

Brownian Dynamics of Globular Proteins  
under Macromolecular Crowding  
as Studied by NMR

Dissertation

zur Erlangung des Doktorgrades der Naturwissenschaften  
(Dr. rer. nat.)

der

Naturwissenschaftlichen Fakultät II  
Chemie, Physik und Mathematik

der Martin-Luther-Universität  
Halle-Wittenberg

vorgelegt von

Matthias Roos, M.Sc.  
geb. am 27. März 1988 in Gelnhausen



Erstgutachter: Prof. Dr. Kay Saalwächter (MLU Halle-Wittenberg)

Zweitgutachter: Prof. Dr. Wolfgang Paul (MLU Halle-Wittenberg)

Drittgutachter: Prof. Dr. Frank Schreiber (Universität Tübingen)

Tag der öffentlichen Verteidigung: 16. Dezember 2016

*"It seems likely [...] that macromolecular crowding in vivo is involved in many aspects of cellular function. Given that crowding was not discovered yesterday, one may wonder why little mention of this phenomenon is found in current textbooks of biochemistry and molecular cell biology."*

R. John Ellis and Allen P. Minton  
*Nature* (2003) | Vol. 425 | pp. 27-28



**ABSTRACT.** *In vivo* molecular motion of biopolymers is known to be strongly influenced by excluded-volume effects caused by the high concentration of organic matter inside cells, usually referred to as crowding conditions. Current literature regarding this topic is diverse, at points even contradictory to one another, and does not permit conclusions on the effect of concentrated proteins on protein diffusivity. In order to advance our understanding of the consequences of macromolecular crowding for protein translational and rotational diffusion, and to establish possible analogies to the behavior of dense colloidal suspensions, pulsed-field gradient (PFG) NMR, NMR relaxometry, X-ray scattering and viscosity measurements were combined for three concentrated proteins in aqueous solution –  $\alpha$ B-crystallin, bovine serum albumin (BSA) and lysozyme. Translational diffusion was also investigated for a protein mixture composed of BSA and a fourth protein, the Src-homology 3 (SH3) domain. The challenges posed by a non-exponential behavior of the protein Brownian tumbling auto-correlation function arising from significant protein-protein interactions under crowding conditions were accessed via NMR relaxometry including measurements at low magnetic fields. The combined set of  $R_1$ , and  $R_{1\rho}$  and  $R_2$  relaxation measurements overcame obstacles encountered in previous NMR studies on (crowded) protein rotational diffusion. Limitations of PFG NMR at high protein concentrations are also addressed. A combination of all approaches enabled a clear, comprehensive physical picture of Brownian motion in highly concentrated protein solutions.

The experimental data on the single-protein solutions demonstrate, on one hand, that *long*-time translational diffusion quantitatively follows the expected increase of macroviscosity upon increasing protein concentration, given that no transient protein-protein binding occurs. The behavior of rotational diffusion, on the other hand, turns out to be generally protein-specific, and spans the full range of limiting cases from fully coupled to completely decoupled from the macroviscosity. Here, the experimental data indicate that anisotropic inter-protein interactions – in particular those of electrostatic nature – give the main factor modulating the (de)coupling between rotational and long-time translational diffusion. Literature data on *short*-time translational diffusion, in turn, show a similar concentration dependence as rotational diffusion. Contrary to the single-protein solutions, for a mixture of BSA and the SH3 domain, presence of crowding-induced, transient inter-protein binding causes translational diffusion to be more retarded than expected from the macroviscosity, accompanied by a significantly increased apparent activation energy as compared to the macroscopic viscosity. Altogether, protein-protein electrostatic interactions and presence or absence of transient inter-protein binding have been identified as crucial factors for the behavior of protein translational and rotational diffusion under crowding conditions, providing an explanation for the diverging results in current literature.

---

**KEY TOPICS:** proteins | crowding | translational and rotational diffusion | transient binding

**SAMPLES:**  $\alpha$ B-crystallin | bovine serum albumin | hen egg white lysozyme | SH3 domain



1	INTRODUCTION:	
	<i>Motivation for understanding the consequences of macromolecular crowding – in vivo conditions revisited</i>	PAGE 11 FF.
2	SCIENTIFIC BACKGROUND:	
	<i>Brownian motion in dense colloidal suspensions</i>	PAGE 19 FF.
2.1	The generalized Stokes-Einstein and Stokes-Einstein-Debye equations: diffusion of concentrated colloids	19
2.2	The rotational auto-correlation function in non-dilute protein solutions: deviation from singly exponential behavior	26
3	EXPERIMENTAL METHODS:	
	<i>NMR spectroscopy and viscosimetry</i>	PAGE 33 FF.
3.1	Basics of NMR	33
3.2	The NMR HSQC experiment	35
3.3	NMR relaxometry	37
3.4	Pulsed-field gradient NMR	50
3.5	Isotope filtering for the measure of protein mixtures	55
3.6	Viscosimetry	56
3.7	Concentration determination	59

#### 4 RESULTS AND DISCUSSION:

##### *The impact of macromolecular crowding on protein diffusivity* \_\_\_\_\_ PAGE 61 FF.

4.1	NMR-detected Brownian dynamics of $\alpha$ B-crystallin over a wide range of concentrations	61
4.2	The “long tail” of the protein tumbling correlation function: observation by $^1\text{H}$ NMR relaxometry in a wide frequency and concentration range	73
4.3	Coupling and decoupling of rotational and translational diffusion of proteins under crowding conditions	89
4.4	Transient binding accounts for apparent violation of the generalized Stokes–Einstein relation in crowded protein solutions	99

#### 5 SUMMARY, CONCLUSIONS AND OUTLOOK:

##### *The key factors that determine the Brownian motion of crowded proteins* \_\_\_\_\_ PAGE 111 FF.

#### S SUPPLEMENT:

##### *Supporting material, bibliography* \_\_\_\_\_ PAGE S-1 FF.

S.1	Notes on spin precession	S-1
S.2	Why using the arithmetically averaged relaxation rate?	S-1
S.3	Limitations of PFG NMR in crowded protein solutions	S-4
S.4	Viscosity determination via a simple capillary	S-11
S.5	Supplementary material to the research articles	S-13
S.6	Bibliography	S-53
S.7	Curriculum vitae, publications, presentations, statutory declaration	S-69



## List of Figures

1.1	The excluded volume effect . . . . .	12
1.2	Illustration of protein Brownian motion under macromolecular crowding . . .	15
1.3	Native structure of the $\alpha$ B-crystallin assembly, bovine serum albumin, lysozyme, and the SH3 domain . . . . .	18
2.1	Schematic representation of $D_t^S$ and $D_t^L$ and appropriate viscosities $\eta_\infty$ and $\eta_0$ respectively, for spherical particles of same size . . . . .	24
2.2	Illustration of protein electrostatic steering . . . . .	28
2.3	Schematic rotational auto-correlation function for non-dilute protein solutions	30
3.1	$^1\text{H}$ NMR spectrum of concentrated $\alpha$ B-crystallin, bovine serum albumin, and lysozyme . . . . .	35
3.2	NMR pulse sequence of the HSQC experiment . . . . .	36
3.3	$^1\text{H}$ - $^{15}\text{N}$ HSQC spectrum of the SH3 domain in absence and presence of bovine serum albumin . . . . .	36
3.4	Energy levels for a two-spin system. . . . .	38
3.5	Spectral density of protein rotational diffusion . . . . .	40
3.6	Frequency dependence of the spectral density for different values of the correlation time $\tau$ . . . . .	41
3.7	Dependence of $R_1$ , $R_{1\rho}$ and $R_2$ on the rotational correlation time . . . . .	42
3.8	Intrinsic distribution of longitudinal relaxation rates in lysozyme . . . . .	43
3.9	Exemplary data set for NMR longitudinal relaxation . . . . .	45
3.10	The principle of field-cycling NMR . . . . .	46
3.11	Exemplary data set for NMR transverse relaxation . . . . .	48
3.12	Effective field strength in the rotating frame and the pulse scheme for on-resonant $R_{1\rho}$ measurements . . . . .	49
3.13	NMR pulse sequences for measuring translational diffusion . . . . .	52
3.14	Exemplary PFG NMR diffusion decays for bovine serum albumin . . . . .	54
3.15	NMR isotope filtering: pulse sequence and exemplary application . . . . .	56
3.16	Shear-rate dependence of the viscosity of BSA solutions . . . . .	58
3.17	Concentration dependence of the viscosity of BSA solutions . . . . .	59

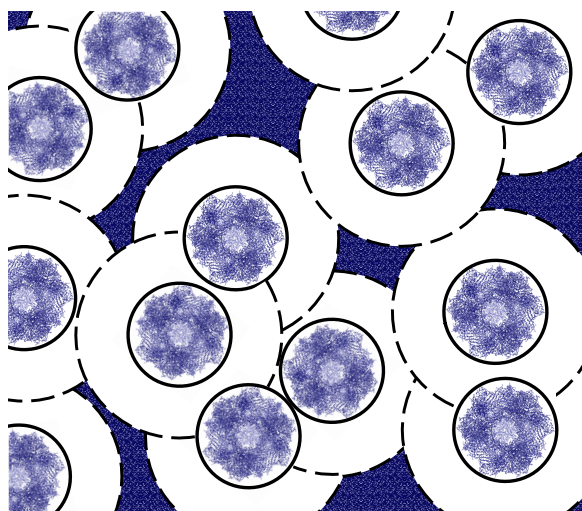
S1	Exemplary PFG NMR diffusion decay including the water signal . . . . .	S-5
S2	The principle of spectral hole burning . . . . .	S-6
S3	Pulse sequence for studying spin polarization exchange between aliphatic protein and residual water protons . . . . .	S-7
S4	Magnetization exchange profiles between protein amid and water protons . . .	S-9
S5	Magnetization exchange time for bovine serum albumin and $\alpha$ B-crystallin as a function of the protein concentration and the temperature . . . . .	S-10
S6	Comparison of magnetization exchange curves using a shaped and a rectan- gular pulse . . . . .	S-11
S7	Viscosity determination using a capillary . . . . .	S-12
S8	Viscosity of bovine serum albumin and lysozyme solutions as determined by a capillary . . . . .	S-12

# 1 Motivation for understanding the consequences of macromolecular crowding – *in vivo* conditions revisited

In a living cell, proteins are surrounded by and interact with a multitude of different macromolecules on intermolecular distances comparable to their size. Typically, the plethora of macromolecules, such as proteins, nucleic acids and polysaccharides, amount to a total mass of several hundreds of grams per liter [Zimmerman and Trach, 1991], providing a concentration which is a factor of tens to hundreds higher than what is usually studied in biophysical and biochemical sciences. Given such an amassment of macromolecules, these occupy 20-40% of the cytoplasmic volume [Ellis, 2001b; Medalia et al., 2002; Ellis and Minton, 2003], and a large volume fraction is excluded for any additional macro-solute (cf. Fig. 1.1). Thus, the high amount of organic matter per volume found in the interior of cells is often referred to as “macromolecular crowding”. Crowding by macro-solutes, however, is not restricted to the cell’s interior: also the extracellular tissue, consisting of a matrix of polysaccharides, collagen and other macromolecules, is subject to crowding effects, and even blood plasma contains about 80 g/l protein [Ellis, 2001a].

An outstanding example for crowding conditions is the eye lens. Here, an extraordinary highly concentrated, polydisperse protein solution serves to achieve a high refractive index and lens transparency. *In vitro* studies of eye lens cell extracts show increasing transmission and reduced light scattering with increasing protein concentration, given that the protein concentration exceeds ~100 g/l [Delaye and Tardieu, 1983; Tardieu, 1998]. Below this value, light scattering increases with protein concentration. The high concentration of proteins is part of biological function, and is further optimized by a concentration gradient from the center to the edge of the eye lens. In the center of the human eye lens, a protein concentration of about 450 g/l can be found [Fagerholm et al., 1981]; the protein concentration inside the glassy fish lens can be even as high as ~1050 g/l [Bloemendal et al., 2004]. The interior of the fibre cells of the eye lens is thus almost completely composed of proteins termed crystallins, which are in some parts as old as the individual: there is neither protein turnover inside the fibre cells, nor do proteins diffuse from one fibre cell to another [Ecroyd and Carver, 2009]. Thus, posterior modifications of a protein’s primary structure such as oxidation, deamidation or cleavage cannot be reversed, and misfolded

## 1. Introduction



**Figure 1.1.:** Excluded volume in concentrated protein solutions, plotted for the example of the eye-lens protein  $\alpha$ B-crystallin. For each of the equal spheres (solid lines) representing the protein, the closest possible distance to the center of another sphere is twice the sphere radius (dashed lines). The space at which the center of another protein may be located (blue area, textured) is strongly restricted.

$\alpha$ B-crystallin was plotted using PyMol and the protein data bank structure 2ygd [Braun et al., 2011].

or aggregated proteins cannot be replaced. Any of such detrimental circumstances can be followed by serious aging effects. Once there is a formation of structures large enough to cause light scattering, lens clouding will occur even if the fraction of these scattering centers may be small. Cataracts, in fact, are the most frequent cause of blindness [Pascolini and Mariotti, 2012], but are not understood on a molecular level: the nature of cataract formation is still a matter of debate. Protein aggregation, amyloid formation, protein crystallization and liquid-liquid phase separation are discussed as sources of cataract formation. The importance of well-balanced protein-protein interactions for keeping the eye lens transparent is further stressed by the fact that lens transparency is highly related to a short-range liquid order of the crystallins [Delaye and Tardieu, 1983]. Such short-range molecular organization is, again, strongly related to crowding, stressing once more the scientific need of understanding the physics of macromolecular crowding.

The combined action of excluded-volume effects and the particularly pronounced protein-protein interactions render protein behavior under crowding conditions essentially different from dilute solutions [Luby-Phelps, 2000] with substantial consequences for cellular function and organization [Zimmerman and Minton, 1993; Zhou et al., 2008]. These effects include altered energetic stabilization of protein folding: non-specific (e.g. electrostatic) protein-protein interactions that tend to destabilize the protein's folded structure are, to different extents, counterbalanced by structure-stabilizing excluded volume effects [Zhou, 2004; Miklos et al., 2011; Schlesinger et al., 2011; Mittal et al., 2015], with the latter favoring compact states (Le Chatelier's principle; see also Hong and Gierasch [2010], and Soranno et al. [2015]). Both kinds of interactions become more pronounced the higher the concentration of the proteins. In this context, altered protein refolding rates [van den Berg et al., 1999; Cheung et al., 2005; Minton, 2005] and the stabilization [Cheung et al., 2005; Miklos et al., 2011; Senske et al., 2014] or destabilization [Danielsson et al., 2015; Mittal et al., 2015] of the compact folded state of a protein have been reported, including an increased tendency of protein association and aggregation [van den Berg et al., 1999; Minton, 2000, 2005; Ellis and Minton, 2006]. Several proteins that are intrinsically disor-

dered under dilute conditions attain a fold structure upon crowding, whereas others do not [Dedmon et al., 2002]. Not only changes in the folding energetics were reported [Hong and Gierasch, 2010], but also reaction rates and reaction equilibria have been discussed to differ significantly – potentially even by orders of magnitude [Minton, 2001] – under crowding conditions from those determined in dilute solution [Minton, 1997, 2006; Ralston, 1990; Ellis, 2001a]. These effects are accompanied by limited protein expression and, ultimately, limited cell growth [Klumpp et al., 2013]. Likewise, macromolecular crowding has been found to decrease intracellular signaling, active protein transport, metabolism and cell fluidity, as is reviewed by Mourão et al. [2014].

As a recent development originating from the compaction of (bio)polymers by the excluded volume effect, molecular sensors [Gnutt et al., 2015; Boersma et al., 2015] have been designed for mapping excluded volume effects inside the living cell, reflecting some degree of spatial heterogeneity [Gnutt et al., 2015]. Theory and simulations on macromolecular crowding rely to a large extent on statistical mechanics, and are qualitatively, but not always sufficiently quantitatively, in accordance with experiments, as reviewed by Elcock [2010]. In fact, just within the last years, increased computational power allowed the field to advance from rigid-body modeling of crowded proteins [Długosz and Trylska, 2011] to performing extensive, rather realistic atomic detail simulations [McGuffee and Elcock, 2006, 2010; Mereghetti et al., 2010; Mereghetti and Wade, 2012; Balbo et al., 2013].

The biochemical consequences of the excluded volume effect have primarily been explored since the 1990s, subsequently followed by considerations of specific as well as non-specific protein-protein interactions. (See, for instance, Minton [2006]; McGuffee and Elcock [2006, 2010]; Feig and Sugita [2012]; Sapir and Harries [2015]; Gnutt and Ebbinghaus [2016], or the review by Hall and Minton [2003] discussing general limitations of hard-sphere modeling.) Earlier [Shastry and Eftink, 1996] as well as recent [Senske et al., 2016] findings on proteins confined in reverse micelles contradict the behavior predicted by *exclusively* considering hard-core excluded volume effects. Instead, the importance of crowding or confinement driven enthalpic effects needs to be taken into account rather than focusing on entropic contributions only [Senske et al., 2014, 2016].

