindner, P. and Zemb, Th. Delta Series, North Holland.
- Goldstein, H. (1977) *Classical Mechanics*. Addison-Wesley, Reading.
- Grishaev, A., Wu, J., Trewhella, J. and Bax A. (2005) Refinement of multidomain protein structures by combination of solution small-angle X-ray scattering and NMR data. *J. Am. Chem. Soc.* **127**(47), 16621-16628.
- Guinier, A. (1939) La Diffraction des Rayons X aux Très Faibles Angles: Applications à l'Etude des Phénomènes Ultra-microscopiques. *Ann. Phys. (Paris)* **12**, 161-236.
- Guinier, A. and Fournet, G. (1955) Small Angle Scattering of X-rays. *John Wiley & Sons.*
- Guo, X.H., Zhao, N.M., Chen, S.H. and Teixeira, J. (1990) Small-angle neutron scattering study of the structure of protein/detergent complexes. *Biopolymers* **29**(2), 335-346.
- Harrison, S.C. (1969) Structure of tomato bushy stunt virus. I. Spherically averaged electron density. *J. Mol. Biol.* **42**, 457-483.
- Hayward, J.A. and Smith, J.C. (2002) Temperature dependence of protein dynamics: computer simulation analysis of neutron scattering properties. *Biophys. J.* **82**(3), 1216-1225.
- Hoppe, W. (1973) The label triangulation method and the mixed isomorphous replacement principle. *J. Mol. Biol.* **78**(3), 581-585.
- Hoppe, W., May, R., Stöckel, P., Lorenz, S., Erdmann, V.A., Wittmann, H.G., Crespi, H.L., Katz, J.J. and Ibel, K. (1975) Neutron Scattering Measurements with the Label Triangulation Method on the 50S Subunit of *E. coli* Risbosomes. *Neutron Scattering for the Analysis of Biological Structures*, Brookhaven Symposia in Biology No. 27, pp. IV38-IV48.

- Hunt, J.F., McCrea, P.D., Zaccaï, G. and Engelman, D.M. (1997) Assessment of the aggregation state of integral membrane proteins in reconstituted phospholipid vesicles using small angle neutron scattering. *J. Mol. Biol.* **273**(5), 1004-1019.
- Jacrot, B. (1976) The study of biological structures by neutron scattering in solution. *Rep. Prog. Phys.* **39**(10), 911-953.
- Jacrot, B. and Zaccai, G. (1981) Determination of Molecular Weight by Neutron Scattering. *Biopolymers* **20**, 2413- 2426.
- Jacrot, B., Cusack, S., Dianoux, A.J. and Engelman, D.M. (1982) Inelastic neutron scattering analysis of hexokinase dynamics and its modification on binding of glucose. *Nature* **300**(5887), 84-86.
- Jasnin, M., Moulin, M., Haertlein, M., Zaccai, G. and Tehei, M. (2008a) In vivo measurement of internal and global macromolecular motions in Escherichia coli. *Biophys. J.* **95**(2), 857-864.
- Jasnin, M., Moulin, M., Haertlein, M., Zaccai, G. and Tehei, M. (2008b) Down to atomic-scale intracellular water dynamics. *EMBO Rep.* **9**(6), 543-547.
- Jensen, M.R., Markwick, P.R., Meier, S., Griesinger, C., Zweckstetter, M., Grzesiek, S., Bernadó, P. and Blackledge, M. (2009) Quantitative determination of the conformational properties of partially folded and intrinsically disordered proteins using NMR dipolar couplings. *Structure* **17**(9), 1169-1185.
- Johs, A., Hammel, M., Waldner, I., May, R.P., Laggner, P. and Prassl, R. (2006) Modular structure of solubilized human apolipoprotein B-100. Low resolution model revealed by small angle neutron scattering. *J. Biol. Chem.* **281**(28), 19732-19739.
- Kirste, R.G., Schulz, G.V. and Stuhrmann, H.B. (1969) Die Konformationsänderung des Pottwal-Mesmyoglobins bei der reversiblen Denaturierung im pH-Bereich 7 bis 1. *Z. Naturf.* **24**b, 1385-1392.
- Kneller, G.R. and Calandrini, V. (2007) Estimating the influence of finite instrumental resolution on elastic neutron scattering intensities from proteins. *J. Chem. Phys.* **126**(12), 125107.
- Koch, M.H.J., Vachette, P. and Svergun, D.I. (2003) Small-angle scattering: a view on the properties, structures and structural changes of biological macromolecules in solution. *Q. Rev. Biophys.* **36**(2), 147-227.
- Kohn, J.E., Millett, I.S., Jacob, J., Zagrovic, B., Dillon, T.M., Cingel, N., Dothager, R.S., Seifert, S., Thiyagarajan, P., Sosnick, T.R., Hasan, M.Z., Pande, V.S., Ruczinski, I., Doniach, S. and Plaxco, K.W. (2004) Random-coil behavior and the dimensions of chemically unfolded proteins. *Proc. Natl. Acad. Sci. USA* **101**(34), 12491-12496.
- Konarev, P.V., Petoukhov, M.V. and Svergun, D.I. (2001) MASSHA - a graphics system for rigid-body modelling of macromolecular complexes against solution scattering data. *J. Appl. Cryst.* **34**, 527-532.
- Korepanova, A., Gao, F.P., Hua, Y., Qin, H., Nakamoto, R.K. and Cross, T.A. (2005) Cloning and expression of multiple integral membrane proteins from Mycobacterium tuberculosis in Escherichia coli. *Protein Sci.* **14**(1), 148-158.
- Köster, L., Rauch, H. and Seymann, L. (1991) *Atomic Data and Nucl. Data Tab.* **49**, 65-120.
- Kratky, O. (1963) X-ray small angle scattering with substances of biological interest in dilute solutions. *Prog. Biophys. Mol. Biol.* **13**, 105-173.
- Kratky, O. and Pilz, I. (1972) Recent advances and applications of diffuse X-ray small-angle scattering on biopolymers in dilute solutions. *Q. Rev. Biophys.* **5**, 481-537.
- Lacapère, J.J., Pebay-Peyroula, E., Neumann, J.M. and Etchebest, C. (2007) Determining membrane protein structures: still a challenge! *Trends Biochem. Sci.* **32**(6), 259-270.
- Lechner, R.E. (2001) Observation-time dependent structural information from quasielastic neutron scattering. *Physica B* **301**, 83-93.
- Lehnert, U., Réat, V., Weik, M., Zaccaï, G. and Pfister, C. (1998) Thermal motions in bacteriorhodopsin at different hydration levels studied by neutron scattering: correlation with kinetics and light-induced conformational changes. *Biophys. J.* **75**(4), 1945-1952.
- le Maire, M., Champeil, P. and Moller, J.V. (2000) Interaction of membrane proteins and lipids with solubilizing detergents. *Biochim. Biophys. Acta.* **1508**(1-2), 86-111.
- Liker, E., Fernandez, E., Izaurralde, E. and Conti, E. (2000) The structure of the mRNA export factor TAP reveals a cis arrangement of a non-canonical RNP domain and an LRR domain. *EMBO J.* **19**(21), 5587-5598.
- Lipari, G. and Szabo, A. (1982) Model-Free Approach to the Interpretation of Nuclear Magnetic Resonance Relaxation in Macromolecules 1. Theory and Range of Validity. *J. Am. Chem. Soc.* **104**(17), 4546-4559.