The most obvious effect of crowding is its impact on Brownian motion: the large extent of excluded volume combined with rather pronounced protein-protein interactions restricts molecular displacements. Ando and Skolnick [2010] showed that *in vivo* molecular motion is dominated by macromolecular crowding effects, with the detailed impact of macromolecular crowding being only marginally understood on a molecular level. This lack of knowledge particularly regards information beyond the pure slowdown of protein diffusivity with increasing protein concentration [Ellis, 2001a; Luby-Phelps, 2000]. Specifically being relevant for biological systems, binding events may retard diffusion beyond pure steric effects [Verkman, 2002], and the multitude of protein-protein interactions may account for rather sophisticated molecular dynamics. The slow-down of protein diffusion,

## 1. Introduction

in turn, impacts biochemical reactivity: diffusion-limited reactions may become slower with increasing concentration of the crowding agent. In contrast, chemical reactions depending on the orientation of the reactants may become more efficient under crowding [Kim and Yethiraj, 2009]. For these orientation-dependent reactions, rotational diffusion combined with attractive and/or repulsive intermolecular forces serve to achieve proper inter-molecular alignments [Shoup et al., 1981]. The extent to which macromolecular crowding reduces rotational relative to translational motion readjusts the time span during which transient contacts allow for orientational justification [Kuttner et al., 2005; Kim and Yethiraj, 2009]. If rotational diffusion is enhanced over translational displacements, more orientations will be achieved during transient contacts, which may positively affect protein reactivity. This effect is of particular interest for specific binding among proteins. In fact, due to a protein's anisotropic reactivity [Barzykin and Shushin, 2001], the fraction to which encounters result into specific binding can be rather small; for instance, only about 2% of the encounters of an actin monomer with the end of an filament finally result in binding [Drenckhahn and Pollard, 1986]. Such a situation underlines the importance of molecular mobility for cellular function and organization.

Ultimately, understanding molecular biology *in vivo* requires elucidating the consequences of macromolecular crowding on protein Brownian diffusion. Systematic studies on protein Brownian motion under macromolecular crowding have remained sparse, and are at points even contradictory to one another. For instance, with increasing protein concentration, protein rotational diffusion was reported to be either less [Zorilla et al., 2007] or to be even more [Wang et al., 2010] retarded than translational diffusion; the former case being supported by atomic detail Brownian dynamics simulations [Mereghetti and Wade, 2012].

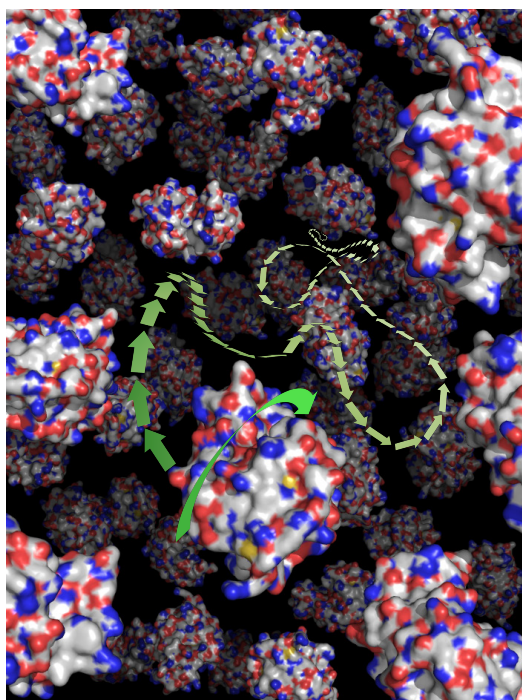
Just as for rotational diffusion, findings on crowded protein translational diffusion are likewise contradictory. Under crowding conditions, protein translational diffusion has been reported as being either slower [Zorilla et al., 2007] or even faster [Li et al., 2009; Wang et al., 2010] than expected from the macroscopic viscosity. An earlier study, in turn, convincingly demonstrates an agreement of the concentration dependence of protein translational diffusion with the macroscopic viscosity [Licinio and Delaye, 1988].

Such discrepancies in experimental results indicate sample specificity with regard to both the studied protein and the crowding agent. Moreover, the protein rotational auto-correlation function has been shown to deviate from mono-exponential behavior for non-dilute samples, challenging experimental studies [Krushelnitsky, 2006]. So-called electrostatic steering effects resulting from the non-uniform charge distribution within each protein are considered to impact re-orientational dynamics in that they render protein rotational diffusion transiently anisotropic [Krushelnitsky, 2006]. That is, on the time scale of rotational motion, some molecular orientations can be considered to be preferred over others, with the preferred alignment resulting from the mean electric field originating

from surrounding proteins. This effect becomes more pronounced the closer the proteins get to each other and is hence of particular relevance for concentrated protein solutions. However, experimental limitations did not allow measurements on the amplitude and time-scale of this effect independently from each other [Krushelnitsky, 2006]. Overall, beyond the general slow-down of protein Brownian diffusion, the detailed behavior of protein Brownian motion under crowding conditions remains unresolved so far.

Specifying the impact of macromolecular crowding on protein translational *and* rotational diffusion is the central objective of this thesis as illustrated in Fig. 1.2.

In the study at hand, crowding was achieved solely by globular, native-state proteins, and not by unstructured (bio-)polymers. These two classes of macromolecules provide a qualitatively different nature of crowding: unlike globular proteins, unstructured (bio-)polymers do not form independent hydrodynamic entities if a critical concentration, the so-called overlap concentration  $c^*$ , is exceeded. Instead, for  $c > c^*$  random coils interpenetrate each other and form an entangled, mesh-like medium that has to be described by polymer physics [Rubinstein and Colby, 2003; Strobl, 2007] rather than by concepts in colloid science. At concentrations mimicking crowding conditions, these effects become relevant. In fact, about 25% of mammalian proteins are predicted to be intrinsically disordered proteins [Dyson and Wright, 2005; Dunker et al., 2008] that are well described by polymer physics scaling laws [Hofmann et al., 2012; Soranno et al., 2015]. The majority of biological macromolecules, however, has a well-defined folded structure. *In vivo* macromolecular crowding can thus be considered to consist of both crowding by folded and



**Figure 1.2.:** Illustration of protein Brownian motion under macromolecular crowding by like proteins. Picture created by repetitive plots of lysozyme (protein data bank ID 1LYZ [Diamond, 1974]) using Pymol.

Although the figure reflects some degree of crowding, the depicted protein concentration ( $\sim 10$  vol%) is still below the highest concentrations studied within this thesis. For lysozyme, concentrations up to  $\sim 20$  vol% ( $\sim 250$  g/l) were studied, i.e, concentrations close to those in the usual cellular environment. The eye lens gives an example of an extraordinarily high protein concentration – here, the overall protein volume fraction is even much higher.

## 1. Introduction

unfolded structures in combination with the excluded volume effect of the cytoskeleton. Unlike protein induced crowding, however, synthetic crowding agents do not properly account for electrostatic interactions, and were recently identified as “poor mimics of the cellular interior” [Smith et al., 2016].

Thus, highly concentrated solutions of *globular proteins* served as model systems for macromolecular crowding conditions, aiming to answer the following questions:

- (i) how do translational and rotational diffusion behave upon increases in protein concentration, in particular, how do they behave with regard to the macroscopic solution viscosity?
- (ii) do translational and rotational diffusion couple to each other under crowding conditions, or are they decoupled?
- (iii) how does the deviation from singly exponential behavior of the rotational auto-correlation function change upon increasing protein concentration?

For this purpose, pulsed-field gradient (PFG) Nuclear Magnetic Resonance (NMR) spectroscopy was combined with NMR relaxometry for studying protein Brownian dynamics on the molecular level. They are advantageous in that they do not require the use of specific tracer molecules, nor do they require structurally modified proteins for signal detection (as required in many techniques, for instance, in fluorescence spectroscopy). In particular, field-cycling NMR was directly applied to protein protons for an unbiased detection of the overall rotational auto-correlation function, including its long-time (i.e., low-frequency) behavior; the latter being required for studying the transient anisotropy of rotational diffusion (the deviation from a singly exponential correlation function). Due to an intrinsically low sensitivity of this technique, field-cycling NMR was to date only rarely directly applied to protein protons [Bertini et al., 2005; Luchinat and Parigi, 2007]. In regards of the non-exponential nature of the rotational auto-correlation function of non-dilute proteins, methodological progress has been achieved, and will be presented in the RESULTS section. Beside these techniques, viscosity was measured with a viscometer requiring only microliter samples. Moreover, small-angle X-ray scattering (SAXS) experiments performed and evaluated by MARIA OTT<sup>‡</sup> complemented the physical picture by probing the structure factor of the protein solutions, allowing estimations for the strength of intermolecular interactions. Polarized fluorescence correlation spectroscopy (FCS), also performed by MARIA OTT<sup>‡</sup>, reinforce the NMR results presented in this thesis.

At any time mentioned in the thesis the term “crowding” refers to crowding by macromolecules. “Crowding” as achieved by adding small or medium size co-solutes, such as glycerol or peptides, will not be addressed. The high volume fraction of macromolecules is referred to as “crowded” or “confining” rather than “concentrated” to account for the

---

<sup>‡</sup>Martin Luther University Halle-Wittenberg, Experimental Polymer Physics



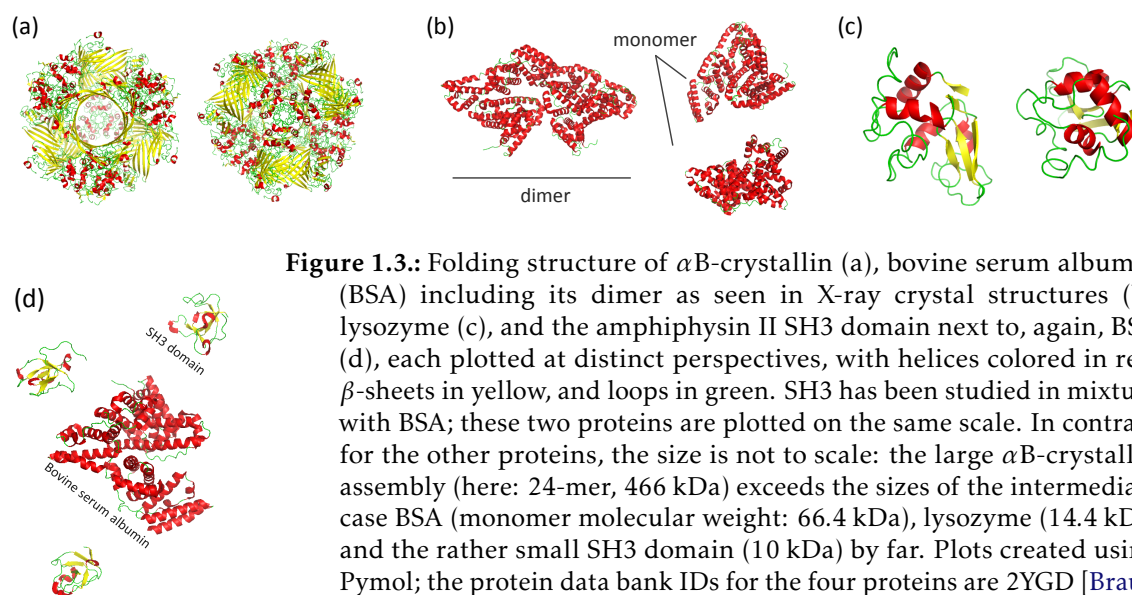
mixture of different macro-solutes present under natural conditions [Ellis, 2001b; Minton, 2001]. A *single* protein, in contrast, hardly occurs *in vivo* at such high concentrations. As an example, in a typical eukaryotic cell the concentration of free actin (i.e, actin that is not incorporated in microfilaments) is about 4 g/l [Bray and Thomas, 1975], although the overall protein concentration even reaches 300-400 g/l [Zimmerman and Trach, 1991]. Hemoglobin within red blood cells represents a famous exception, as it occurs at a concentration of even about 350 g/l [Zimmerman and Minton, 1993]. The eye lens contains different types of crystallin proteins each at high concentrations. Of these proteins the behavior of concentrated  $\alpha$ B-crystallin will be investigated in this thesis.

Accounting for the multitude of different proteins dispersed in the cytoplasm, and in the context of an increased interest in the excluded volume effect, the term “crowding” developed as a key word for mimicking the high volume fraction of macromolecules *in vivo*.

*In vivo* experiments are rewarding for quantifying biological process under native conditions, but are difficult to interpret due to the multitude of influencing factors. Thus, a “bottom-up” approach appears to be most efficient for understanding the mechanisms related to (bio-)macromolecular crowding [Elcock, 2010]. In particular, for assessing the above questions, and thus for achieving insights beyond the simple slow-down of Brownian motion, one *has to* rely on *in vitro* studies to allow for a well defined experimental parameter space. Thus, the studies presented here were performed *in vitro*, but aim at understanding the physics relevant for *in vivo* conditions. Thereby, concentrated solutions of a single protein (“homo-crowding”) were investigated first, followed by a study on protein mixtures (“hetero-crowding”) as the next step. In these studies,  $\alpha$ B-crystallin ( $\alpha$ Bc), hen egg white lysozyme (LYZ, HEWL), bovine serum albumin (BSA), and the Src-homology 3 (SH3) domain were used as test proteins, with fold structures plotted in Fig. 1.3.

*The thesis is organized as follows:* In the next chapter, Brownian dynamics of concentrated particles will be shortly discussed from the point of view of colloidal science where the generalized Stokes-Einstein and Stokes-Einstein-Debye equation for translational and rotational diffusion, respectively, will be introduced. Aspects of particular relevance for protein diffusion will be discussed there as well, addressing the current knowledge on the impact of the so-called “electrostatic steering effect” on protein rotational diffusion. After that, the experimental methods used for addressing the above questions will be presented in chapter 3, followed in chapter 4 by the results obtained. These are presented in a cumulative form of four research articles published in peer-reviewed journals. Finally, the last chapter summarizes the results obtained, and gives an outlook for potential subsequent studies.

## 1. Introduction



**Figure 1.3.:** Folding structure of  $\alpha$ B-crystallin (a), bovine serum albumin (BSA) including its dimer as seen in X-ray crystal structures (b), lysozyme (c), and the amphiphysin II SH3 domain next to, again, BSA (d), each plotted at distinct perspectives, with helices colored in red,  $\beta$ -sheets in yellow, and loops in green. SH3 has been studied in mixture with BSA; these two proteins are plotted on the same scale. In contrast, for the other proteins, the size is not to scale: the large  $\alpha$ B-crystallin assembly (here: 24-mer, 466 kDa) exceeds the sizes of the intermediate case BSA (monomer molecular weight: 66.4 kDa), lysozyme (14.4 kDa) and the rather small SH3 domain (10 kDa) by far. Plots created using Pymol; the protein data bank IDs for the four proteins are 2YGD [Braun et al., 2011], 4F5S [Bujacz, 2012], 1LYZ [Diamond, 1974], and 1bb9 [Owen et al., 1998], respectively.

# 2 Brownian motion in dense colloidal suspensions

### 2.1. The generalized Stokes-Einstein and Stokes-Einstein-Debye equations: diffusion of concentrated colloids

Today's understanding of Brownian motion relies on the pioneering work by [Einstein \[1905\]](#), [Smoluchowski \[1906\]](#) and [Langevin \[1908\]](#), who independently from each other showed that the mean-square displacement of a particle dispersed in a viscous medium scales with time  $t$  as

$$\langle |\mathbf{r}(t) - \mathbf{r}(0)|^2 \rangle = 6 D_t t \quad . \quad (2.1)$$

Here,  $D_t$  is the particle's (isotropic) translational (self-)diffusion coefficient, and  $\mathbf{r}(t) = (x, y, z)^T$  is its time-dependent position. For one- or two-dimensional diffusion, the prefactor of 6 in eq. (2.1) has to be simply replaced by 2 and 4, respectively. The Einstein-Smoluchowski relationship, in turn, relates the diffusion coefficient to the translational mobility  $\mu_t$  of the particle,

$$D_t = \mu_t k_B T \quad , \quad (2.2)$$

where  $k_B T$  is the thermal energy, composed of the Boltzmann constant and the absolute temperature. Analogously, for rotation about a *single* axis, the mean-square angular deviation of the particle's orientation angle  $\theta(t)$  is

$$\langle |\theta(t) - \theta(0)|^2 \rangle = 2 D_r t \quad , \quad (2.3)$$

$$D_r = \mu_r k_B T \quad , \quad (2.4)$$

where  $D_r$  and  $\mu_r$  are the rotational diffusion coefficient and rotational mobility, respectively. Exactly these translational and rotational mobilities,  $\mu_t$  and  $\mu_r$ , make the difference between dilute and concentrated solutions, and give rise to open questions when studying crowding effects – especially with regard to the open point of coupling or decoupling between rotational and translational diffusion in crowded protein solutions (i.e., sensitivity to the

## 2. Scientific background

same or distinct viscous effects, respectively, and thus also a shared or distinct concentration dependence of  $D_t$  and  $D_r$ ).

The mobility is determined by the viscous drag to which the tracer particle is exposed. The viscous drag of translational and rotational diffusion originates from a dissipative force  $\mathbf{F}$  and dissipative torque  $\mathbf{M}$ , respectively, which for laminar flow in a homogeneous medium both scale linearly with the translational velocity  $\mathbf{V}$  and its angular counterpart  $\mathbf{\Omega}$ , respectively:

$$\mathbf{F} = -f_t \mathbf{V} \quad , \quad \mathbf{M} = -f_r \mathbf{\Omega} \quad . \quad (2.5)$$

Here, the friction coefficients are usually determined by

$$f_t = 6\pi\eta R_H \quad \text{and} \quad f_r = 8\pi\eta R_H^3 \quad , \quad (2.6)$$

where  $\eta$  denotes the viscosity of the surrounding medium, and  $R_H$  is the radius of a smooth sphere. In this way, size and shape of the particle are expressed in terms of a sphere of equivalent hydrodynamic dimension. Given that the Reynolds number is low, as is the case for Brownian motion, the mobility scales inversely with that friction,  $\mu_{t,r} = 1/f_{t,r}$ .