- Lipfert, J. and Doniach, S. (2007) Small-Angle X-ray Scattering from RNA, Proteins, and Protein Complexes. *Annu. Rev. Biophys. Biomol. Struct.* **36**, 307-327.
- Magazù, S., Maisano, G., Migliardo, F. and Mondelli, C. (2004) Mean-square displacement relationship in bioprotectant systems by elastic neutron scattering. *Biophys. J.* **86**(5), 3241-3249.
- Mareuil, F., Sizun, C., Pérez, J., Schoenauer, M., Lallemand, J.Y. and Bontems, F. (2007) A simple genetic algorithm for the optimization of multidomain protein homology models driven by NMR residual dipolar coupling and small angle X-ray scattering data. *Eur. Biophys. J.* **37**(1), 95-104.
- Marino, M., Zou, P., Svergun, D., Garcia, P., Edlich, C., Simon, B., Wilmanns, M., Muhle-Goll, C. and Mayans, O. (2006) The Ig doublet Z1Z2: a model system for the hybrid analysis of conformational dynamics in Ig tandems from titin. *Structure* **14**(9), 1437-1447.
- Masson, S., Kern, T., Le Gouëllec, A., Giustini, C., Simorre, J.P., Callow, P., Vernet, T., Gabel, F. and Zapun, A. (2009) Central domain of DivIB caps the C-terminal regions of the FtsL/DivIC coiled-coil rod. *J. Biol. Chem.* **284**(40), 27687-27700.
- Mattinen, M.L., Pääkkönen, K., Ikonen, T., Craven, J., Drakenberg, T., Serimaa, R., Waltho, J. and Annila, A. (2002) Quaternary structure built from subunits combining NMR and small-angle x-ray scattering data. *Biophys. J.* **83**(2), 1177-1183.
- May, R.P. (1991) Label Triangulation. In *Neutron, X-ray and Light Scattering.* Eds. Lindner, P. and Zemb, Th. Elsevier.
- Middendorf, H.D., Randall, J.T. and Crespi, H.L. (1984) Neutron Spectroscopy of hydrogenous and biosynthetically deuterated proteins. In: *Neutron in Biology*. Ed: Schoenborn, B.P. Plenum Press.
- Millett, I.S., Doniach, S. and Plaxco, K.W. (2002) Toward a taxonomy of the denatured state: small angle scattering studies of unfolded proteins. *Adv. Protein Chem.* **62**, 241-262.
- Mittag, T. and Forman-Kay, J.D. (2007) Atomic-level characterization of disordered protein ensembles. *Curr. Opin. Struct. Biol.* **17**(1), 3-14.
- Nierhaus, K.H., Lietzke, R., May, R.P., Nowotny, V., Schulze, H., Simpson, K., Wurmbach, P. and Stuhrmann, H.B. (1983) Shape determinations of ribosomal proteins *in situ*. *Proc. Natl. Acad. Sci. USA* **80**(10), 2889-2893.
- Niimura, N. (1999) Neutrons expand the field of structural biology. *Curr. Opin. Struct. Biol.* **9**, 602-608.
- Paciaroni, A., Stroppolo, M.E., Arcangeli, C., Bizzarri, A.R., Desideri, A. and Cannistraro, S. (1999) Incoherent neutron scattering of copper azurin: a comparison with molecular dynamics simulation results. *Eur. Biophys. J.* **28**(6), 447-456.
- Paciaroni, A., Cinelli, S. and Onori, G. (2002) Effect of the environment on the protein dynamical transition: a neutron scattering study. *Biophys. J.* **83**(2), 1157-1164.
- Palmer, A.G. 3rd. (2001) NMR probes of molecular dynamics: overview and comparison with other techniques. *Annu. Rev. Biophys. Biomol. Struct.* **30**, 129-155.
- Pebay-Peyroula, E., Garavito, R.M., Rosenbusch, J.P., Zulauf, M. and Timmins, P.A. (1995) Detergent structure in tetragonal crystals of OmpF porin. *Structure* **3**(10), 1051-1059.
- Pérez, J., Zanotti, J.M. and Durand, D. (1999) Evolution of the internal dynamics of two globular proteins from dry powder to solution. *Biophys. J.* **77**(1), 454-469.
- Pérez, J., Vachette, P., Russo, D., Desmadril, M. and Durand, D. (2001) Heat-induced unfolding of neocarzinostatin, a small all-beta protein investigated by small-angle X-ray scattering. *J. Mol. Biol.* **308**(4), 721-743.
- Petoukhov, M.V., Eady, N.A., Brown, K.A. and Svergun, D.I. (2002) Addition of missing loops and domains to protein models by x-ray solution scattering. *Biophys. J.* **83**(6), 3113-3125.
- Petoukhov, M.V. and Svergun, D.I. (2005) Global rigid body modeling of macromolecular complexes against small-angle scattering data. *Biophys. J.* **89**(2), 1237-1250.
- Petoukhov, M.V. and Svergun, D.I. (2007) Analysis of X-ray and neutron scattering from biomacromolecular solutions. *Curr. Opin. Struct. Biol.* **17**(5), 562-571.
- Petrescu, A.J., Receveur, V., Calmettes, P., Durand, D., Desmadril, M., Roux, B. and Smith, J.C. (1997) Small-angle neutron scattering by a strongly denatured protein: analysis using random polymer theory. *Biophys. J.* **72**(1), 335-342.
- Petrescu, A.J., Receveur, V., Calmettes, P., Durand, D. and Smith, J.C. (1998) Excluded volume in the configurational distribution of a strongly-denatured protein. *Protein Sci.* **7**(6), 1396-1403.
- Plaxco, K.W. and Gross, M. (2001) Unfolded, yes, but random? Never! *Nat. Struct. Biol.* **8**(8), 659-660.

- Prestegard, J.H. (1998) New techniques in structural NMR--anisotropic interactions. *Nat. Struct. Biol.* **5** Suppl., 517-522.
- Putnam, C.D., Hammel, M., Hura, G.L. and Tainer, J.A. (2007) X-ray solution scattering (SAXS) combined with crystallography and computation: defining accurate macromolecular structures, conformations and assemblies in solution. *Q. Rev. Biophys.* **40**(3), 191-285.
- Rahman, A., Singwi, K.S. and Sjölander, A. (1962) Theory of Slow Neutron Scattering by Liquids. *Phys. Rev.* **126**(3), 986-996.
- Randall, J., Middendorf, H.D., Crespi, H.L. and Taylor, A.D. (1978) Dynamics of protein hydration by quasi-elastic neutron scattering. *Nature* **276**, 636-638.
- Rawiso, M., Duplessix, R. and Picot, C. (1987) Scattering Function of Polysterene. *Macromolecules* **20**, 630-648.
- Réat, V., Zaccai, G., Ferrand, M. and Pfister, C. (1997) Functional Dynamics in Purple Membrane. In: *Biological Macromolecular Dynamics. Proceedings of a Workshop on Inelastic and Quasielastic Neutron Scattering in Biology.* Eds. Cusack, S., Büttner, H., Ferrand, M., Langan, P. and Timmins, P. Adenine Press.
- Réat, V., Patzelt, H., Ferrand, M., Pfister, C., Oesterhelt, D. and Zaccai, G. (1998) Dynamics of different functional parts of bacteriorhodopsin: H-2H labeling and neutron scattering. *Proc. Natl. Acad. Sci. USA* **95**(9), 4970-4975.
- Réat, V., Dunn, R., Ferrand, M., Finney, J.L., Daniel, R.M. and Smith, J.C. (2000) Solvent dependence of dynamic transitions in protein solutions. *Proc. Natl. Acad. Sci. USA* **97**(18), 9961-9966.
- Receveur, V., Calmettes, P., Smith, J.C., Desmadril, M., Coddens, G. and Durand, D. (1997) Picosecond dynamical changes on denaturation of yeast phosphoglycerate kinase revealed by quasielastic neutron scattering. *Proteins* **28**(3), 380-387.
- Russo, D., Hura, G. and Head-Gordon, T. (2004) Hydration dynamics near a model protein surface. *Biophys. J.* **86**(3), 1852-1862.
- Salmon, P.S. (1987) The dynamics of water molecules in ionic solution: I. The application of quasi-elastic neutron scattering to the study of translational diffusive proton motion. *J. Phys. C: Solid State Phys.* **20**, 1573-1587.
- Schellman, J. (1958) The Factors Affecting the Stability of Hydrogen-bonded Polypeptide Structures in Solution. *J. Phys. Chem*. **62**, 1485–1494.
- Schurtenberger, P. (2002) Static Properties of Polymers. In: *Neutrons, X-rays and Light.* Eds. Lindner, P. and Zemb, Th. Delta Series, North Holland.
- Sears, V.F. (1975) Slow Neutron Multiple Scattering. *Adv. Phys.* **24**(1), 1-45.
- Serdyuk, I.N., Zaccai, N.R. and Zaccai, G. (2007) *Methods in Molecular Biophysics.* Cambridge University Press.
- Settles, M. and Doster, W. (1996) Anomalous Diffusion of Adsorbed Water. A Neutron Scattering Study of Hydrated Myoglobin. *Faraday Discuss.* **103**, 269-280.
- Settles, M. and Doster, W. (1997) Iterative Calculation of the Vibrational Density of States from Incoherent Neutron Scattering Data with the Account of Double Scattering. In: *Biological Macromolecular Dynamics. Proceedings of a Workshop on Inelastic and Quasielastic Neutron Scattering in Biology.* Eds. Cusack, S., Büttner, H., Ferrand, M., Langan, P. and Timmins, P. Adenine Press.
- Shortle, D. and Ackerman, M.S. (2001) Persistence of native-like topology in a denatured protein in 8 M urea. *Science* **293**(5529), 487-489.
- Smith, J.C. (1991) Protein dynamics: comparison of simulations with inelastic neutron scattering experiments. *Q. Rev. Biophys.* **24**(3), 227-291.
- Springer, T. (1977) Molecular Rotations and Diffusion in solids, in Particular Hydrogen in Metals. In: *Dynamics of Solids and Liquids by Neutron Scattering.* Eds.: S.W. Lovesey, T. Springer. Springer Verlag Berlin, Heidelberg, New York. p. 284f.
- Stadler, A.M., Embs, J.P., Digel, I., Artmann, G.M., Unruh, T., Büldt, G. and Zaccai, G. (2008) Cytoplasmic water and hydration layer dynamics in human red blood cells. *J. Am. Chem. Soc.* **130**(50), 16852-16853.
- Stadler, A.M., Digel, I., Embs, J.P., Unruh, T., Tehei, M., Zaccai, G., Büldt, G. and Artmann, G.M. (2009) From powder to solution: hydration dependence of human hemoglobin dynamics correlated to body temperature. *Biophys. J.* **96**(12), 5073-5081.