The above equations take as a basis that the particle diffuses in a *continuous* medium. Provided that the tracer particle is much larger than the surrounding host particles, the description of the host fluid as a homogeneous medium is indeed a well-suited approximation. In the infinitely dilute limit, the medium viscosity equals the solvent viscosity  $\eta_{\text{sol}}$  such that the Stokes-Einstein (SE) and Stokes-Einstein-Debye (SED) relationship for translational and rotational diffusion, respectively, read

$$D_{t,0} = \frac{k_B T}{6\pi\eta_{\text{sol}} R_H} \quad \text{and} \quad D_{r,0} = \frac{k_B T}{8\pi\eta_{\text{sol}} R_H^3} \quad . \quad (2.7)$$

Interestingly, the Stokes-Einstein(-Debye) equation provides sufficient results even on a molecular level, in particular when introducing modified prefactors that account for apparent slip of the host particles on the surface of the tracer sphere [Hu and Zwanzig, 1974; Dote et al., 1981; Ould-Kaddour and Levesque, 2000],

$$f'_t = 6\nu_t \pi \eta_{\text{sol}} R_H = \nu_t f_t \quad \text{and} \quad f'_r = 8\nu_r \pi \eta_{\text{sol}} R_H^3 = \nu_r f_r \quad . \quad (2.8)$$

For perfect slip,  $\nu_t = 2/3$  and  $\nu_r = 0$  [Hu and Zwanzig, 1974]. Here, solvent molecules are not tagged along with the moving sphere, meaning that there would be no viscous friction for rotational diffusion. For translational diffusion, in contrast, viscous friction arises even under slip boundary conditions, as surrounding molecules must be displaced during movement of the tracer sphere. Experiments on molecular systems are in accordance with  $\nu_t \approx 2/3$ , and provide partial slip for rotational diffusion ( $0 < \nu_r < 1$ ) [Edward, 1970; Koenderink et al., 2002]. In addition, simulations indicated slip boundary conditions for translational diffusion in molecular systems [Ould-Kaddour and Levesque, 2000], though

## 2.1. The generalized Stokes-Einstein and Stokes-Einstein-Debye equations

contradictions with regard to the impact of the particle's mass were noticed [Walser et al., 1999]. Stick boundary conditions ( $v_{t,r} = 1$ , implied in eqs. (2.6) and (2.7)), in turn, assume that the first layer of surrounding molecules adheres to the diffusing sphere. This causes energy dissipation, and thus retards the motion of the tracer sphere. In reality, both rotation and translation of the particle will be accompanied by velocity gradients of surrounding solvent molecules, for which the above boundary conditions account in a simplified manner.

The flow of solvent around each diffusing colloidal particle mediates hydrodynamic interactions (HIs). HI is of relevance even when the (effective) hard-sphere radius of the particle<sup>1</sup> exceeds the hydrodynamic radius by several times [Cichocki and Felderhof, 1991]. For a non-dilute colloidal suspension, this situation accounts for an indirect, dissipative force between the colloids, i.e., colloid-colloid hydrodynamic friction. If the colloid is also charged, electrosteric friction will arise next to the bare solvent friction and the colloid-colloid HI [McPhie and Nägele, 2007]. Surrounding ions of the solvent ("micro-ions") and the charged colloids themselves ("macro-ions") cannot react instantaneously to the (Brownian) displacement of the charged colloid. This somewhat delayed feedback of surrounding ions to the motion of the charged colloid is accompanied by an electro-steric relaxation force [McPhie and Nägele, 2007] and, hence, electro-viscous friction. Since the macro-ions usually carry more charges than the (co- and counter-)ions of the solvent and since the realignment of macro-ions is slower than that of the micro-ions, the electro-viscous friction is in general dominated by the macro-ions and their dynamics [McPhie and Nägele, 2007].

With respect to the success of the SE and SED relationship in molecular systems, one may wonder about the validity of the SE and SED relationship for concentrated colloidal suspensions: both for molecular systems and concentrated colloidal suspensions, the tracer particle moves in an environment occupied by particles of similar or same size as the tracer particle, with a multitude of direct and indirect multi-body interactions. With ongoing time the inhomogeneity of the environment is progressively averaged out both by the translational displacement of the tracer particle itself and by the structural reconfiguration of the surrounding: the continual re-organization of the particle alignment, mediated by Brownian motion and averaged over time, reintroduces a continuous environment. Then, the Stokes-Einstein(-Debye) relationship may be expected to hold with regard to the macroscopic dispersion viscosity even for concentrated colloidal systems. This effect also provides the basis for the applicability of the Stokes-Einstein relationship to molecular liquids. In contrast, in the *short*-time limit, neither translational or rotational diffusion are subject to multiple encounters with neighboring particles, nor does sufficient averaging apply with respect to the inhomogeneous surrounding. A sensitivity to large-scale (i.e. predominately macroscopic) properties of the solution can be strongly questioned here. Instead, for particles much larger than the solvent molecules, short-time diffusion takes place within a local environment *enclosed* by surrounding particles. Here, solvent-related

---

<sup>1</sup>The radius at which the direct particle-particle interaction potential is effectively decayed.

## 2. Scientific background

properties and local hydrodynamics [Tokuyama and Oppenheim, 1994; Roosen-Runge et al., 2011] dominate the friction relevant for Brownian motion, where particle encounters only marginally contribute to the viscous friction. As a conceptual expression for the viscous drag within such a local environment and to distinguish it from the viscosity on a macroscopic scale, the denotation of a “micro-viscosity” evolved.

Consequently, which limiting case applies depends on the experimental time window  $\Delta\tau_{\text{tw}}$  compared to the time span required for significant reconfiguration of the concentrated suspension. For particles of the same size, the critical time scale separating between both scenarios can be denoted as the period required for diffusion over a typical inter-particle distance  $\xi$ , i.e.,  $\tau_{\text{crit}} \sim \xi^2/6D_{t,0}$  [Koenderink and Philipse, 2000]. At times much smaller than  $\tau_{\text{crit}}$  (i.e.,  $\langle r^2 \rangle \ll \xi^2$ ), translational displacements only take place within the local environment enclosed by surrounding particles, with no significance of caging effects [Doliwa and Heuer, 1998; Weeks and Weitz, 2002b]. Thus, the short-time diffusion coefficient can indeed be close to that in dilute solution [Feig and Sugita, 2012]. Note, however, that diffusivity over short times still implies that the considered time scale is larger than that of the momentum relaxation of the colloidal particle; otherwise, displacements are not diffusive but ballistic. In the ballistic regime,  $\langle r^2 \rangle = \bar{v}^2 t^2$  holds, where  $\bar{v}$  denotes the particles’ mean velocity. In contrast, at times much longer than  $\tau_{\text{crit}}$  (i.e.,  $\langle r^2 \rangle \gg \xi^2$ ), the particle is subject to a multitude of encounters with other entities. Given that the concentration of colloidal particles is high enough to cause caging effects, in the long-time limit of translational diffusion the tracer particle escapes the macromolecular/colloidal cage multiple times. Both of the latter scenarios promote a hindrance of free diffusion and render long-time diffusion slower than short-time diffusion. However, as long as  $\Delta\tau_{\text{tw}} \ll \tau_{\text{crit}}$  or  $\Delta\tau_{\text{tw}} \gg \tau_{\text{crit}}$  the diffusion coefficient stays constant within the corresponding time regime. In these regimes, translational displacements in fact behave truly diffusively ( $\langle r^2 \rangle \propto t$ ) [Doliwa and Heuer, 1998; Weeks and Weitz, 2002a,b].

In-between these regimes, where  $\langle r^2 \rangle \sim \xi^2$ , the diffusion coefficient is time-dependent,  $D(t) \propto t^{\alpha-1}$ , accompanied with displacements  $\langle r^2 \rangle \propto t^\alpha$ . Such behavior is referred to as anomalous diffusion ( $\alpha < 1$ : sub-diffusion,  $\alpha > 1$ : super-diffusion). In agreement with a diffusion coefficient that decreases with time, colloidal particles usually show sub-diffusive behavior at time scales at which particle caging comes into play [Doliwa and Heuer, 1998; Weeks and Weitz, 2002a]. Super-diffusion, in turn, is reminiscent of Levy flights, and is usually not observed for protein diffusion in usual suspensions.<sup>2</sup> The presence of anomalous diffusion has been shown for living cells [Saxton, 1994; Banks and Fradin, 2005; Dix and Verkman, 2008]; in particular, this effect is of importance for very large particles such as macromolecular complexes, but is less significant for protein diffusion [Parry et al., 2014]. The issue of anomalous protein diffusion will not be addressed within this thesis;

---

<sup>2</sup>Levy flights are, for instance, reminiscent of the spread of diseases. Given the world-wide human traveling activities, an infectious disease can cross over from one infected, local area to another, presently uninfected one, allowing for local clusters with large spatial “jumps” in-between.

## 2.1. The generalized Stokes-Einstein and Stokes-Einstein-Debye equations

here, diffusion will solely be assessed within the short-time<sup>3</sup> and long-time regime, in which  $\langle r^2 \rangle \propto t$  holds (“normal” diffusion).

From the rheological point of view, *short-time* translational and rotational diffusion coefficients (denoted  $D_t^S$  and  $D_r^S$ , respectively) are discussed to be linked to the viscosity measured in the limit of high shear rates, denoted as  $\eta_\infty(\phi)$ ,

$$D_t^S = \frac{k_B T}{6\pi\eta_\infty(\phi)R_H} \quad , \quad D_r^S = \frac{k_B T}{8\pi\eta_\infty(\phi)R_H^3} \quad . \quad (2.9)$$

In the limit of a long experimental time window, instead, the same equations are considered to apply with regard to the zero-shear viscosity  $\eta_0(\phi)$ ,

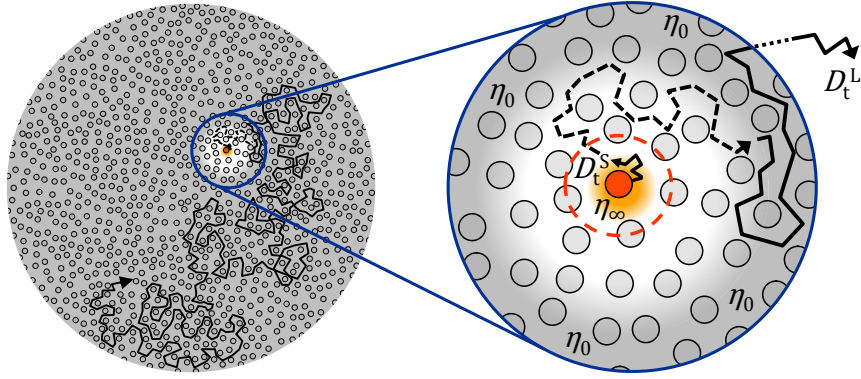
$$D_t^L = \frac{k_B T}{6\pi\eta_0(\phi)R_H} \quad , \quad D_r^L = \frac{k_B T}{8\pi\eta_0(\phi)R_H^3} \quad , \quad (2.10)$$

now defining the *long-time* diffusion coefficients  $D_t^L$  and  $D_r^L$  [Koenderink and Philipse, 2000; Nägele, 2003]. In the above equations,  $\phi$  denotes the volume fraction of dispersed colloidal particles, i.e., their concentration. These relationships are referred to as the *generalized* Stokes-Einstein (GSE) and the *generalized* Stokes-Einstein-Debye (GSED) equations. The high-frequency limit of the viscosity  $\eta_\infty(\phi)$  gives the bulk dissipation of a fast shear oscillation of low amplitude, and is mainly governed by local hydrodynamic interactions and the solvent; direct inter-particle interactions affect  $\eta_\infty(\phi)$  only indirectly via its influence on the equilibrium micro-structure of the solution [Koenderink et al., 2002]. In contrast,  $\eta_0(\phi)$  is determined by slow (i.e. low frequency) steady-shear experiments, for which both hydrodynamic interactions and direct interactions between the colloidal particles are of relevance [Koenderink et al., 2002]. For translational diffusion, the transition from short-time to long-time diffusion is schematically visualized in Fig. 2.1, along with the displacements over which  $\eta_\infty$  and  $\eta_0$  become relevant ( $\langle r^2 \rangle \ll \xi^2$  and  $\langle r^2 \rangle \gg \xi^2$ , respectively; cf. eqs. (2.9) and (2.10)).

In regard to (long-time) *translational* diffusion, the validity of the GSE equation was shown for concentrated hard-sphere systems [Segre et al., 1995; Banchio et al., 1999a,b] and even for soft colloids [Gupta et al., 2015]; however, for charged colloids deviations were reported [Banchio et al., 1999b; Koenderink and Philipse, 2000; Nägele, 2003]. For short-time *rotational* diffusion of colloids Koenderink et al. [2002] showed experimentally that the scaling relation  $D_r^S \propto \eta_\infty(\phi)$  is sufficiently accurate, though not exact, when apparent slip of the host particles is introduced ( $\nu_r = 0.22$  in eq. (2.8)). Other studies, in contrast, report mismatching results as long as the tracer/host size ratio is not large [Koenderink et al., 2001; Zhang and Nägele, 2002; Koenderink et al., 2003]. For rod-like particles (short-time) rotational diffusion was found to couple to  $\eta_0(\phi)$  [Koenderink et al., 2003; Kleshchanok et al., 2012]. Note that a large axis ratio of an ellipsoidal or rod-like particle induces a rotation-translation coupling [Han et al., 2006].

<sup>3</sup>Short-time translational diffusion is accessible by e.g. neutron scattering techniques, and is provided by literature data; see Roosen-Runge et al. [2011].

## 2. Scientific background



**Figure 2.1.:** Schematic representation of an arbitrarily chosen translational trajectory (solid and dashed polygonal line) of a tracer particle (red) dispersed in a solution of like spheres, together with a distinction between  $D_t^S$  and  $D_t^L$ , and the appropriate viscosity  $\eta_\infty$  and  $\eta_0$ , respectively. The tracer particle is caged by surrounding host particles (red dashed circle), but can move almost freely within a very local environment (orange). The “transition zone” (white) between short-time and long-time diffusion is underrepresented here and may involve much more particles with which encounters take place. Long-time diffusion is obtained for displacements much larger than the average cage size, accompanied by a multitude of inter-particle collisions. Then, friction can be expected to scale with  $\eta_0$  (gray area).

### Aspects on protein diffusion

For concentrated protein solutions the applicability of the above equations is not yet justified. Protein-specific interactions and protein complex formation give rise to a potentially sophisticated behavior of crowded protein Brownian motion, whereas colloid-science concepts usually rely on isotropic, inert particles. Within the framework of colloidal-science concepts, the slow-down of protein translational diffusion has been modeled by considering an inhomogeneous environment combined with steric hindrance meaning that regions of sufficiently large vacancies must be present for translational displacements to take place. This simple *ansatz* results in an exponential decrease of protein translational diffusivity with the excluded-volume fraction  $\phi_{\text{ex}}$ ,

$$D_t^L \propto \exp[-\gamma_v \phi_{\text{ex}} / (1 - \phi_{\text{ex}})] \quad , \quad \phi_{\text{ex}} = \rho V_{\text{ex}} \quad , \quad (2.11)$$

where  $\gamma_v$  is a constant relating to the effort of vacancy formation [O’Leary, 1987], and  $\rho$  and  $V_{\text{ex}}$  are the protein solution density and the protein exclusion volume (in liters per gram of protein), respectively. O’Leary observed good agreement of eq. (2.11) with experimental data on hemoglobin up to very high concentrations, hence providing a successful description of protein translational diffusion using hard-sphere modeling. More advanced models from Scaled Particle Theory [Muramatsu and Minton, 1988; Han and Herzfeld, 1993] are in agreement with a (roughly) exponential decrease of the translational diffusion coefficient with increasing particle concentration. As is relevant for the often non-spherical shape of a protein, it was also found that spherical host particles retard



## 2.1. The generalized Stokes-Einstein and Stokes-Einstein-Debye equations

diffusion less than elongated ones of the same volume fraction [Han and Herzfeld, 1993]. Nevertheless, inter-particle interactions beyond pure steric effects may be quite significant for protein diffusion. Given the spread in literature results with regard to applicability [Licinio and Delaye, 1988; Wang et al., 2010] or invalidity [Li et al., 2009; Wang et al., 2010; Zorilla et al., 2007] of the GSE/GSED relationship for different proteins and crowding agents, it seems that strength and nature of inter-particle interactions are indeed critical factors beyond excluded volume. Notably, *in vivo* translational diffusivity of the green fluorescent protein is significantly more retarded than expected from sterical hindrance alone, with a 14-fold decrease of the long-time diffusion coefficient as compared to dilute solution. From that finding, the importance of reversible binding among proteins and other entities has been proposed [Konopka et al., 2006]. As will be shown in the RESULTS section, transient binding among the proteins indeed represents the key factor whether GSE behavior applies or not.

Experiments on protein rotational diffusion, in turn, are often discussed in terms of the previously mentioned micro-viscosity [Luby-Phelps, 2000]. Indeed, for a suspension of only weakly interacting proteins, the probe molecule may rotate rather freely within the cage formed by the surrounding colloids [Długosz and Antosiewicz, 2014]. Then, rotational diffusion would be rather insensitive to concentration effects, and is enhanced relative to *long-time* translational diffusion that is retarded by steric hindrance [Kim et al., 2011]. However, potentially substantial protein-protein interactions may be considered to enhance the rotational viscous drag beyond that of the above scenario. In fact, it has recently been shown that even rather local displacements ( $\tau_{tw} < 10$  ns) can be affected by interactions with the protein's environment [Gupta et al., 2016]. In this study, crowding has been mimicked by synthetic polymer chains. For proteins crowded by proteins, this effect may be even more pronounced, in particular due to charge-charge interactions. Here, the viscous drag may become large enough to even match that of the (macroscopic) zero-shear viscosity  $\eta_0$ . Then, provided that translational diffusion just as well follows the concentration dependence of the macro-viscosity, rotational and long-time translational diffusion would be coupled, and a shared or at least similar concentration dependence of these two quantities would be observed.

The results presented herein suggest that protein rotational diffusion spans the full range between the limiting cases of full coupling and full decoupling from the macro-viscosity, where “full decoupling” relates to hard-sphere behavior. In-between these limiting cases, intermediate behavior has also been found. Comparing among the proteins studied, these data are in line with the SAXS data provided by MARIA OTT reporting on the strength of particle-particle interactions. Taken together, the presented results indicate that anisotropic protein-protein interactions give the key factor for coupling or decoupling of rotational and translational diffusion.