- Stöckel, P., May, R., Strell, I., Cejka, Z., Hoppe, W., Heumann, H., Zillig, W., Crespi, H.L., Katz, J.E. and Ibel, K. (1979) Determination of Intersubunit Distances and Subunit Shape Parameters in DNA-Dependent RNA Polymerase by Neutron Small-Angle Scattering. *J. Appl. Cryst.* **12**, 176-185.
- Stuhrmann, H.B. (1970a) Interpretation of small-angle scattering functions of dilute solutions and gases. A representation of the structures related to a one-particle scattering function. *Acta Cryst.* **A26**, 297-306.
- Stuhrmann, H.B. (1970b) Ein neues Verfahren zur Bestimmung der Oberflächenform und der inneren Struktur von gelösten globulären Proteinen aus Röntgenkleinwinkelmessungen. *Z. Phys. Chem.* (Frankfurt), **72**(4-6), 177-184.
- Stuhrmann, H.B. (1974) Neutron Small-Angle Scattering of Biological Macromolecules in Solution. *J. Appl. Cryst.* **7**, 173-178.
- Svergun, D.I., Semenyuk, A.V. and Feigin, L.A. (1988) Small-Angle Scattering Data Treatment by the Regularization Method. *Acta Cryst.* **A44**, 244-250.
- Svergun, D.I. (1992) Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J. Appl. Cryst.* **25**, 495-503.
- Svergun D.I., Barberato, C. and Koch, M.H.J. (1995) CRYSOL : a Program to Evaluate X-ray Solution Scattering of Biological Macromolecules from Atomic Coordinates. *J. Appl. Cryst.* **28**, 768-773.
- Svergun, D.I., Richard, S., Koch, M.H., Sayers, Z., Kuprin, S. and Zaccai, G. (1998) Protein hydration in solution: experimental observation by x-ray and neutron scattering. *Proc. Natl. Acad. Sci. USA* **95**(5), 2267-2272.
- Svergun, D.I. (1999) Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. *Biophys J.* **76**(6), 2879-2886.
- Svergun, D.I. and Koch M.H.J. (2002) Advances in structure analysis using small-angle scattering in solution. *Curr. Opin. Struct. Biol.* **12**(5), 654-660.
- Tanford, C. (1964) Isothermal Unfolding of Globular Proteins in Aqueous Urea Solutions. *J. Am. Chem. Soc*. **86**, 2050–2059.
- Tardieu, A., Vachette, P., Gulik, A. and le Maire, M. (1981) Biological macromolecules in solvents of variable density: characterization by sedimentation equilibrium, densimetry, and X-ray forward scattering and an application to the 50S ribosomal subunit from Escherichia coli. *Biochemistry* **20**(15), 4399-4406.
- Tehei, M., Madern, D., Pfister, C. and Zaccai, G. (2001) Fast dynamics of halophilic malate dehydrogenase and BSA measured by neutron scattering under various solvent conditions influencing protein stability. *Proc. Natl. Acad. Sci. USA* **98**(25), 14356-14361.
- Tehei, M., Franzetti, B., Madern, D., Ginzburg, M., Ginzburg, B.Z., Giudici-Orticoni, M.T., Bruschi, M. and Zaccai G. (2004) Adaptation to extreme environments: macromolecular dynamics in bacteria compared in vivo by neutron scattering. *EMBO Rep.* **5**(1), 66-70.
- Tehei, M., Franzetti, B., Wood, K., Gabel, F., Fabiani, E., Jasnin, M., Zamponi, M., Oesterhelt, D., Zaccai, G., Ginzburg, M. and Ginzburg, B.Z. (2007) Neutron scattering reveals extremely slow cell water in a Dead Sea organism. *Proc. Natl. Acad. Sci. USA* **104**(3), 766-771.
- Teixeira, J., Bellissent-Funel, M., Chen, S.H. and Dianoux, A.J. (1985) Experimental determination of the nature of diffusive motions of water molecules at low temperatures. *Phys. Rev. A.* **31**(3), 1913-1917.
- Tjandra, N. and Bax, A. (1997) Direct Measurement of Distances and Angles in Biomolecules by NMR in Dilute Liquid Crystalline Medium. *Science* **278**, 1111-1114.
- Tjandra, N., Garrett, D.S., Gronenborn, A.M., Bax, A. and Clore, G.M. (1997) Defining long range order in NMR structure determination from the dependence of heteronuclear relaxation times on rotational diffusion anisotropy. *Nat. Struct. Biol.* **4**(6), 443-449.
- Trantham, E.C., Rorschach, H.E., Clegg, J.S., Hazlewood, C.F., Nicklow, R.M. and Wakabayashi, N. (1984) Diffusive properties of water in *Artemia* cysts as determined from quasi-elastic neutron scattering spectra. *Biophys. J.* **45**(5), 927-938.
- Tsai, A.M., Neumann, D.A. and Bell, L.N. (2000) Molecular dynamics of solid-state lysozyme as affected by glycerol and water: a neutron scattering study. *Biophys. J.* **79**(5), 2728-2732.
- Tugarinov, V. and Kay, L.E. (2003) Quantitative NMR studies of high molecular weight proteins: application to domain orientation and ligand binding in the 723 residue enzyme malate synthase G. *J. Mol. Biol.* **327**(5), 1121-1133.

- van Hove, L. (1954) Correlations in Space and Time and Born Approximation Scattering in Systems of Interacting Particles. *Phy. Rev.* **95**(1), 249-262.
- Vineyard, G.H. (1958) Scattering of slow Neutrons by a Liquid. *Phys. Rev.* **110**(5), 999-1010.
- Volino, F. and Dianoux, A.J. (1980) Neutron Incoherent Scattering Law for Diffusion in a Potential of Spherical Symmetry - General Formalism and Application to Diffusion inside a Sphere. *Mol. Phys.* **41**(2), 271-279.
- Wallin, E. and von Heijne, G. (1998) Genome-wide analysis of integral membrane proteins from eubacterial, archaean, and eukaryotic organisms. *Protein Sci.* **7**(4), 1029-1038.
- Wang, Z.Y., Muraoka, Y., Nagao, M., Shibayama, M., Kobayashi, M. and Nozawa, T. (2003) Determination of the B820 subunit size of a bacterial core light-harvesting complex by small-angle neutron scattering. *Biochemistry* **42**(39), 11555-11560.
- White, S.H. (2009) Biophysical dissection of membrane proteins. *Nature* **459**, 344-346.
- Wood, K., Frölich, A., Paciaroni, A., Moulin, M., Härtlein, M., Zaccai, G., Tobias, D.J. and Weik, M. (2008a) Coincidence of dynamical transitions in a soluble protein and its hydration water: direct measurements by neutron scattering and MD simulations. *J. Am. Chem. Soc.* **130**(14), 4586-4587.
- Wood, K., Grudinin, S., Kessler, B., Weik, M., Johnson, M., Kneller, G.R., Oesterhelt, D. and Zaccai, G. (2008b) Dynamical heterogeneity of specific amino acids in bacteriorhodopsin. *J. Mol. Biol.* **380**(3), 581-591.
- Yuzawa, S., Ogura, K., Horiuchi, M., Suzuki, N.N., Fujioka, Y., Kataoka, M., Sumimoto, H. and Inagaki, F. (2004) Solution structure of the tandem Src homology 3 domains of p47phox in an autoinhibited form. *J. Biol. Chem.* **279**(28), 29752-29760.
- Zaccai, G. and Jacrot, B. (1983) Small angle neutron scattering. *Annu. Rev. Biophys. Bioeng.* **12**, 139-157.
- Zaccai, G. (2000) How soft is a protein? A protein dynamics force constant measured by neutron scattering. *Science* **288**(5471), 1604-1607.
- Zanotti, J.M., Bellissent-Funel, M.C. and Parello, J. (1999) Hydration-coupled dynamics in proteins studied by neutron scattering and NMR: the case of the typical EF-hand calcium-binding parvalbumin. *Biophys. J.* **76**(5), 2390-2411.
- Zimmer, J., Doyle, D.A. and Grossmann, J.G. (2006) Structural characterization and pH-induced conformational transition of full-length KcsA. *Biophys. J.* **90**(5), 1752-1766.
- Zorn, R. (2009) On the evaluation of neutron scattering elastic data. *Nucl. Instr. Meth. Phys. Res. A* **603**, 439-445.
- Zuiderweg, E.R. (2002) Mapping protein-protein interactions in solution by NMR spectroscopy. *Biochemistry* **41**(1), 1-7.

Appendix

Supplementary equations for chapter 1

$$B(\theta,\phi) = 4(R_1+R_2)^2\left[\begin{array}{l}\sin^2\theta\left(a_B\cos^2\phi + b_B\sin^2\phi + c_B\sin\phi\cos\phi\right) + d_B\cos^2\theta \\ +\sin\theta\cos\theta(e_B\cos\phi + f_B\sin\phi)\end{array}\right]$$
$$+ 4(R_1+R_2)[\sin\theta(g_B\cos\phi + h_B\sin\phi) + i_B\cos\theta],$$
$$a_B = \sum_{i\in K_1, j\in K_2}\left(x_i^2 + x_j^2\right),\quad b_B = \sum_{i\in K_1, j\in K_2}\left(y_i^2 + y_j^2\right),\quad c_B = 2\sum_{i\in K_1, j\in K_2}\left(x_iy_i + x_jy_j\right),$$
$$d_B = \sum_{i\in K_1, j\in K_2}\left(z_i^2 + z_j^2\right),\quad e_B = 2\sum_{i\in K_1, j\in K_2}\left(x_iz_i + x_jz_j\right),\quad f_B = 2\sum_{i\in K_1, j\in K_2}\left(y_iz_i + y_jz_j\right),$$
$$g_B = \sum_{i\in K_1, j\in K_2}\left(x_i^3 + y_i^2x_i + z_i^2x_i - x_j^3 - y_j^2x_j - z_j^2x_j\right),$$
$$h_B = \sum_{i\in K_1, j\in K_2}\left(x_i^2y_i + y_i^3 + z_i^2y_i - x_j^2y_j - y_j^3 - z_j^2y_j\right),$$
$$i_B = \sum_{i\in K_1, j\in K_2}\left(x_i^2z_i + y_i^2z_i + z_i^3 - x_j^2z_j - y_j^2z_j - z_j^3\right),$$