## 2. Scientific background

The protein concentrations studied in this work are not as high as to be close to the colloidal glass transition [Pusey and van Megen, 1987; Pusey, 2008; Hunter and Weeks, 2012]. In the vicinity of the colloidal glass transition, with caging effects being very pronounced, collective motions become essential, and dynamical heterogeneity arises [Doliwa and Heuer, 1998; Kegel and van Blaaderen, 2000] due to locally different structural relaxation times of the dense, almost glassy suspension [Weeks and Weitz, 2002b; Hunter and Weeks, 2012]. Similar conditions hold for undercooled fluids, with the emergence of non-Gaussian displacements [Edmond et al., 2012; Schober and Peng, 2016]. Under such conditions, the behavior of rotational relative to translational diffusion changes, i.e., translational diffusion appears to be enhanced in comparison to rotational diffusion and shear viscosity [Chang and Sillescu, 1997; Edmond et al., 2012]. This phenomenon is known as the “translation-rotation paradox” [Stillinger and Hodgdon, 1994], and directly relates to ensemble effects combined with these dynamical inhomogeneities [Chang and Sillescu, 1997; Sillescu, 1999]. To be noted, the colloidal glass transition of a dense protein suspension should not be mixed up with the “protein glass transition” [Ringe and Petsko, 2003]. The protein glass transition denotes the arrested internal motion of a protein (decreased amplitude and increased time scale of atomic fluctuations) at temperatures of around or below 200 K. It is discussed as being contingent on restricted (“frozen”) mobility of solvent molecules bound to the protein [Vitkup et al., 2000; Paciaroni et al., 2002; Tournier et al., 2003], but is not at all linked to the arrested Brownian motion of a protein at high protein concentrations.

### 2.2. The rotational auto-correlation function in non-dilute protein solutions: deviation from singly exponential behavior

Experiments that address rotational diffusion are often sensitive to the (normalized) auto-correlation function of the time-dependent orientation angle  $\theta$  of the particle,

$$C_r^{(\ell)}(t) = \frac{\langle P_\ell(\cos(\theta(\tau))) \cdot P_\ell(\cos(\theta(\tau+t))) \rangle_\tau}{\langle [P_\ell(\cos(\theta(\tau)))]^2 \rangle_\tau} , \quad (2.12)$$

where  $\langle \cdot \rangle_\tau$  is an average over time  $\tau$ , and  $P_\ell(x)$  is, for most cases, either the first or the second Legendre polynomial,

$$P_1(x) = x \quad , \quad P_2(x) = \frac{1}{2}(3x^2 - 1) \quad . \quad (2.13)$$

The rotational auto-correlation function (RACF) describes the loss of orientational memory with ongoing time, and can, for instance, be accessed via dielectric spectroscopy (DS;  $\ell = 1$ ), polarized fluorescence correlation spectroscopy (FCS,  $\ell = 2$ ), or NMR relaxometry ( $\ell = 2$ ).

## 2.2. Rotational diffusion in non-dilute protein solutions

For isotropic rotational motion in which all orientations are equally populated with time, the RACF decays from  $C_r^{(\ell)}(0) = 1$  to zero,

$$\lim_{t \rightarrow \infty} C_r^{(\ell)}(t) \propto \left\langle P_\ell(\cos(\theta(\tau))) \right\rangle_\tau^2 = 0 \quad ,$$

and equals an exponential decay for isolated particles [Jones, 1988],

$$C_r^{(\ell)}(t) = \exp\left(-|t|/\tau_r^{(\ell)}\right) \quad . \quad (2.14)$$

Here, an infinitely dilute solution and the absence of an orientating external field has been assumed. For non-dilute suspensions, the orientational correlation function can be expected “to show complicated non-exponential behavior” [Jones, 1988] at long times, while at short times an exponential relation may still be valid [Jones, 1988].

In the following, only  $t \geq 0$  will be considered, and the absolute value of  $t$  can be skipped. Eq. (2.14) relates the loss of orientational correlation to the rotational correlation time  $\tau_r^{(\ell)}$ , that in turn relates to the rotational diffusion coefficient via

$$\tau_r^{(\ell)} = \frac{1}{\ell(\ell+1) D_r} \quad . \quad (2.15)$$

As follows from eqs. (2.3) and (2.15), the rotational correlation time  $\tau_r^{(1)}$  (DS) refers to the time required for the molecule to rotate, on average, one radian ( $\approx 57^\circ$ ). Analogously, the root mean-squared angular displacement after a time span of  $\tau_r^{(2)}$  (FCS, NMR) is  $3^{-1/2}$  rad ( $\approx 33^\circ$ ).

If the rotational motion is restricted to a certain range of angles or subject to a preferential orientation (such that not all orientation angles are equally distributed), the rotational auto-correlation function will not decay to zero, but to a finite value,

$$\lim_{t \rightarrow \infty} C_r^{(\ell)}(t) =: S_\ell^2 \leq 1 \quad . \quad (2.16)$$

Such treatment is required when studying the internal (backbone or side-chain) dynamics of a protein. It should be noted that the value of the (generalized) order parameter  $S_\ell^2$  depends on the choice of  $\ell = 1, 2$ .

In this thesis, rotational diffusion will be accessed by NMR relaxometry. Henceforth, the index  $\ell = 2$  will be skipped for ease of notation.

Different particles contributing to the overall rotational auto-correlation function  $C_r(t)$  are accounted for by the sum of the individual correlation functions  $C_{r,k}(t)$ , i.e.

$$C_r(t) = \frac{1}{N} \left( C_{r,1} + C_{r,2} + \dots + C_{r,N} \right) \quad . \quad (2.17)$$

## 2. Scientific background

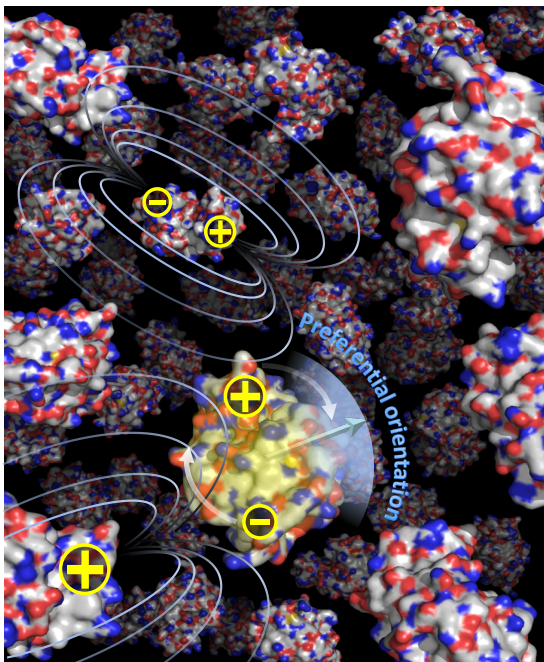
In contrast, for  $M$  superimposed, but statistically and physically independent motions of a single particle, the overall correlation function  $C_r(t)$  consists of a product of the individual correlation functions  $\tilde{C}_i$ ,  $i = \{1, 2, \dots, M\}$ , usually each with its own correlation time  $\tau_i$ :

$$C_r(t) = \tilde{C}_1 \cdot \tilde{C}_2 \cdot \dots \cdot \tilde{C}_M \quad . \quad (2.18)$$

Such treatment is commonly applied for incorporating both internal motions of a protein and the overall Brownian tumbling. However, if these motions are not independent from each other, or are not well separated by different time scales ( $\tau_1 \ll \tau_2 \ll \dots \ll \tau_M$ ), cross-correlations will occur. Correct analysis of motions not decoupled from each other requires advanced models for precise data analysis [Meirovitch et al., 2010], with potentially physically misleading parameters when not accounting for this aspect [Tugarinov et al., 2001; Meirovitch et al., 2006].

Studying non-dilute protein solutions, the RACF has been shown to deviate from singly exponential behavior (that is, eq. (2.14)) beyond the impact of a non-spherical shape of the protein or protein association effects [Krushelnitsky, 2006]. This deviation has been pointed out to be most significant for the long-time decay of the RACF. Such non-ideal behavior suggests an impact of non-negligible protein-protein interactions on protein rotational diffusion and/or the presence of geometrical constraints for protein re-orientational dynamics under crowding conditions.

The deviation of the protein RACF from mono-exponential behavior has been attributed to so-called electrostatic steering effects among neighboring proteins [Krushelnitsky, 2006]. Proteins usually have a non-uniform charge distribution that gives rise to an anisotropic electrostatic field around each protein. In a mean-field approach, the dipole



**Figure 2.2.:** Illustration of protein electrostatic steering for LYZ. Positive and negative charges are plotted here for emphasizing their effect only, but do not rely on the real charge distribution within the protein. Each protein is subject to the electrostatic field arising from surrounding proteins (mean-field approximation), as is here indicated for the slightly yellowish colored protein by emphasizing the dipolar interaction with two close-by LYZ molecules. The figure relies on the same protein constellation as presented in Fig. 1.2.

## 2.2. Rotational diffusion in non-dilute protein solutions

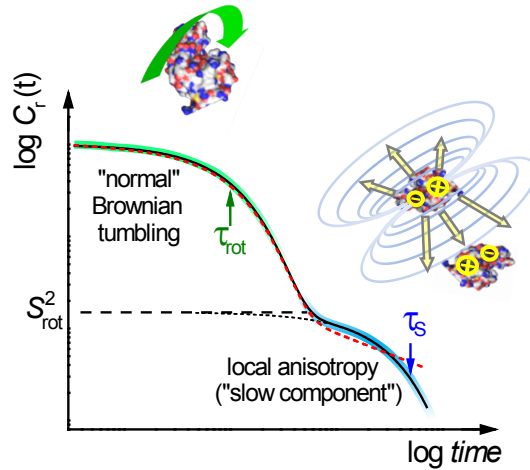
moment of any protein tends to orient along this mean electric field resulting from all the surrounding proteins (cf. Fig. 2.2), which in turn also affects the alignment of the surrounding particles (“feedback effect”). The interaction potential of protein mutual electrostatic steering is comparable to the thermal energy [Ermolina et al., 1993; McClurg and Zukoski, 1998] which is sufficient for introducing local anisotropy (in terms of a preferred orientation of the molecules with respect to each other), but is not sufficient for causing a strict orientational order (as for liquid crystallinity). As has been shown in theoretical physics, mutual alignment effects are a general feature of the solution equilibrium structure of colloids (or proteins) with electrostatic charge patchiness [McClurg and Zukoski, 1998; Grant, 2001; Kurut et al., 2012; Yigit et al., 2015]. As a consequence of electrostatic steering, during the rotational diffusion process of the probe protein the distribution of the orientation angle may stay to some extent non-uniform with time, that is, a slightly preferential orientation remains on average. As a result, the RACF will not decay in a singly exponential manner. Instead, a transient residual ordering may *phenomenologically* be described via an (apparent) rotational order parameter  $S_{\text{rot}}^2$  [Krushelnitsky, 2006]. As discussed below, the denotation of an order parameter is not free of objections here regarding its physical significance; nonetheless, it serves for a straightforward (and minimal) parametrization of the non-exponential nature of the RACF.

With ongoing time the micro-environment is subject to random fluctuations, and the direction of the local electric field changes. Accumulated random fluctuations of the preferred orientation angle finally reintroduce an isotropic average, and  $C_r(t) \rightarrow 0$  ( $t \rightarrow \infty$ ) holds again. Rearrangement of the micro-environment is considered to be primarily mediated by translational diffusion, and requires times much longer than those of Brownian rotation [Ermakova et al., 2002].

The separation of time scales suggests that the overall protein tumbling RACF may be decomposed into two components, accounting for the “usual” Brownian rotation ( $\tau_{\text{rot}}$ ) and a “lifetime” of the transient rotational anisotropy ( $\tau_S$ ), see Fig. 2.3. Both processes are *superimposed* but are intrinsically not statistically independent from each other: the isotropization of each protein’s local environment, causing the RACF to finally decay to zero, relies again on Brownian motion. Since, however,  $\tau_S \gg \tau_{\text{rot}}$ , the long-time re-organization of the environment becomes to a large extent uncorrelated from Brownian rotation on the time scale of  $\tau_{\text{rot}}$ . Despite the potential limitations, following the concept in eq. (2.18), the normalized RACF of non-dilute proteins may tentatively be approximated by the product of the correlation function for Brownian rotation,  $\tilde{C}_{\text{rot}}$ , and that of the isotropization process which is also referred to as the “slow component” of rotational diffusion,  $\tilde{C}_S$ , [Krushelnitsky, 2006]

$$\begin{aligned} C_r = \tilde{C}_{\text{rot}} \tilde{C}_S &= \left[ (1 - S_{\text{rot}}^2) \exp(-t/\tau_{\text{rot}}) + S_{\text{rot}}^2 \right] \exp(-t/\tau_S) \quad (t \geq 0) \\ &\approx (1 - S_{\text{rot}}^2) \exp(-t/\tau_{\text{rot}}) + S_{\text{rot}}^2 \exp(-t/\tau_S) \quad (\tau_{\text{rot}} \ll \tau_S) \quad . \quad (2.19) \end{aligned}$$

## 2. Scientific background



**Figure 2.3.:** Schematic RACF for non-dilute protein solutions. Besides the “usual” Brownian tumbling of correlation time  $\tau_{\text{rot}}$  (colored green), a “slow tail” of the protein rotational auto-correlation function has been observed experimentally (colored blue) that likely arises from so-called electrostatic steering effects [Krushelnitsky, 2006]. This “slow tail” is phenomenologically described by the order parameter  $S_{\text{rot}}^2$  and the correlation time  $\tau_S$  assuming an exponential behavior of the slow component of rotational diffusion (black lines). An exponential-like decay of the latter is a simplistic assumption without further physical justification; the actual shape of the RACF at longer times may differ from such behavior, e.g., via a power law (red dashed line).

This treatment is closely related to the so-called model-free approach invented by Lipari and Szabo [1982a,b] (to be discussed in the next chapter).

Simulating the rotational dynamics of particles having a dipole moment, the concept of the so-called electro-static steering has been confirmed, and a good agreement of the bi-exponential approach (eq. (2.19)) with the simulated RACF has been obtained [Ermakova et al., 2002]. In this simulation, it was also shown that the strength of the dipole moment only weakly affects  $\tau_{\text{rot}}$ , whereas  $\tau_S$  increases with increasing dipole moment. Moreover,  $\tau_S$  was also found to be longer for fixed positions of the particles (but keeping rotational freedom) than in the case of particles subject to random translational displacements. Such finding is in line with the “lifetime”  $\tau_S$  of local ordering effects being mostly dominated by translational motions.

In the limit of very dilute protein solutions, rotational diffusion can be considered to be not affected by surrounding particles, i.e.,  $S_{\text{rot}}^2 \rightarrow 0$ , and only  $\tau_{\text{rot}}$  governs the RACF. Then, the inverse value of  $\tau_{\text{rot}}$  directly relates to the initial slope of the RACF,

$$-\left. \frac{d}{dt} C_r(t) \right|_{t=0} = \frac{1}{\tau_{\text{rot}}} \quad . \quad (2.20)$$

## 2.2. Rotational diffusion in non-dilute protein solutions

In the presence of a slow component of rotational diffusion, the initial slope is

$$\frac{1 - S_{\text{rot}}^2}{\tau_{\text{rot}}} + \frac{S_{\text{rot}}^2}{\tau_{\text{S}}} =: \frac{1}{\tau_{\text{r}}} = 6 D_{\text{r}} \quad (\ell = 2) \quad . \quad (2.21)$$

Eq. (2.21) defines a mean correlation time  $\tau_{\text{r}}$  on the basis of a harmonic average, i.e. it defines a mean diffusion rate  $D_{\text{r}}$ . As usual for time domain data, a mean value is defined via the corresponding frequencies  $\tau_1^{-1}, \dots, \tau_N^{-1}$  rather than averaging over  $\tau_1, \dots, \tau_N$  themselves. In this way, eq. (2.21) provides the *instantaneous* re-orientation diffusion coefficient that is determined by both processes, the usual Brownian tumbling and the slow component of rotational diffusion. Moreover, this effective rotational correlation time  $\tau_{\text{r}}$  (diffusion coefficient  $D_{\text{r}}$ ) is defined in the same way as for other techniques relying on the initial slope of the RACF (as, for instance, polarized FCS).

After all, one should still keep in mind that the above modeling is of a phenomenological nature. Eq. (2.19) rather serves as a parametrization of the non-exponential nature of the RACF than as an advanced model providing a comprehensive physical description. For proteins of different charge distributions, the intrinsic shape of the RACF may differ, and the parameters describing the slow component of rotational diffusion ( $S_{\text{rot}}^2, \tau_{\text{S}}$ ) may not be suitable for comparison among different proteins. In fact, the decomposition into two components, slow and fast, with physically meaningful values of  $S_{\text{rot}}^2$  and  $\tau_{\text{S}}$ , may not always be justified: given that the RACF has no well-developed plateau in-between the fast and the slow component, the actual value of the order parameter is dependent on the fitting model used to mimic the shape of the RACF at longer times. The order parameter may be even physically ill-defined in case that the deviation from singly exponential behavior relates to a power law (see again Fig. 2.3, red line). Having this limitation in mind, the value of  $S_{\text{rot}}^2$  should be interpreted in terms of an apparent amplitude  $a_{\text{S}}$  reflecting the deviation from a singly exponential behavior rather than discussing a physically well-based order parameter. In the same way, the time scale reflecting the long-lasting decay of the RACF,  $\tau_{\text{S}}$ , is an intrinsically model-based parameter. Nevertheless, such parametrization mimics the general shape of the RACF, and thus provides a minimal model for addressing its non-exponential nature in non-dilute protein solutions. Studying the concentration dependence of the “slow tail” of rotational diffusion reflects – within the limitations of the model applied – the trend with which the non-exponential nature of the RACF changes upon crowding. Also note that the initial slope of the RACF as defined in eq. (2.21) is a physically well-defined quantity (the instantaneous rotational diffusion coefficient), and is safely determined during data fitting. As will be shown in the RESULTS section, neglecting the slow tail of rotational diffusion can provide strongly misleading results for non-dilute protein solutions.