$$C(\theta,\phi) = 24(R_1+R_2)^4\left[\begin{array}{l}\sin^2\theta\left(a_C\cos^2\phi + b_C\sin^2\phi + c_C\sin\phi\cos\phi\right) + d_C\cos^2\theta \\ +\sin\theta\cos\theta(e_C\cos\phi + f_C\sin\phi)\end{array}\right]$$
$$+ 24(R_1+R_2)^3[\sin\theta(g_C\cos\phi + h_C\sin\phi) + i_C\cos\theta]$$
$$+ 16(R_1+R_2)^3\left[\begin{array}{l}\sin^3\theta\left(j_C\cos^3\phi + k_C\sin^3\phi + l_C\sin\phi\cos^2\phi + m_C\sin^2\phi\cos\phi\right) \\ +n_C\cos^3\theta + \sin^2\theta\cos\theta\left(o_C\cos^2\phi + p_C\sin^2\phi + q_C\cos\phi\sin\phi\right) \\ +\sin\theta\cos^2\theta(r_C\cos\phi + s_C\sin\phi)\end{array}\right],$$
$$a_C = a_B + \sum_{i\in K_1, j\in K_2}\left(x_i^2 + x_j^2\right)\frac{r_{ij}^2}{(R_1+R_2)^2},\quad b_C = b_B + \sum_{i\in K_1, j\in K_2}\left(y_i^2 + y_j^2\right)\frac{r_{ij}^2}{(R_1+R_2)^2},$$
$$c_C = c_B + 2\sum_{i\in K_1, j\in K_2}\left(x_iy_i + x_jy_j\right)\frac{r_{ij}^2}{(R_1+R_2)^2},\quad d_C = d_B + \sum_{i\in K_1, j\in K_2}\left(z_i^2 + z_j^2\right)\frac{r_{ij}^2}{(R_1+R_2)^2},$$
$$e_C = e_B + \sum_{i\in K_1, j\in K_2}\left(x_iz_i + x_jz_j\right)\frac{r_{ij}^2}{(R_1+R_2)^2},\quad f_C = f_B + \sum_{i\in K_1, j\in K_2}\left(y_iz_i + y_jz_j\right)\frac{r_{ij}^2}{(R_1+R_2)^2},$$
$$g_Cg_B + \sum_{i\in K_1, j\in K_2}\left(x_i^3 + y_i^2x_i + z_i^2x_i - x_j^3 - y_j^2x_j - z_j^2x_j\right)\frac{1}{2}\frac{r_{ij}^2}{(R_1+R_2)^2},$$
$$h_C = h_B + \sum_{i\in K_1, j\in K_2}\left(x_i^2y_i + y_i^3 + z_i^2y_i - x_j^2y_j - y_j^3z_j^2y_j\right)\frac{1}{2}\frac{r_{ij}^2}{(R_1+R_2)^2},$$
$$i_C = i_B + \sum_{i\in K_1, j\in K_2}\left(x_i^2z_i + y_i^2z_i + z_i^3 - x_j^2z_j - y_j^2z_j - z_j^3\right)\frac{1}{2}\frac{r_{ij}^2}{(R_1+R_2)^2},$$
$$j_C = \sum_{i\in K_1, j\in K_2}\left(x_i^3 - x_j^3\right),\quad k_C = \sum_{i\in K_1, j\in K_2}\left(y_i^3 - y_j^3\right),\quad l_C = 3\sum_{i\in K_1, j\in K_2}\left(x_i^2y_i - x_j^2y_j\right),$$
$$m_C = 3\sum_{i\in K_1, j\in K_2}\left(y_i^2x_i - y_j^2x_j\right),\quad n_C = \sum_{i\in K_1, j\in K_2}\left(z_i^3 - z_j^3\right),\quad o_C = 3\sum_{i\in K_1, j\in K_2}\left(x_i^2z_i - x_j^2z_j\right),$$
$$p_C = 3\sum_{i\in K_1, j\in K_2}\left(y_i^2z_i - y_j^2z_j\right),\quad q_C = 6\sum_{i\in K_1, j\in K_2}\left(x_iy_iz_i - x_jy_jz_j\right),\quad r_C = 3\sum_{i\in K_1, j\in K_2}\left(z_i^2x_i - z_j^2x_j\right),$$
$$s_C = 3\sum_{i\in K_1, j\in K_2}\left(z_i^2y_i - z_j^2y_j\right),$$

Curriculum vitae

Personal data

Name	Frank Gabel
Date of birth	July 2nd 1974
Place of birth	Pforzheim (Germany)
Nationality	German
Marital status	married, two children
Address	117, rue de la Procession, 38500 La Buisse, France

Professional data

Employer	CEA Grenoble
Status	Ingénieur-Chercheur (E2) since October 1st 2006
Address	Institut de Biologie Structurale, 41 rue Jules Horowitz, 38027 Grenoble, France
Fax	+33 (0)4 38 78 54 94
Telephone	+33 (0)4 38 78 95 73
Email	frank.gabel@ibs.fr

Secondary education

1985-1994 Naturwissenschaftliches Kepler Gymnasium (Pforzheim, Germany)
1994 A levels ("Abitur") with the mark "very good" (1.1)
1989-1994 Multiple participations at national selection rounds of the "*International Physics Olympiad*" 1994-1995 Civil service at the "Arbeiter-Samariter-Bund" (Pforzheim) caring for disabled people

University education

1995-1999	Undergraduate and graduate studies of physics at Karlsruhe Technical University
1998	Bachelor of Science in Physics ("Vordiplom") with the mark "very good" (1.1) 3 month internship at BOSCH Research Centre (Schillerhöhe, Stuttgart)
1999	Master of Science in Physics ("Hauptdiplom") with the mark "very good" (1.1)
1999-2000	French Diplôme d'Etudes Approfondies (DEA) „*Physique de la Matière et du Rayonnement*" at Joseph Fourier University (Grenoble) "Stage DEA" at the Institut de Biologie Structurale (IBS) under the supervision of Dr. Giuseppe Zaccaï "*Etude des membranes biologiques par diffusion de neutrons*"
2001	German graduate diploma in physics with mention "very good" (1.1) "*Messung thermischer Fluktuationen in photosynthetischen Membranen mittels Neutronenstreuung*". Supervisor Dr. Giuseppe Zaccaï
2000-2003	PhD thesis in biophysics at the IBS under the supervision of Dr. Giuseppe Zaccaï "*The effect of solvent composition, inhibition, and structure on cholinesterase molecular dynamics: A neutron scattering study*"

Academic career

2004-2006	Postdoc position at the EMBL Heidelberg (NMR group Dr. Michael Sattler). "*Combining NMR and Small Angle Scattering Restraints for Structural Refinement in Solution*"
2006-	Permanent position as "Ingénieur-Chercheur CEA" at IBS Grenoble

Fellowships

1999-2000	HERAEUS (German industrial foundation)
2000-2003	MENRT (French ministry PhD fellowship)
2005	DFG Fellowship "*Strukturbestimmung eines Protein-RNA Komplexes mittels Kernspinresonanz und Kleinwinkelstreuung*"

Current funding

2008-	Participant in four French ANRs
2008-	Participant in NMI3

Teaching experience

October 2007	Lecturer at HERCULES Specialized Course HSC6 "*Advances and new applications of synchrotron radiation for structural biology*" (Grenoble, France)
2008-	Lecturer at Master 2 course "*Méthodes Physiques et Physico-Chimie*" at UJF University (Grenoble, France)
June 2008	Introductory lecture in the "*Structure and Dynamics of Biological Macromolecules*" Master 2 Module at the Doctoral School CSDV (Grenoble, France)
September 2008	Lecturer at Les Houches Predoc session "*Biologie Structurale*" (Les Houches, France)
2009-	Lecturer at HERCULES Session B Biology "*Small angle neutron scattering"* and *"Protein dynamics by neutron scattering*" (Grenoble, France)
2009-	Lecturer at PSB Course "*Biological Small Angle Scattering with X-rays and Neutrons*" (Grenoble, France)

Invitations as speaker to national and international conferences

May 2004	*12^{e} Journées de neutronique "Neutrons et Biologie"* (Praz-sur-Arly, France)
April 2006	*ILL Millenium Symposium* (Grenoble, France)
October 2006	*Ecole Eau* (Roscoff, France)
January 2007	*Japanese-French seminar "Structural dynamics of Proteins"* (Grenoble, France)
February 2009	*"Neutrons and X-rays meet Biology"* (Berlin, Germany)
March 2009	*NESY Winterschool* (Planneralm, Austria)
June 2009	*"Neutrons in Biology"* (Lund, Sweden)
September 2009	*FEBS Course "Sofia School of Protein Science"* (Sofia, Bulgaria)

Participation in panels and other scientific activity

2007-2009	Member of ILL Subcommittee 8 "*Biology*"
2007-	External coordinator of several BAGs on *D22*
2007-	Representative of IBS in the PSB "*Small angle scattering platform*"
2009-	Co-organizer of PSB Course on "*Biological Small Angle Scattering with X-rays and Neutrons*" (Grenoble, France)
2010-	French ANR reviewer ("expert") in the Life Sciences

Peer review for scientific journals

Referee for *Phys. Rev. Lett., J. Mol. Biol., Biophys. J., Phys. Rev. E, Eur. Biophys. Lett.*

Supervision of PhD students

2007-2008 PhD co-supervisor of Dr. Soizic Masson at IBS ("*Structural study of a complex of three proteins of the Pneumococus division, DivB, DivIC and FtsL*")

2007-2008 Supervision of small angle scattering part of thesis "*Structural, biophysical and enzymatic characterization of a piezo-thermophilic peptidase complex from the deep sea archeon* Pyrococcus horikoshii" at IBS (Dr. Eva Rosenbaum)

Participation in thesis committees

2008 Participant as co-supervisor of Dr. Soizic Masson: "*Structural study of a complex of three proteins of the Pneumococus division, DivB, DivIC and FtsL*"

2009 "Examinateur" for Dr. Andreas Stadler: "*Molecular Dynamics of Red Blood Cells*"

Awards

2000 "Prix DEA" awarded by the Deutsch-Französische Hochschulkolleg

2004 "Prix Thèse" awarded by the Deutsch-Französische Hochschulkolleg

Memberships

Deutsche Physikalische Gesellschaft (DPG)
Société Française de Biophysique (SFB)

Languages

German (native speaker)
English, French, Italian (fluent)
Spanish (intermediary level)
Portuguese, Russian (notions)

Personnel interests

Interests include running, hiking, cooking and languages

Publication list 2002-2010

* = corresponding author

22) Salvay, A.G., Gabel, F., Fabiano, A.-S., Santos, J., Howard, E.I. and Ebel, C. (2010). Characterization of fish type III antifreeze protein in solution: structure and function. *Biophys. J.* **[in press]**

21) Falb, M., Amata, I., Gabel, F., Simon, B. and Carlomagno, T. (2010). The K-turn motif is not a constitutive RNA structural element: a combined NMR and SANS study. *Nucl. Acids Res.* **[in press]**

20) Wood, K., Tobias, D., Kessler, B., Gabel, F., Oesterhelt, D., Mulder, F., Zaccai, G. and Weik, M. (2010). The low temperature inflection observed in neutron scattering measurements of proteins is due to methyl rotation: direct evidence using isotope labelling and molecular dynamics simulations. *J. Am. Chem. Soc.* **132**(14), 4990-4991.