# 3 Nuclear Magnetic Resonance spectroscopy and viscosimetry

The study of crowded protein diffusivity presented in this thesis relies on Nuclear Magnetic Resonance (NMR) spectroscopy experiments, with pulsed-field gradient (PFG) NMR for the measure of translational diffusion and NMR relaxometry for studying Brownian tumbling. In the following, an introduction to these techniques will be provided. Key aspects of the viscosity measurements will be presented as well. Experiments and techniques that do not relate to my own work (recombinant protein expression, small-angle X-ray scattering (SAXS) and polarized fluorescence correlation spectroscopy (FCS)) are not discussed here.

### 3.1. Basics of NMR

Several atomic nuclei carry an intrinsic property whose mathematical behavior can be described in terms of an angular momentum called spin. As a quantum-mechanical property, the spin vector  $\mathbf{S}$  is quantized with a total angular momentum of  $\|\mathbf{S}\| = [\mathcal{I}(\mathcal{I} + 1)]^{1/2} \hbar$  and a  $z$ -component given by  $S_z = m_s \hbar$ , where  $m_s \in \{-\mathcal{I}, -\mathcal{I} + 1, \dots, \mathcal{I} - 1, \mathcal{I}\}$ . The nuclear-spin quantum number  $\mathcal{I}$  takes a non-negative integer (even nucleon number, bosons) or a non-negative half-integer (odd number of nucleons, fermions) value. This situation relates to  $(2\mathcal{I} + 1)$  sub-levels that are split by an energy gap of  $\Delta E = -\gamma \hbar B_0$  (Zeeman effect), where  $\gamma$  is the (positive or negative) magnetogyric ratio of the nucleus,  $B_0$  is the strength of an external magnetic field, and  $\hbar$  is the reduced Planck constant. For  $B_0 = 0$  the sub-levels are degenerate. If  $B_0 \neq 0$ , transitions between the sub-states will occur if the energy provided or dissipated by the spin system matches  $\hbar \omega_0 = -\gamma \hbar B_0$ , where  $\omega_0$  is the resonance frequency (referred to as Larmor frequency) with which an electromagnetic wave may be irradiated. Typically, the strength of the magnetic field in NMR spectroscopy is as high as several Tesla,<sup>1</sup> providing a resonance frequency in the radio frequency (rf) range. Most NMR

<sup>1</sup>For comparison, the earth's magnetic field strength is between approximately 25 and 65  $\mu\text{T}$ , with the tendency of a high field strength at the poles and a lower field strength at the equator.

### 3. Experimental methods

spectroscopic experiments on biological systems are carried out on  $^1\text{H}$ ,  $^{13}\text{C}$  or  $^{15}\text{N}$  nuclei, each with a spin of  $\mathcal{I} = 1/2$ .

The nuclear spin gives rise to a magnetic moment

$$\boldsymbol{\mu} = \gamma \mathbf{S} \quad . \quad (3.1)$$

The sum of all magnetic moments is a macroscopic magnetic dipole moment or a macroscopic magnetization  $\mathbf{M} = \frac{1}{V} \sum_i \boldsymbol{\mu}_i$ , where  $V$  is a reference volume. In classical electrodynamics a magnetic moment exposed to a magnetic field  $B_0$  is subject to a torque  $\mathbf{D}$ ,

$$\mathbf{D} = \mathbf{M} \times \mathbf{B}_0 \quad , \quad (3.2)$$

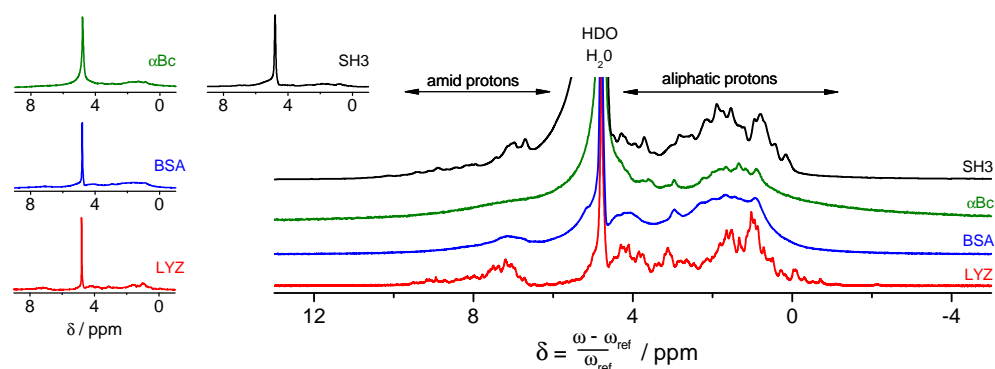
which leads to a precession of the magnetic moment about the direction of the magnetic field,  $\mathbf{B}_0 = B_0 \hat{\mathbf{e}}_z$ . From quantum mechanics it can be easily concluded that the precession frequency equals the Larmor frequency,

$$\omega_0 = -\gamma \mathbf{B}_0 \quad . \quad (3.3)$$

Application of eq. (3.2) requires a spin *ensemble* (limiting case of classical physics); for a single spin, no conclusion on its spin state besides  $S_z = m_s \hbar$  can be drawn. For further comments on a quantum-mechanical basis, see the appendix, section S.1.

A spread of resonance frequencies (shielding effects) originates from the small local fields  $\mathbf{B}_{\text{loc}}$  arising from the electrons surrounding the nucleus which are sensitive to the local chemical environment. For the multitude of spins in chemically equivalent positions (sub-ensembles), this situation is accompanied by individual precession frequencies  $\omega_{0,i} = |\gamma \mathbf{B}_0 + \gamma \mathbf{B}_{\text{loc},i}|$ . The variation of the resonance frequency is referred to as chemical shift  $\delta$  and is quite small, with values in the range of ppm (parts per million) only.

A resonant rf pulse (carrier frequency of  $\omega_0$ ) induces not only transitions among the spin sub-states, but also creates spin coherence. *Macroscopically*, on-resonant rf irradiation serves to rotate the sample magnetization by an angle of  $\phi = \omega_1 t_p$  about the  $x$ - or  $y$ -axis, where  $\omega_1 = |\gamma| B_1$  is the nutation frequency dependent on the strength of the irradiating rf field, denoted as  $B_1$ , and  $t_p$  is the pulse duration. Note that transverse magnetization does not correspond to magnetization in terms of a spin population difference in the lab frame, but results exclusively from a spin coherence. The precession of the spin coherence evolves as an oscillating transverse magnetic field by which a voltage is induced in the NMR coil which serves for signal detection. With time, the coherence is lost, such that the detected signal decays. The detected signal is referred to as Free Induction Decay (FID), the Fourier transform of which gives the individual frequencies the FID is composed of. This results in an NMR spectrum, see Fig.3.1.



**Figure 3.1.:**  $^1\text{H}$  NMR spectrum of 180 g/l  $\alpha\text{B}$ -crystallin ( $\alpha\text{Bc}$ ), as well as 130 g/l BSA, 130 g/l LYZ, and 130 g/l SH3, each dissolved in  $\text{D}_2\text{O}$  ( $T = 22^\circ\text{C}$ ,  $\omega_0 = 400$  MHz). Regions of amide and aliphatic protons are marked. The peak of high intensity (cf. also small graphs) belongs to residual solvent protons (HDO,  $\text{H}_2\text{O}$ ) and was calibrated to 4.8 ppm.

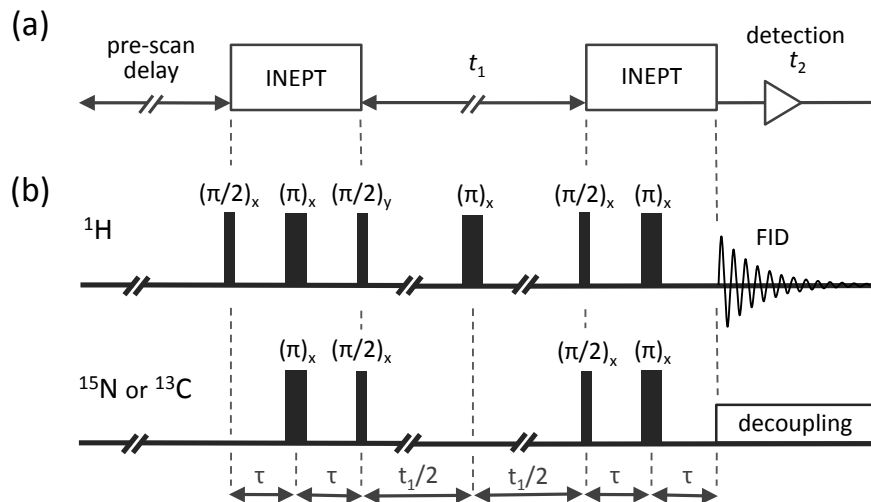
### 3.2. The NMR HSQC experiment

For proteins, the multitude of different proton sites gives many superimposed peaks in an  $^1\text{H}$  NMR spectrum, see again Fig. 3.1. For signal decomposition, protein NMR often relies on multi-dimensional spectra in which the proton chemical shift is plotted against another spin-codable property.

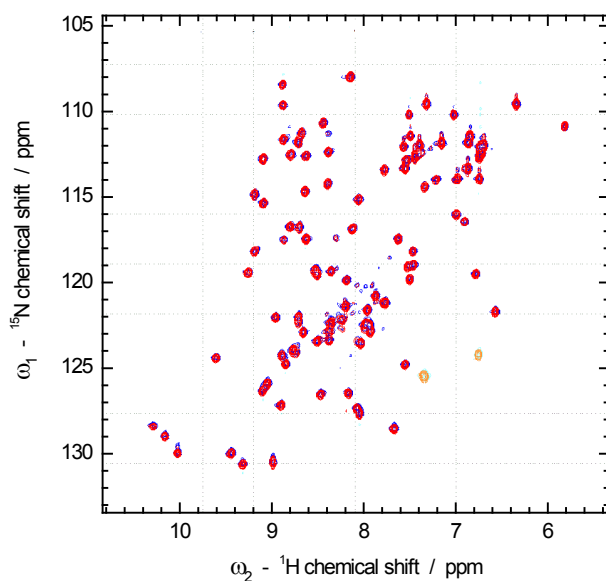
In the results section, the hetero-nuclear single quantum coherence (HSQC) experiment [Bodenhausen and Ruben, 1980] will be used. The basic scheme of an HSQC experiment (see Fig. 3.2a) is as follows: Via INEPT (insensitive nuclei enhanced by polarization transfer; Morris and Freeman [1979]), spin polarization is transferred from the proton spins to the hetero-nucleus, and then evolves during the time delay  $t_1$ . During this period, the signal evolution resembles an FID, and can thus be used to encode chemical-shift information of the hetero-nucleus. Directly after this evolution period, another INEPT is used to transfer the spin polarization back to the proton spins and the resulting proton signal is detected. The signal finally obtained is an FID (evolution time  $t_2$ ) of which the initial (i.e., total) intensity depends on the previous choice of  $t_1$ . By repeating the experiment for different values of  $t_1$  and performing a Fourier transform in both  $t_1$  and  $t_2$ , a two-dimensional spectrum is obtained in which the proton chemical shift is linked (correlated) to that of the close-by hetero-nucleus. In this way, the chemical shift of the hetero-nucleus is detected indirectly. The accordant pulse sequence is shown in Fig. 3.2b.

### 3. Experimental methods

HSQC experiments result in kind of a spectral fingerprint of the protein and can be used to identify protein reaction sites. For this purpose, one usually compares an HSQC spectrum in presence and absence of the protein's ligand. Fig. 3.3 shows an HSQC spectrum of uniformly  $^{15}\text{N}$ -enriched SH3 with and without adding unlabeled BSA. For this particular sample, no significant chemical shift differences were obtained indicating no specific protein-protein binding. Fig. 3.3 directly relates to the results discussed in paper #4.



**Figure 3.2.:** NMR pulse sequence of the HSQC experiment, including its basic scheme (a) and the detailed pulse sequence (b). For complete polarization transfer within the INEPT blocks,  $\tau = 1/(4J)$  must hold, where  $J$  is the scalar coupling strength (in units of Hz).



**Figure 3.3.:**  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectrum ( $\omega_0/2\pi = 800$  MHz) of 3 g/l uniformly  $^{15}\text{N}$ -enriched SH3 in absence (red) and presence (blue) of 19.8 g/l unlabeled BSA, according to a molar ratio of about 1:1 for the two proteins. For both samples, the pH was adjusted to 7.4 using 225 mM sodium phosphate buffer. Comparing the HSQC spectra with each other, no significant changes are obtained. Peaks in orange or bright blue belong to negative intensities resulting from a mirror image of peaks caused by the Fourier transform.

### 3.3. NMR relaxometry

If the spin ensemble is perturbed, spin relaxation will occur. Studying spin relaxation enables the determination of the time scale and amplitude of re-orientational molecular motions that take place in the (sub-)ns to  $\mu$ s regime. Three types of spin relaxation are of importance: (i) longitudinal relaxation, which re-establishes the spin state population in thermal equilibrium (Curie magnetization), (ii) transverse relaxation, which is the loss of spin coherence (loss of “transverse magnetization”), and (iii) spin relaxation in the rotating frame, denoting relaxation of spin-locked coherence during (on- or slightly off-resonant) rf-field irradiation, for which relaxation occurs along the irradiated rf field  $\mathbf{B}_1$ .

The transition of a spin ensemble from the initially distorted to the equilibrium state (intensities  $I_0 = I(0)$  and  $I_\infty = \lim_{t \rightarrow \infty} I(t)$ , respectively) usually follows an exponential behavior,

$$I(t) = I_0 + (I_\infty - I_0) \left( 1 - \exp(-R_x t) \right) \quad , \quad (3.4)$$

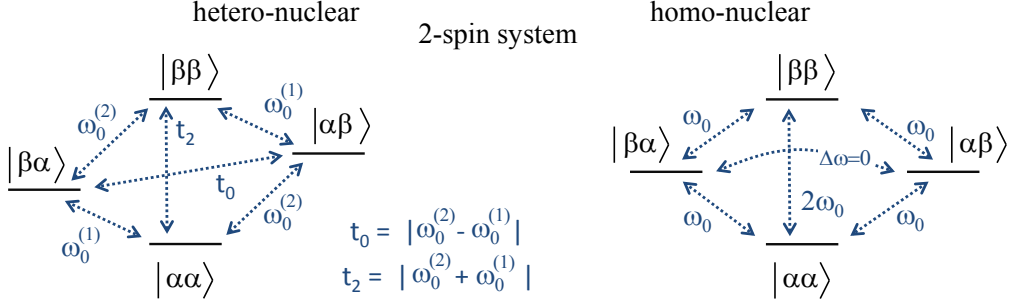
in which  $R_x = 1/T_x$  is the longitudinal ( $R_1$ ), transverse ( $R_2$ ), or rotating-frame ( $R_{1\rho}$ ) relaxation rate ( $T_x$ : time).

The physics of spin relaxation is as follows. The energy levels of a (hetero-nuclear) spin-1/2 pair are split by energy gaps  $\omega_0^{(1)}$ ,  $\omega_0^{(2)}$  and  $\omega_0^{(1)} \pm \omega_0^{(2)}$  (each in units of  $\hbar$ ), see Fig. 3.4. For like spins,  $\omega_0^{(1)}$  and  $\omega_0^{(2)}$  are equal, such that the energy gaps for single and double quantum transitions are  $\omega_0$  and  $2\omega_0$ , respectively. For zero-quantum transitions, no energy gap occurs in the homo-nuclear case. If the energy levels are modulated with a frequency that matches these resonance frequencies, spin transitions will be initiated, with the transition probability per unit of time being calculated by Fermi’s golden rule of quantum mechanics. With these spin transitions, the spin ensemble will re-establish its thermal equilibrium. Note that unlike longitudinal relaxation, transverse relaxation can take place even without changes in the spin population along the z-direction: zero-quantum transitions in the homo-nuclear case do not alter the spin population of the considered spin species (e.g.  $^1\text{H}$ ), yet the energy-conserving, simultaneous mutual exchange of spin polarization (“flip-flop”) is accompanied by a loss of phase coherence. Stochastic fluctuations in the z-component of  $\mathbf{B}_{\text{loc}}$  also cause an intrinsic, accumulative loss of spin coherence, where the effects of chemical exchange are discussed in the experimental section on transverse relaxation.

For (protein) protons, fluctuating local magnetic fields and the modulation of the energy levels mainly originate from dipolar couplings,<sup>2</sup> the angular dependence of which is described by spherical harmonics of second order,  $Y_{2m}$ , where  $m = \{-2, -1, 0, 1, 2\}$ . Spin flip-flops are likewise mediated by dipolar couplings. Thermal motions initiate stochastic fluctuations in the dipolar coupling strength that are described via the auto-correlation

<sup>2</sup>To be mentioned for completeness, for spins with  $I > 1/2$  electric quadrupole couplings are relevant.

### 3. Experimental methods



**Figure 3.4.:** Energy levels for a two-spin system (each spin with  $\mathcal{I} = 1/2$ ), where  $|\alpha\rangle = |\frac{1}{2}, \frac{1}{2}\rangle$  and  $|\beta\rangle = |\frac{1}{2}, -\frac{1}{2}\rangle$  denote the spin-up and spin-down state, respectively. Energies are given in units of  $\hbar$  and should not be mixed up with transition frequencies.

function of these spherical harmonics,  $C_{2m}(t) \propto \langle Y_{2m}(\tau) \cdot Y_{2m}(\tau+t) \rangle_\tau \cong A_m \exp(-t/\tau_{r,m}^{(2)})$ . The Fourier transform of the auto-correlation functions gives the frequencies the modulation of dipolar couplings relates to, and is referred to as the spectral density of motion  $J(\omega)$ . To be noted, for isotropic samples the correlation functions  $C_{2m}(t)$  are identically and can be expressed in terms of the 2<sup>nd</sup> Legendre polynomial,  $Y_{20}(\theta) = P_2(\cos \theta)$  [Lipari and Szabo, 1982a].