19) Masson, S., Kern, T., Le Gouëllec, A., Giustini, C., Simorre, J.-P., Callow, P., Vernet, T., Gabel, F. and Zapun, A. (2009). The central domain of DivIB caps the 1 C-terminal regions of the FtsL/DivIC coiled-coil rod. *J. Biol. Chem.* **284**(40), 27687-27700.

18) Breyton, C., Gabel, F., Abla, M., Pierre, Y., Lebaupain, F., Durand, G., Popot, J.-L., Ebel, C. and Pucci, B. (2009). Micellar and biochemical properties of (hemi)fluorinated surfactants are controlled by the size of the polar head. *Biophys. J.* **97**(4), 1077-1086.

*17) Gabel, F., Ringkjøbing Jensen, M., Zaccai, G. and Blackledge, M. (2009). Quantitative Modelfree Analysis of Urea Binding to Unfolded Ubiquitin using a Combination of Small Angle X-ray and Neutron Scattering. *J. Am. Chem. Soc.* **131**(25), 8769-8771.

16) Durá, M.A., Rosenbaum, E., Larabi, A., Gabel, F., Vellieux, F.M.D. and Franzetti, B. (2009). The structural and biochemical characterizations of a novel TET peptidase complex from *Pyrococcus horikoshii* reveal an integrated peptide degradation system in hyperthermophilic Archea. *Mol. Microbiol.* **72**(1), 26-40.

*15) Gabel, F., Masson, P., Froment, M.-T., Doctor, B. P., Saxena, A., Silman, I., Zaccai, G. and Weik, M. (2009). Direct correlation between molecular dynamics and enzymatic stability: a comparative neutron scattering study of native human butyrylcholinesterase and of its "aged" soman conjugate. *Biophys. J.* **96**(4), 1489-1494.

14) Frölich, A., Gabel, F., Jasnin, M., Lehnert, U., Oesterhelt, D., Stadler, A. M., Tehei, M., Weik, M., Wood, K. and Zaccai, G. (2009). From shell to cell: neutron scattering studies of biological water dynamics and coupling to activity. *Faraday Discuss.* **141**, 117-130.

13) Gabel, F., Simon, B., Nilges, M., Petoukhov, M., Svergun, D. and Sattler, M. (2008). A structure refinement protocol combining NMR residual dipolar couplings and small angle scattering restraints. *J. Biomol. NMR* **41**(4), 199-208.

12) Wood, K., Plazanet, M., Gabel, F., Kessler, B., Oesterhelt, D, Zaccai, G. and Weik, M. (2008). Dynamics of hydration water in deuterated purple membranes explored by neutron scattering. *Eur. Biophys. J.* **37**(5), 619-626.

11) Wood, K., Plazanet, M., Gabel, F., Kessler, B., Oesterhelt, D., Tobias, D.J., Zaccai, G. and Weik, M. (2007). Coupling of protein and hydration-water dynamics in biological membranes. *Proc. Natl. Acad. Sci. USA,* **104**(46), 18049-18054.

*10) Gabel, F. and Bellissent-Funel, M.-C. (2007). *C-phycocyanin* hydration water dynamics in the presence of trehalose: an incoherent elastic neutron scattering study at different energy resolutions *Biophys. J.* **92**(11), 4054-4063.

9) Tehei, M., Franzetti, B., Wood, K., Gabel, F., Fabiani, E., Jasnin, M., Zamponi, M., Oesterhelt, D., Zaccai, G., Ginzburg, M. and Ginzburg, B.-Z. (2007). Neutron scattering reveals extremely slow cell water in a Dead Sea organism. *Proc. Natl. Acad. Sci. USA* **104**(3), 766-771.

8) Gabel, F., Wang, D., Madern, D., Sadler, A., Dayir, K., Zamanian Daryoush, M., Schwahn, D., Zaccai, G., Lee, X. and Williams, B. R. G. (2006). Dynamic flexibility of human double-stranded RNA activated PKR molecules in solution. *J. Mol. Biol.* **359**(3), 610-623.

7) Pimienta, G., Gabel, F., Zanier, K., Conti, E. and Sattler, M. (2006). Chemical shift backbone assignments of TAP-N, the cargo-binding region of the protein TAP. *J. Biomol. NMR* **36** Suppl 5, 23.

*6) Gabel, F., Simon, B. and Sattler, M. (2006). A target function for quaternary structural refinement by combining small angle scattering and NMR orientational restraints. *Eur. Biophys. J.* **35**(4), 313-327.

*5) Gabel, F. (2005). L'influence du solvant sur la dynamique interne de la butyrylcholinestérase et sur la dynamique de l'eau d'hydratation : une étude par diffusion élastique incohérente de neutrons. *J. Phys. IV.* **130**, 133-140.

4) Gabel, F., Weik, M., Doctor, B. P., Saxena, A., Fournier, D., Brochier, L., Renault, F., Masson, P., Silman, I. and Zaccai, G. (2005). Effects of soman inhibition and of structural differences on cholinesterase molecular dynamics: a neutron-scattering study. *Biophys. J.* **89**(5), 3303-3311.

*3) Gabel, F. (2005). Protein dynamics in solution and powder measured by incoherent elastic neutron scattering: the influence of Q-range and energy resolution. *Eur. Biophys. J.* **31**(1), 1-12.

2) Gabel, F., Weik, M., Doctor, B. P., Saxena, A., Fournier, D., Brochier, L., Renault, F., Masson, P., Silman, I. and Zaccai, G. (2004). The influence of solvent composition on global dynamics of human butyrylcholinesterase powders: a neutron-scattering study. *Biophys. J.* **86**(5), 3152-3165.

1) Gabel, F., Bicout, D., Lehnert, U., Tehei, M., Weik, M. and Zaccai, G. (2002). Protein dynamics studied by neutron scattering. *Quart. Rev. Biophys.* **35**, 327-367.