Studying proton relaxation in proteins with a natural abundance of  $^{13}\text{C}$  and  $^{15}\text{N}$ , homo-nuclear couplings are far more dominant over hetero-nuclear couplings. Chemical shift anisotropy gives another source for spin relaxation; although, for proton spins, this effect is usually negligible as compared to relaxation mediated by dipole-dipole couplings. Homo-nuclear dipolar couplings thus dominate protein  $^1\text{H}$  relaxation rates, where

$$R_1 = \frac{2}{3} K_{\text{HH}} \left( J(\omega_0) + 4J(2\omega_0) \right) , \quad (3.5)$$

$$R_2 = \frac{1}{3} K_{\text{HH}} \left( 3J(0) + 5J(\omega_0) + 2J(2\omega_0) \right) , \quad (3.6)$$

$$R_{1\rho}^{\text{on}} = \frac{1}{3} K_{\text{HH}} \left( 3J(2\omega_1) + 5J(\omega_0) + 2J(2\omega_0) \right) . \quad (3.7)$$

Here,  $K_{\text{HH}}$  relates to the effective spin-spin coupling strength. As mentioned above, for isotropic media  $J_{m=0}(0)$ ,  $J_{m=1}(\omega_0)$  and  $J_{m=2}(2\omega_0)$  are identical except of the frequency probed, as shown in the above equations. Note that locally anisotropic motions will still provide this identity *if* averaging over different local settings reintroduces an isotropic angular distribution [Kurbanov et al., 2011], as, for instance, in the case of powder averaging in polycrystalline materials. For macroscopically anisotropic media, however,  $J_{m=0}(\omega)$ ,  $J_{m=1}(\omega)$  and  $J_{m=2}(\omega)$  can be structurally different since  $C_{2m}(t)$  may yield different results for different values of  $m = \{0, \pm 1, \pm 2\}$ . A Markov process can still be expected to provide an exponential decay of  $C_{2m}(t)$ , yet the correlation time (and order parameters) may be dependent on the value of  $m$ .

For completeness, it should be mentioned that the equations for spin relaxation due to hetero-nuclear dipolar couplings are basically of the same nature, but with sensitivity to those frequencies (and amplitudes) that relate to the hetero-nuclear case (cf. again Fig. 3.4).

Note that eq. (3.7) holds for on-resonant  $R_{1\rho}$  relaxation experiments, i.e., an on-resonant spin-lock pulse. For off-resonance, see Jones [1966] or, for a more compact expression, Krushelnitsky and Reichert [2004]. The equation for  $R_{1\rho}^{\text{on}}$  resembles that for  $R_2$ , except that the spectral density is not probed in its zero-frequency limit but at twice the nutation frequency of the irradiated rf field. With  $\omega_1 = |\gamma|B_1$  being in the range of some tens of kHz (for reasoning see the section on  $R_{1\rho}$  experiments),  $R_{1\rho}$  relaxation serves for probing the low-frequency region of the spectral density.  $R_2$  basically probes  $J(\omega = 0)$ , and  $R_1$  usually covers the MHz regime.

For an exponential auto-correlation function (Debye process), the spectral density is

$$J(\omega) \sim \frac{\tau}{1 + (\omega\tau)^2} \quad , \quad (3.8)$$

with  $\tau$  being the correlation time characteristic of the motion, and  $J(0) = \tau$ . Eq. (3.8) gives the simplest form of the spectral density; dynamics that are not well represented by a single mode but involve an (intrinsic) distribution of correlation times usually require a modified form of the spectral density [Beckmann, 1988]. Non-Debye dynamics are reminiscent of supercooled liquids [Blochowicz et al., 2003] and synthetic polymers [Kruk et al., 2012].

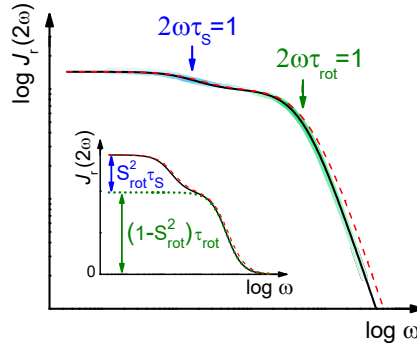
In a mono-disperse solution of a globular protein at a concentration far below the colloidal glass transition, all particles feature the same properties, with the same rotational correlation time for all molecules. Such a scenario is well represented by a simple Lorentzian, i.e., eq. (3.8). If there are distinct proteins, or given the case of protein association, each sub-species is represented again by a Lorentzian, and rotational dynamics of the overall dispersion are well described by a weighted sum of these Lorentzians.

Non-negligible protein-protein interactions gives rise to an intrinsically non-exponential shape of the RACF, and are of particular relevance under crowding conditions. Relying on the bi-exponential approach in eq. (2.19), the spectral density of protein rotational motion under crowding conditions reads ( $S_{\text{rot}}^2 \cong a_s$ )

$$J_r(\omega) = \left(1 - S_{\text{rot}}^2\right) \frac{\tau_{\text{rot}}}{1 + (\omega\tau_{\text{rot}})^2} + S_{\text{rot}}^2 \frac{\tau_s}{1 + (\omega\tau_s)^2} \quad . \quad (3.9)$$

Fig. 3.5 displays the spectral density in eq. (3.9) for the case of  $S_{\text{rot}}^2 = 0$  and  $S_{\text{rot}}^2 \neq 0$ , and compares  $J(2\omega)$  to the  $R_1$  characteristics (eq. (3.5)). As reflected in this graph, the frequency-dependence of  $R_1$  is a well-suited tool to directly access the spectral density and the RACF. On a logarithmic scale, the inflection point of the Lorentzian spectral density (eq. (3.8)) occurs at the condition of  $\omega\tau \sim 1$ . This inflection point serves as a safe estimate of the correlation time  $\tau_{\text{rot},s}$ , provided of course that  $\omega$  is varied within a suitable frequency range.

### 3. Experimental methods



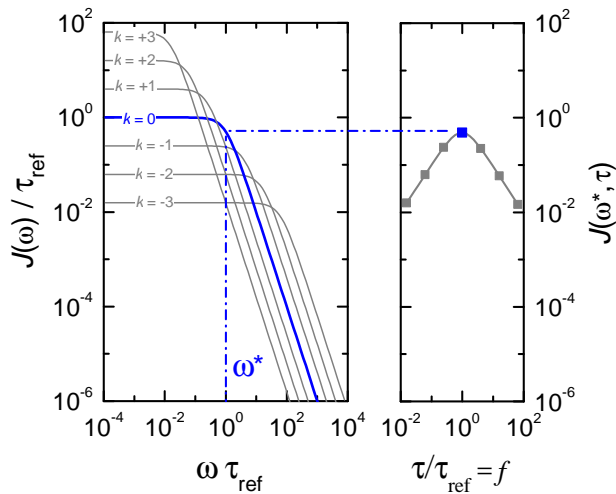
**Figure 3.5.:** Spectral density of protein rotational diffusion. The here shown data set directly results from the Fourier transform of the RACF shown in Fig. 2.3. Longitudinal relaxation rates  $R_1$  have a similar frequency dependence as  $J_r(2\omega)$ , as is seen from also plotting  $(1J_r(\omega) + 4J_r(2\omega))/5$  (red dashed line). On a logarithmic scale in  $\omega$ , the difference between  $J_r(\omega)$  and  $J_r(2\omega)$  causes only a horizontal shift along the frequency axis. The inset reflects the case of  $S_{\text{rot}}^2 \cong a_S = 0$  (green dotted line) on a linear scale in  $J_r(2\omega)$ , from which the amplitudes of the “fast” and “slow” component of rotational diffusion can be compared to each other.

The temperature dependence of spin relaxation rates gives another way of assessing (rotational) correlation times. The rotational correlation time changes with temperature, e.g., via an Arrhenius behavior,

$$\tau(T) \propto \exp\left(\frac{E_A}{k_B T}\right), \quad (3.10)$$

where  $E_A$  is an activation energy. For a fixed frequency  $\omega = \omega^*$  but a variable value of  $\tau$ ,  $J(\omega^*, \tau)$  has a maximum if  $\omega^* \tau^* \sim 1$ . This effect is illustrated in Fig. 3.6 by comparing the frequency dependence of  $J(\omega, \tau)$  for different values of  $\tau$ . For a fixed frequency  $\omega^*$ ,  $R_1$  relaxation rates will behave in a similar way, with a maximum of  $R_1$  occurring again at  $\omega^* \tau^* \sim 1$ . (More precisely, the maximum occurs at  $\omega^* \tau^* = 0.615$  due to the combined action of  $J(\omega)$  and  $J(2\omega)$ .) The temperature dependence of longitudinal relaxation rates may serve to identify the temperature  $T = T^*$  at which this condition is fulfilled. As  $\omega^*$  is precisely known from the experimental setup, this situation provides a safe estimate of  $\tau$  at  $T = T^*$ . For illustration, Fig. 3.7 displays  $R_1$ ,  $R_{1\rho}$  and  $R_2$  relaxation rates for a variable correlation time  $\tau$  that has been multiplied by the fixed frequency  $\omega = \omega^*$ . Note the similar characteristics of  $R_{1\rho}$  relaxation rates as compared to  $R_1$ . If an Arrhenius behavior of  $\tau$  is valid (eq. (3.10)),  $\ln(\tau)$  will be directly proportional to  $1/T$ . Then, the  $x$ -axes in Fig. 3.7 can be replaced by plotting  $1/T$  on a linear scale without changing the characteristics of the relaxation curves. Fig. 3.7 thus directly reflects the temperature dependence of relaxation rates that are sensitive to an Arrhenius process. Apart from the condition of  $\omega^* \tau^* \sim 1$ , the slope of the relaxation rates directly relates to the activation energy  $E_A$ . Also note that noisy experimental data recorded within a rather narrow temperature range may not permit the distinction between Arrhenius and non-Arrhenius behavior of rotational correlation times.





**Figure 3.6.:** Frequency dependence of the spectral density (left) for different values of  $\tau$ , with  $\tau = f \tau_{\text{ref}}$ ,  $f = 4^k$ ,  $k = \{-3, -2, -1, \dots, +2, +3\}$ . Note that all curves are plotted relative to  $\tau_{\text{ref}}$ . Given a fixed frequency  $\omega = \omega^*$  (vertical blue line),  $J(\omega^*)$  has a maximum if  $\tau$  fulfills the condition  $\omega^* \tau^* \sim 1$ . This condition is fulfilled here for  $\tau_{\text{ref}} = \tau^*$ . The right-handed graph shows the value of  $J(\omega^*, \tau)$  for each of the curves as a function of  $\tau / \tau_{\text{ref}} = f$ .

Most applications of NMR relaxometry aim for site-specific information on the *internal* dynamics of the protein. Initiated by the “model-free approach” by Lipari and Szabo [1982a,b], a large number of NMR relaxometry studies have been published in which protein internal dynamics are modeled on the basis of

$$C_{\text{int},r}(t) = C_{\text{int}}(t) \exp(-t/\tau_{\text{rot}}) \quad , \quad C_{\text{int}}(t) = S_{\text{int}}^2 + (1 - S_{\text{int}}^2) \exp(-t/\tau_{\text{int}}) \quad (3.11)$$

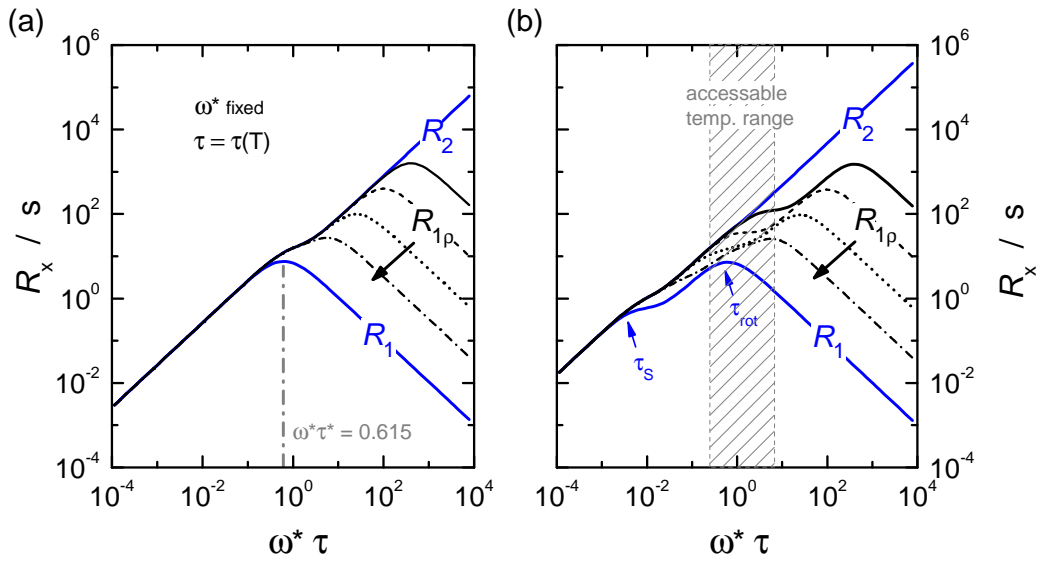
$$\begin{aligned} &\iff \\ J_{\text{int},r}(\omega) &= \left(1 - S_{\text{int}}^2\right) \frac{\tau}{1 + \omega^2 \tau^2} + S_{\text{int}}^2 \frac{\tau_{\text{rot}}}{1 + \omega^2 \tau_{\text{rot}}^2} \quad , \quad (3.12) \\ &1/\tau := 1/\tau_{\text{rot}} + 1/\tau_{\text{int}} \quad . \end{aligned}$$

Here,  $C_{\text{int}}(t)$  denotes the correlation function of internal motions for each resolved position inside the protein, parametrized by a single correlation time  $\tau_{\text{int}}$  and the generalized order parameter  $S_{\text{int}}^2$  that accounts for the spatial restrictions of protein internal dynamics. The model-free approach is “model free” in that knowledge on the detailed physics of motion is not required; nevertheless, the value of the order parameter can be linked to physical constraints: for wobbling within a cone of semiangle  $\theta_c$ , for instance,  $S_{\text{int}}^2$  is determined by [Lipari and Szabo, 1980]

$$S_{\text{int}}^2 = \left( \frac{1}{2} \cos \theta_c (1 + \cos \theta_c) \right)^2 \quad . \quad (3.13)$$

For rigid residues (mostly those of  $\alpha$ -helices or  $\beta$ -sheets),  $S_{\text{int}}^2$  commonly has a value of around 0.8; for rather flexible residues (such as in mobile loops),  $S_{\text{int}}^2$  drops down to values of 0.6 or below. An exemplary data set can already be found in the early work by Clore et al. [1990a], followed by more sophisticated modeling of protein internal dynamics [Clore et al., 1990b]. The model-free approach established as a key methodology for studying protein internal dynamics [Fenwick and Dyson, 2016] with at present more than 100 citations per year of the pioneering work by Lipari and Szabo [1982a] and a total of about 3000 citations by 2016. Objections on the physical value of this approach have been raised by Freed et

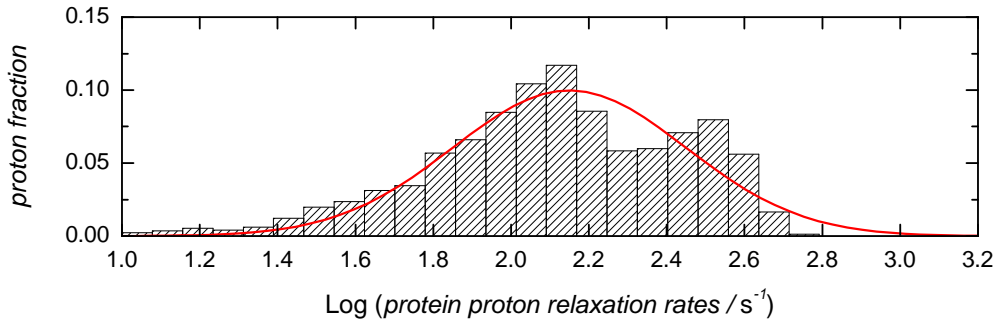
### 3. Experimental methods



**Figure 3.7.:** Dependence of the spin relaxation rates  $R_1$ ,  $R_{1\rho}$  and  $R_2$  on the value of  $\tau$  relative to the fixed frequency  $\omega^*$  for a spectral density composed of a single Lorentzian (a), and for the spectral density of the bi-modal approach in eq. (3.9) (b). Dependence on the value of  $K_{HH}$  (cf. eqs. (3.5)-(3.7)) the curves may be vertically shifted. For  $R_{1\rho}$ , the spin-lock frequency  $\omega_1$  (cf. eq. (3.7)) is in the kHz regime (black solid line).  $R_{1\rho}$  curves approaching  $R_1$  are plotted for illustration only (dashed/dotted lines) and rely on high values of  $\omega_1$  that cannot be achieved by real instrumentation. The above figures also reflect the temperature dependence of spin relaxation curves, see text. As an exemplary illustration a temperature range has been marked in (b) that *may* be accessible without harm for the protein. For this marked temperature range, longitudinal relaxation rates  $R_1$  would be predominantly sensitive to  $\tau_{rot}$  (right-handed maximum), whereas  $R_{1\rho}$  resolves slow dynamics (left-handed shoulder of the  $R_{1(\rho)}$  curves).

al.: given the rather complex nature of protein internal dynamics, the generalized order parameter may not always be physically representative [Tugarinov et al., 2001; Meirovitch et al., 2006].

In all these models for investigating protein internal dynamics, knowledge on the protein's global tumbling time is a prerequisite for estimating internal motions. For the condition of  $\tau_{int} \lesssim 100$  ps and  $\tau_{rot} \gtrsim 1$  ns, with  $R_2$  being not increased by chemical exchange effects (see below),  $\tau_{int}$  and  $S_{int}^2$  have only a minor contribution to the value of  $R_1/R_2$  [Clore et al., 1990a]. Thus,  $\tau_{rot}$  is usually estimated from fitting an appropriate model to the  $R_1/R_2$  ratio of the most rigid residues. Additional benefit comes from the fact that  $R_1/R_2$  is essentially independent from the *a priori* unknown value of  $K_{HH}$ : the coupling constant  $K_{HH}$  can be easily calculated in absence of any internal dynamics, yet its actual value can in fact be reduced by internal motions faster than those resolved by the experimental time window (consider, for instance, fast chemical bond vibrations).



**Figure 3.8.:** Intrinsic distribution of longitudinal relaxation rates in LYZ (bar diagram) as approximated by a log-normal distribution (red line) when fitting the magnetization build-up (or decay) curves as shown in Fig. 3.9. Experimental data reproduced from [Luchinat and Parigi \[2007\]](#), where LYZ was dissolved in D<sub>2</sub>O at a concentration of 2.8 mM (~ 40 g/l) at pH 3.5.

The study of protein internal dynamics requires a high magnetic field strength in order to achieve spectral resolution. However, estimating the global rotational correlation time  $\tau_{\text{rot}}$  from high-field NMR data is merely a compromise for avoiding additional measurements at low magnetic fields. At low magnetic fields, we have  $\omega_0 \tau_{\text{rot},S} \sim 1$  which gives a significantly increased sensitivity to the value of  $\tau_{\text{rot},S}$ . This increased sensitivity of  $R_1$  to global protein dynamics is achieved at the cost of a low signal intensity and a lack of spectral resolution. Due to a lack of spectral resolution, the sample's integral (proton) signal must be evaluated. Different protein proton sites have different dipolar couplings and are subject to different internal dynamics, which is accompanied by a spread of relaxation times. Consequently, the *integral* protein proton relaxation curve is intrinsically multi-exponential,

$$I(t) = I_0 + (I_\infty - I_0) \left( 1 - \sum_{i=1}^N p_i \exp(-R_x^{(i)} t) \right) \quad , \quad \sum_{i=1}^N p_i = 1 \quad . \quad (3.14)$$

Here,  $R_x^{(i)}$  denotes the  $R_1$ ,  $R_{1\rho}$  or  $R_2$  relaxation rate of the  $i^{\text{th}}$  protein proton site. For fitting the NMR relaxometry curves, the distribution  $p_i$  of protein proton relaxation rates may be roughly approximated by a log-normal distribution, as demonstrated in Fig. 3.8 on the example of  $R_1$  relaxation rates in lysozyme [[Luchinat and Parigi, 2007](#)]. Note that the intrinsically multi-exponential nature of the spin relaxation curve should not be mixed up with the deviation of the protein RACF from singly exponential behavior.

As follows from eq. (3.14), the initial slope of the normalized relaxation curve reports on the arithmetic average of relaxation rates,  $R_x^{(\text{av})}$ :

$$\left. \frac{\frac{d}{dt} I(t)}{I_\infty - I_0} \right|_{t=0} = \sum_{i=1}^N p_i R_x^{(i)} =: R_x^{(\text{av})} \quad . \quad (3.15)$$

### 3. Experimental methods

This arithmetic average relates to the mean spectral density of motion and will be used for data analysis. For more details on this topic, see the Supporting Information, section S.2.

For evaluating a shared data set composed of  $R_1$ ,  $R_2$  and  $R_{1\rho}$  measurements, it is required that the same protons contribute to each experiment. Otherwise,  $R_1$ ,  $R_2$  and  $R_{1\rho}$  rates are not directly comparable to each other. Fortunately, this aspect is automatically fulfilled by relying on the integral intensity of *all* the protein protons.

In the following, a basic description of the *technical* aspects of measuring spin relaxation will be provided. Special emphasis on multi-component relaxometry data analysis will be part of the RESULTS section, in particular article #2, and will not be discussed in this chapter.

#### (i) *Measuring longitudinal relaxation:*

##### *Variable field strength via field-cycling NMR*

In thermal equilibrium, the spin population follows a Boltzmann distribution. Hence, for spin-1/2 nuclei, the equilibrium population number of spin-down ( $N_\beta$ , upper energy level) relative to spin-up ( $N_\alpha$ , lower energy level) reads

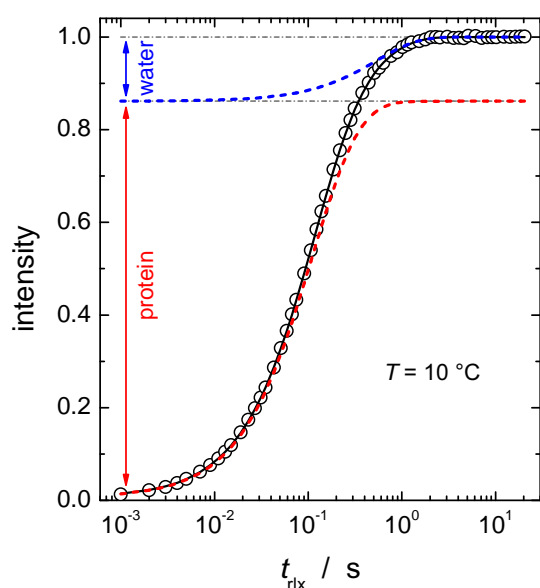
$$\left. \frac{N_\beta}{N_\alpha} \right|_{\text{eq.}} = \exp(-\hbar\omega_0/k_B T) \quad . \quad (3.16)$$

Apart from extraordinarily low temperatures, both states are almost equally populated since  $\hbar\omega_0 \ll k_B T$ ; the excess of spins in the lower energetic state is only in the range of parts per million (ppm). Only the difference in the population states is accessible by NMR experiments, which is the reason of the intrinsically low sensitivity of NMR spectroscopy.

To measure longitudinal relaxation, the equilibrium population of spin states needs to be perturbed. There are two ways of doing so: (i) use of pulses (commonly, the saturation/inversion recovery experiment) or (ii) switching the magnetic field (field-cycling NMR).

In most applications, longitudinal relaxation rates are accessed by saturating the spin population, i.e, the same number of spin-up and spin-down states, with no sample magnetization occurring. Here, the intensity of the detected NMR signal is evaluated as a function of the evolution period  $t_{\text{rlx}}$ , cf. Fig. 3.9. Saturation of the spin population can be achieved by a (train of  $N \geq 1$ )  $(\pi/2)$ -pulse(s). Consequently, the accordant basic pulse sequence simply reads  $[\tau - (\pi/2)]_N - t_{\text{rlx}} - (\pi/2) - \{\text{detection}\}$ , and is referred to as the saturation recovery experiment.<sup>3</sup>

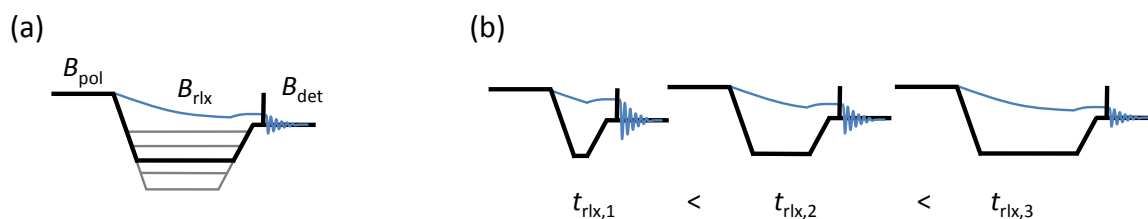
<sup>3</sup>For the inversion recovery experiment, the initial  $(\pi/2)$ -pulse is replaced by a  $(\pi)$ -pulse, or, for compensating pulse imperfections, by a so-called composite pulse,  $(\pi/2)_x - (\pi)_y - (\pi/2)_x$ .



**Figure 3.9.:** Exemplary data set (open symbols) for longitudinal relaxation, measured at 20 MHz Larmor frequency via a saturation-recovery experiment on BSA (213 g/l,  $T = 10^\circ\text{C}$ ), using 320 scans per data point. The black line shows the fitted overall magnetization recovery, composed of the protein (red dashed line) and the water (blue dashed line) signal (see chapter 4, paper #2 for details). For presentation only, the fitted water signal was shifted upwards towards higher intensities. The equilibrium intensity of protein and residual water protons is indicated by vertical arrows and the horizontal dashed-dotted lines, respectively.

At low magnetic fields for accessing longer time scales ( $\omega_0$  is small) the NMR signal is too weak for acquisition using a reasonable number of scans (i.e., relying on a limited measurement period). For this purpose, but also for having an adjustable value of the Larmor frequency  $\omega_0$ , field-cycling (FC) NMR has been developed in which the equilibrium (Curie) magnetization changes during the measurement by application of quick adjustments of the magnetic field; see Fig. 3.10a. Here, an increased sample magnetization is created by incorporating a magnetic field stronger than that one at which relaxation is later on studied. The basic concept is as follows: after the sample was given enough time to be in equilibrium, the polarization field (e.g.  $\omega_0^{(\text{pol})}/2\pi = 20$  MHz) is ramped down within milliseconds towards the relaxation field (adjustable field strength,  $\omega_0^{(\text{rlx})} < \omega_0^{(\text{pol})}$ , e.g.  $\omega_0^{(\text{rlx})} = 10$  MHz). Re-establishment of the new equilibrium magnetization is then studied by variation of the time span during which the relaxation field (and hence longitudinal relaxation of the sample) is active; see Fig. 3.10b. After the relaxation delay, and for signal acquisition only, the magnetic field is ramped up again towards the detection field (here,  $\omega_0^{(\text{det})}/2\pi$  was fixed to 16 and 19 MHz dependent on the setup and experimental conditions). Then, a  $(\pi/2)$ -pulse is applied to now record the FID. The (initial) intensity of the recorded FID reflects the sample magnetization remaining after the adjustable relaxation period. In usual applications, the magnetization ramps are short compared to the spin relaxation time  $T_1$ . If the spin relaxation time is comparable to or even faster than the field switching time, the data analysis of a multi-exponential decay will occur to be problematic: protein proton sites of fast longitudinal relaxation may not contribute to the relaxation decay, and a biased signal will be obtained.

### 3. Experimental methods



**Figure 3.10.:** The principle of FC NMR. By adjusting the relaxation field  $B_{rlx}$ , relaxation at different Larmor frequencies can be accessed (a). The initial intensity of the signal is determined by  $B_{pol}$ ; detection is performed at  $B_{pol}$  after a  $(\pi/2)$ -pulse. By variation of the duration of the relaxation field  $t_{rlx}$ , the spin relaxation curve is obtained (b). In comparison to the time scale of spin relaxation, the duration of the magnetic-field ramps is exaggerated here, at least with respect to usual applications.

Field-cycling NMR can also be applied in the way that the relaxation field is higher than the polarization field. Then, a magnetization build-up curve is detected (very similar, but not totally equal to a saturation-recovery experiment; cf. Fig. 3.9). The larger the difference between the polarization and the relaxation field, the better spin relaxation can be resolved.

FC instrumentation does not rely not on superconductive magnets, but on electromagnets combined with strong, quickly switchable electric currents, with the equipment designed in such a way that only very short electrical ring-down delays are required. For an overview on FC NMR instrumentation and applications using switchable coil magnets, see the reviews by Kimmich [1980], Kimmich and Anardo [2004] and Fujara et al. [2014]. Limited strength and homogeneity of the detection field of these magnets do not provide sufficient spectral resolution; thus, one *has to* rely on the *integral* proton signal. Within the last years, however, shuttling devices have been developed that combine high-field NMR magnets with relaxation measurements at low magnetic fields [Redfield, 2003, 2012]. Here, the sample is moved away from and back again to the center of the high magnetic field. A drawback of these sample shuttling techniques is that increased field switching times must be conceded (currently, approx. 300 ms), which is contrary to the fast spin relaxation times of proteins under crowding conditions ( $R_1 \sim 1 - 10$  ms (protons), depending on temperature and protein concentration; see paper #2 of the RESULTS section). Moreover, quickly moving and stopping the sample can distort the behavior of the “slow component” of rotational diffusion, which is considered to be related to the re-alignment of the local environment of each protein [Krushelnitsky, 2006]. Fortunately, for Brownian dynamics, discriminating among different protein protons is not required.

Here the presented FC NMR experiments were performed in cooperation with the research group of ERNST RÖSSLER<sup>†</sup> in Bayreuth, Germany. After having clarified all relevant measurement parameters and after having completed a first measurement series, the experiments were continued by MARIUS HOFMANN<sup>†</sup>, who is greatly acknowledged for his

<sup>†</sup>University of Bayreuth, Experimental Physics II

effort. Performing the FC NMR experiments for a large number of relaxation fields and temperatures, the overall measuring period took several months. The accumulated data were analyzed several times for identifying the most suitable way of data analysis.

(ii) *Measuring transverse relaxation:*

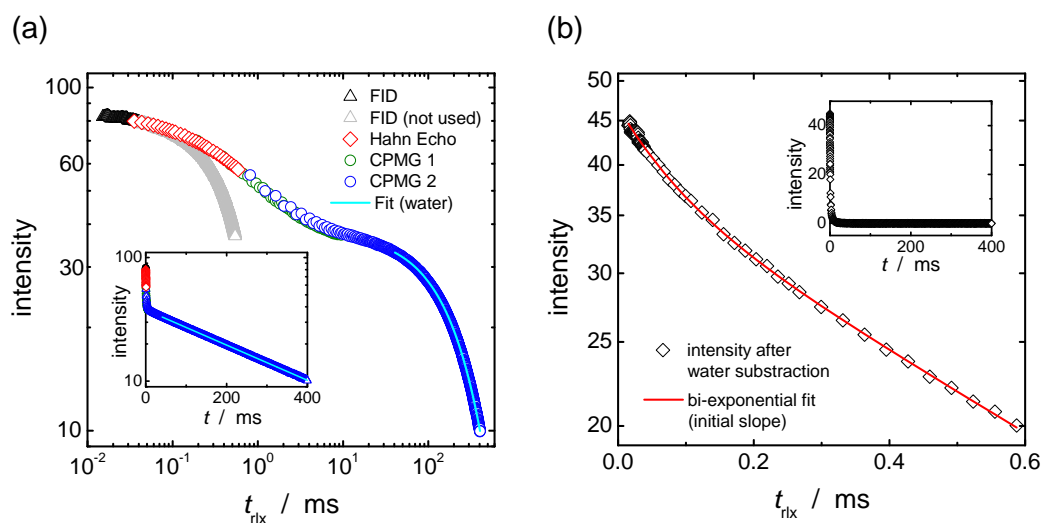
*Combined action of FID, Hahn echo, and the CPMG experiment*

Transverse relaxation, that is the loss of spin coherence created after an NMR pulse, originates from two sources: inhomogeneity of the outer magnetic field  $\mathbf{B}_0$ , providing a spread of Larmor frequencies, and intrinsic transverse relaxation emerging from spin-spin interactions. For estimating rotational diffusion rates, only intrinsic relaxation is of relevance. In contrast, the magnetic field inhomogeneity depends on instrumentation and is generally kept as low as possible.<sup>4</sup> The impact of residual field inhomogeneity can be easily circumvented by the use of  $(\pi)$ -pulses applied to the “transverse magnetization” in the middle of the evolution period  $t_{rlx}$  (the so-called Hahn-Echo,  $\frac{1}{2}t_{rlx} - (\pi) - \frac{1}{2}t_{rlx}$ ); then, only intrinsic spin-spin relaxation remains. The intrinsic transverse relaxation of the protein occurs on a time scale much faster than the intrinsic transverse relaxation of the solvent (water) molecules. To overcome the different time scales involved, but in particular for precisely resolving the initial slope of the protein relaxation decay, a set of different experiments were combined: (a) the usual FID the initial part of which is unaffected by field inhomogeneity effects, providing the very beginning of the  $R_2$  decay; (b) a Hahn echo for accessing the protein-dominated loss of spin coherence; and (c) the Carr-Purcell-Meiboom-Gill (CPMG) pulse sequence for an effective measure of the water component. Unlike the Hahn echo the CPMG experiment consists of a train of  $(\pi)$ -pulses combined with stroboscopic detection of the NMR signal (i.e., short detection of the integral signal intensity in-between the pulses). In this way, many points can be detected within a *single* CPMG experiment. The Hahn echo, in contrast, only provides information on a single data point. Varying the time span  $\frac{1}{2}t_{rlx}$  before and after the  $(\pi)$ -pulse, however, the Hahn Echo accesses a time scale shorter than that of the CPMG experiment in which only a multiple of the shortest possible relaxation delay can be resolved. A combined set of these experiments is shown in Fig. 3.11a. On the basis of these experiments, and after subtraction of the water signal, the intrinsically multi-exponential protein relaxation decay can be resolved, see Fig. 3.11b.

All measurements of transverse relaxation were performed at 20 MHz Larmor frequency. Although signal detection at higher magnetic fields requires much shorter measurement times due to an intrinsically better signal-to-noise ratio (remember, in particular, the Boltzmann population of spin states, eq. (3.16)), low-field NMR  $R_2$  measurements still remain beneficial: chemical exchange effects [Luz and Meiboom, 1963] that cause faster transverse relaxation are not of importance here. If the local chemical environment of

<sup>4</sup>Except for Magnetic Resonance Imaging (MRI) and measurements of translational diffusion, which both require static or pulsed magnetic-field gradients.

### 3. Experimental methods

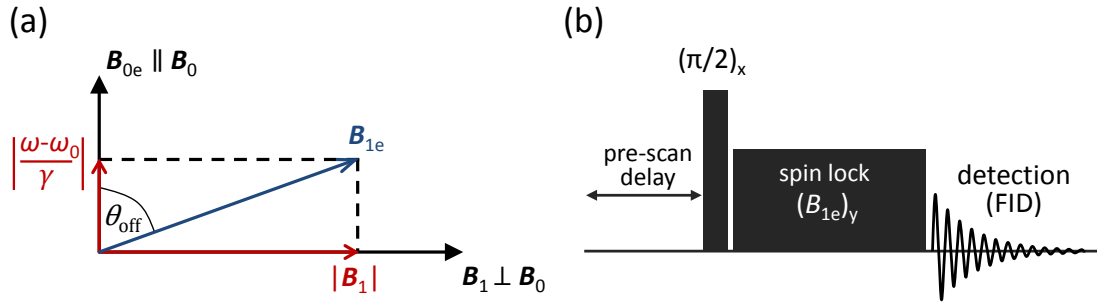


**Figure 3.11.:** Exemplary data set for transverse relaxation (BSA 213 g/l,  $T=20^\circ\text{C}$ ,  $\omega_0 = 20\text{ MHz}$ ), composed of four experiments (a). Note the faster loss of coherence of the simple FID (gray points) as compared to the same decay when residual field inhomogeneity effects are compensated (colored points). The CPMG experiment using a short inter-pulse spacing (green) is only used to bridge the gap in-between the Hahn echo (red) and the second CPMG using a longer inter-pulse delay (blue). The solvent signal is mono-exponential (cf. inset for a semi-logarithmic plot) and can be easily subtracted, such that the  $R_2$  decay of solely the protein protons can be addressed (b). Here, the (initial part of the) protein decay can be reproduced by either a bi-exponential fit or by the use of a log-normal distribution of relaxation times (the latter is even sufficient to reproduce the full protein decay). For achieving a proper signal-to-noise ratio, 128 repetitive scans were used for FID detection and the CPMG experiments, and 64 scans for the Hahn echo.

a spin changes, e.g. via internal dynamics of the protein, the chemical shift of the spin will change as well. Such change of the Larmor frequency can only be compensated if the inter-pulse delay between the ( $\pi$ )-pulses is shorter than the characteristic time scale of this effect; otherwise, re-focussing by the ( $\pi$ )-pulse remains incomplete, and the so-called chemical exchange effect causes a shortening of the  $R_2$  decay that is not related to the Brownian dynamics of the protein. For instance, chemical exchange of water protons with amide protons of  $\sim 100\text{ g/l}$  BSA causes the water transverse relaxation rate to increase by a factor of about two at 100 MHz Larmor frequency [Hills et al., 1989]. To overcome this effect, inter-pulse spacings shorter than  $100\ \mu\text{s}$  are required. However, using such short inter-pulse delays, a large number of ( $\pi$ )-pulses is needed to reach the long times<sup>5</sup> required for detection of the long-lasting water decay, in turn causing sample heating via energy dissipation. Moreover, decreasing the inter-pulse delay down to values comparable to the ( $\pi$ )-pulse duration ( $\sim 10\ \mu\text{s}$ ) one approaches an  $R_{1\rho}$  experiment rather than reliably measuring  $R_2$ . At low field strength, however, chemical shift differences are small, and chemical exchange effects are negligible (for demonstration, see again Hills et al. [1989]).

<sup>5</sup>Up to the range of seconds for high-field NMR, e.g., for  $\sim 100\text{ g/l}$  BSA in  $\text{H}_2\text{O}$  at  $23^\circ\text{C}$  and a Larmor frequency of 100 MHz,  $R_{2,\text{H}_2\text{O}} \approx 2\text{ s}^{-1}$  [Hills et al., 1989].





**Figure 3.12.:** Effective field strength in the rotating frame,  $\mathbf{B}_{1e}$  (a), and the pulse scheme for the  $R_{1\rho}$  measurements (b). For off-resonant  $R_{1\rho}$  measurements, a slightly modified version of the above pulse sequence is used: here the off-resonant spin-lock pulse is flanked by two pulses of the same off-resonance, the first one of which aligns the magnetization vector along the direction of  $\mathbf{B}_{1e}$ , whereas the second turns the magnetization vector back to the direction of  $\mathbf{B}_0$ . This is followed by an on-resonant  $(\pi/2)$ -pulse and signal detection; see also the Supporting information for paper #1. Off-resonant spin-lock pulses are usually described via the off-resonance angle  $\theta = \theta_{\text{off}}$  between  $\mathbf{B}_{1e}$  and the static magnetic field  $\mathbf{B}_0$  (z-direction). For on-resonant rf irradiation,  $\theta_{\text{off}} = 90^\circ$ .

(iii) *Measuring relaxation in the rotating frame:*

*Bridging the gap in-between  $R_2$  and  $R_1$*

$R_{1\rho}$  measurements reflect the spin relaxation during a spin-lock pulse of strength  $B_1$ , with usually  $B_1 \ll B_0$ . During spin-lock the rf field is irradiated with the same phase as the spin coherence. In this way, any coherence that has the same phase as the spin-lock pulse will be constantly refocused and is held along the axis of the rf pulse. Spin relaxation now occurs along  $\mathbf{B}_1$  instead of  $\mathbf{B}_0$ . In order to handle this situation, one usually relies on the so-called “rotating frame” approach: by considering a coordinate frame that rotates with the same frequency  $\omega$  as the spin precesses, the *effective*, rotating-frame  $B_0$  field vanishes,  $B_{0e} = |(\omega - \omega_0)/\gamma| = 0$ . Then, only the irradiated rf field  $\mathbf{B}_{1e} = \mathbf{B}_1$  remains active and gives the field along which spin relaxation occurs. For that reason,  $R_{1\rho}$  relaxation is also referred to as “relaxation in the rotating frame”. Off-resonance effects can be treated in a similar way. Here,  $\omega \neq \omega_0$  goes with the residual precession of the spin ensemble which conforms to a residual  $B_0$  field. Assuming  $\mathbf{B}_1$  points along the  $x$ -direction, the effective overall magnetic field reads  $\mathbf{B}_{1e} = |(\omega - \omega_0)/\gamma| \hat{\mathbf{e}}_z + B_1 \hat{\mathbf{e}}_x = B_{0e} \hat{\mathbf{e}}_z + B_1 \hat{\mathbf{e}}_x$ , where  $\hat{\mathbf{e}}_x$  and  $\hat{\mathbf{e}}_z$  give the unity vectors in  $x$ - and  $z$ -direction, respectively. The angle between  $\mathbf{B}_{1e}$  and  $\mathbf{B}_0$  is referred to as the off-resonance angle  $\theta_{\text{off}}$  and serves to describe the experimental conditions; see also Fig. 3.12a.

$R_{1\rho}$  measurements are well suited for resolving the spectral density in the kHz regime, and bridge the gap between  $R_1$  (usually MHz regime) and  $R_2$  (providing  $J(0)$ ). In fact, relaxation at magnetic fields that are as low as several kHz can also be accessed by modern field-cycling instrumentation. In recent developments sufficiently stable fields

### 3. Experimental methods

for relaxation regimes of several Hz have been achieved [Kresse et al., 2011, 2014]. Such measurements are achieved by large magnetic field jumps between the polarization, the relaxation and the detection field. Given a multi-exponential relaxation decay, a biased estimate of the mean  $R_1$  might be obtained. In particular, biased results were obtained for BSA at lower magnetic fields; this data were not used for further data analysis.  $R_{1\rho}$  measurements were of additional value to verify the field-cycling  $R_1$  data measured at low magnetic fields. Note that in  $R_{1\rho}$  measurements, the change in the applied magnetic fields is almost instantaneous ( $\mu\text{s}$  delays only), with no need of magnetic field ramps (cf. Fig. 3.12b).

For the applications here,  $\omega_1/2\pi$  is in the range of 20–40 kHz: the value of  $\omega_1$  should be larger than the spectral width of the NMR signal ( $\delta \lesssim 25$  ppm, i.e.,  $\lesssim 10$  kHz at 400 MHz  $^1\text{H}$  resonance frequency) to avoid additional off-resonance effects, but should also be low enough to allow for long-lasting  $B_1$  irradiation without risk of damage to the NMR pulse generator. To access higher frequencies, off-resonant  $B_1$  irradiation should be used.

### 3.4. Pulsed-field gradient NMR

Pulsed-field gradient (PFG) NMR enables a precise and straightforward determination of translational self-diffusion coefficients. It evolved from the pioneering work of Stejskal and Tanner [1965] and is based upon the efficiency of nuclear spin echoes (SE) in presence of (temporarily) inhomogeneous magnetic fields [Hahn, 1950; Carr and Purcell, 1954]. In this technique, a position-dependent magnetic field encodes spin positions, the same basis of MRI.

The pulse scheme (cf. Fig. 3.13a,b) of the PFG NMR SE pulse sequence is as follows: having first excited a single-quantum coherence by applying a  $(\pi/2)$ -pulse, the coherence starts to dephase because of intrinsic transverse relaxation, chemical shift effects, and, importantly, due to the pulsed gradient of strength  $\mathbf{g} = \nabla\mathbf{B}_0$  and duration  $\delta$ . The pulsed field gradient causes a position-dependent evolution of the phase angle of the spin coherence according to

$$\Delta\phi_{\text{PFG}}(\mathbf{r}_i) = \omega t \Big|_0^\delta = \gamma \mathbf{g} \cdot \mathbf{r}_i \delta \quad . \quad (3.17)$$

In eq. (3.17) a rectangular gradient pulse (no ramp times; constant amplitude  $|\mathbf{g}| = g$  during time span  $\delta$ ) is assumed. For non-rectangular gradient pulses, however, this relationship still holds, given that gradient strength and duration are described by

$$\int_0^\delta \tilde{g}(t) dt \stackrel{!}{=} g\delta \quad \Leftrightarrow \quad g = \langle \tilde{g}(t) \rangle_\delta \quad . \quad (3.18)$$

After having switched off the gradient pulse, one waits for an overall time delay  $\Delta$  during which translational diffusion occurs. Note that the diffusion delay is defined as the time span from the first gradient pulse to the second, the latter used for decoding the diffusion signal. After half of the diffusion delay, a  $(\pi)$ -pulse is applied. The  $(\pi)$ -pulse causes an inversion (altered sign) of the phase angle. Given that the gradient strength and duration acting on the spin ensemble the second time are the same as they were during the first gradient pulse, a (fully refocused) spin echo will occur. A spin echo means that chemical shift differences and gradient pulses are canceled out; that is, in absence of translational displacements, all spin sub-ensembles  $i = \{1, \dots, N\}$  have the same phase angle,<sup>6</sup>

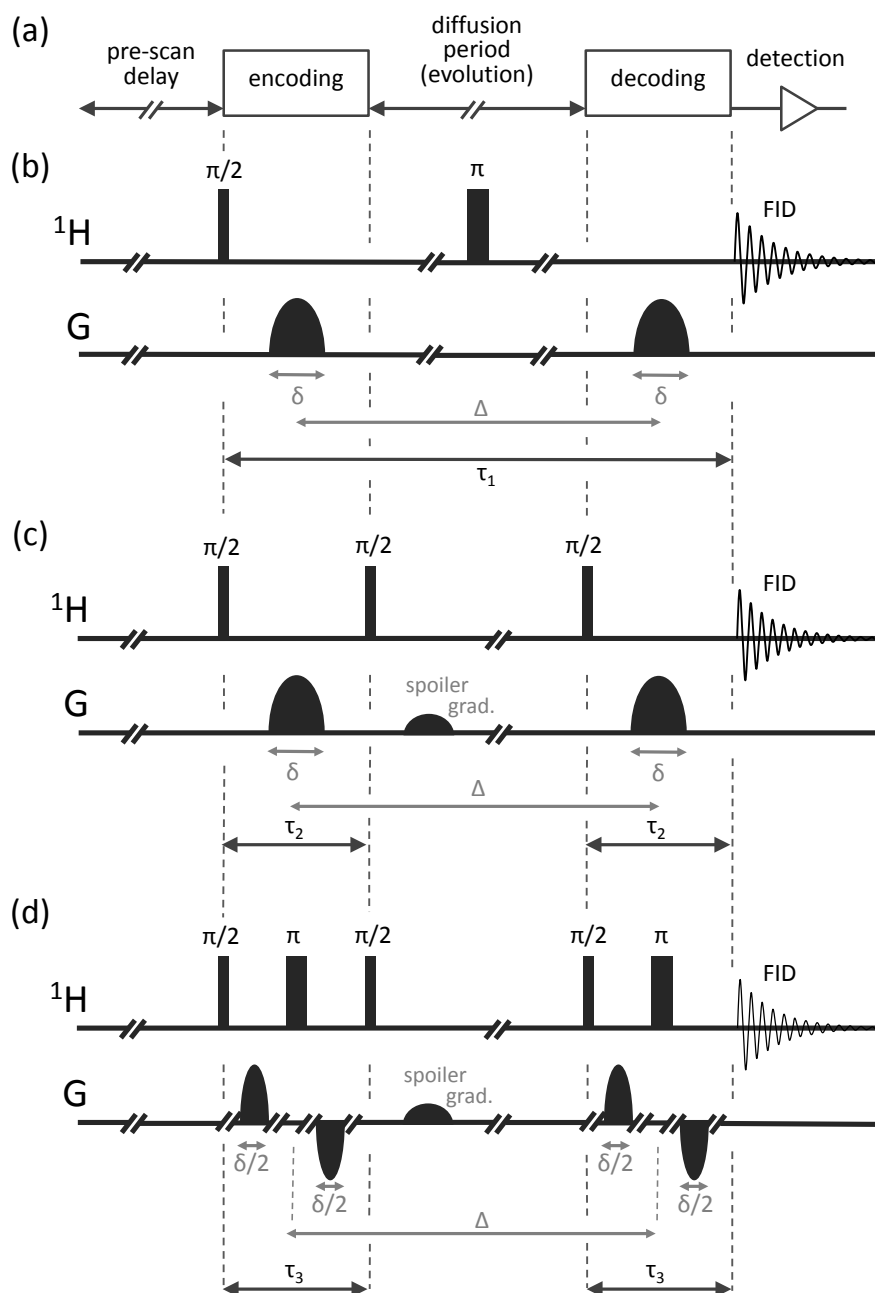
$$\begin{aligned} \Delta\phi_i(\mathbf{r}_i(t), \mathbf{r}_i(t + \Delta)) = \omega t \Big|_{t \rightarrow t + \Delta} &= \left( \gamma \mathbf{g} \cdot \mathbf{r}_i(t) \delta + \omega_{0,i} \frac{\Delta}{2} \right) - \left( \omega_{0,i} \frac{\Delta}{2} + \gamma \mathbf{g} \cdot \mathbf{r}_i(t + \Delta) \delta \right) \\ &= -\gamma \mathbf{g} \cdot \Delta \mathbf{r} \delta \quad , \quad \Delta \mathbf{r} = \mathbf{r}_i(t + \Delta) - \mathbf{r}_i(t) \quad (3.19) \\ &= 0 \quad \text{if} \quad \Delta \mathbf{r} = 0 \quad . \end{aligned}$$

If the position of each spin stays unchanged, the spin echo will have its maximum intensity. If translational diffusion occurs then  $\mathbf{r}_i(t) \neq \mathbf{r}_i(t + \Delta)$  and the accumulated phase angle will be different from spin to spin (in terms of sub-ensembles). As a consequence, the larger the (diffusion-mediated) displacement of the spins the more reduced the measurable spin coherence. As such, the attenuation of the NMR signal with ongoing diffusion provides a direct measurement of the ensemble-averaged mean-square displacement of the probe molecules, in turn relating to the self-diffusion coefficient.

Quantification of translational self-diffusion via the SE pulse sequence works sufficiently well for small molecules, as these are subject to slow transverse relaxation ( $R_2 \approx R_1$ ) only. In contrast, for large molecules such as proteins, the intrinsic relaxation rate  $R_2$  can be quite short, such that the signal finally obtained can be appreciably reduced, if not weakened beyond detection. A minor modification, referred to as the Stimulated Echo (STE) pulse scheme [Tanner, 1970], serves to avoid this issue: instead of relying on a coherent state of the spin ensemble during the diffusion delay  $\Delta$ , the information is “stored” along the  $z$ -axis, i.e., for spin-1/2, the information is stored in terms of spin-down and spin-up states during the diffusion delay, leading to a position-dependent  $\pm z$ -magnetization. In practice, this improvement is achieved by splitting the  $(\pi)$ -pulse into two  $(\pi/2)$ -pulses; see Fig. 3.13c. Stored along the  $z$ -axis, the spin-encoded information is subject to longitudinal relaxation ( $R_1$ ) instead of being sensitive to  $R_2$ . For proteins (and similar macro-molecules)  $1/T_1 = R_1 \ll R_2 = 1/T_2$ , such that the relaxation-mediated signal attenuation is significantly reduced by the STE scheme. A disadvantage of the STE scheme is a 50% reduction in signal due to the fact that the second pulse (cf. Fig. 3.13c) only reconverts half of the spin

<sup>6</sup>Neglecting intrinsic dephasing via spin-spin interactions, i.e., intrinsic  $T_2$  relaxation.

### 3. Experimental methods



**Figure 3.13.:** NMR pulse sequences for measuring translational diffusion. The here shown schemes apply to  $^1\text{H}$ , but can likewise be applied to other spin-carrying nuclei. The basic scheme (a) is the same for the SE pulse sequence (b) as for the monopolar (c) and the bipolar (d) STE experiment. Note that  $\tau_1 > 2\tau_3 > 2\tau_2$ , where  $\tau_1$  denotes the time span during which the spin system is subject to transverse relaxation.

coherence back to a  $\pm z$ -magnetization; the other half is dephased by a spoiler gradient to not interfere with the measuring result.

To avoid the influence of longitudinal spin relaxation on the *shape* of the diffusion decay one usually keeps all time constants fixed, but varies the strength of the pulsed gradient. Stronger gradients increase the sensitivity to translational displacements. With increasing gradient strength  $g = |\mathbf{g}|$  the decay of the signal intensity  $I_{\text{PFG}}$  provides the (long-time) self-diffusion coefficient  $D_t^{(L)}$  via [Stejskal and Tanner, 1965]

$$I_{\text{PFG}}(g) = I_{\text{PFG}}(0) \exp\left(-\gamma^2 g^2 D_t^{(L)} \delta^2 (\Delta - \delta/3)\right) . \quad (3.20)$$

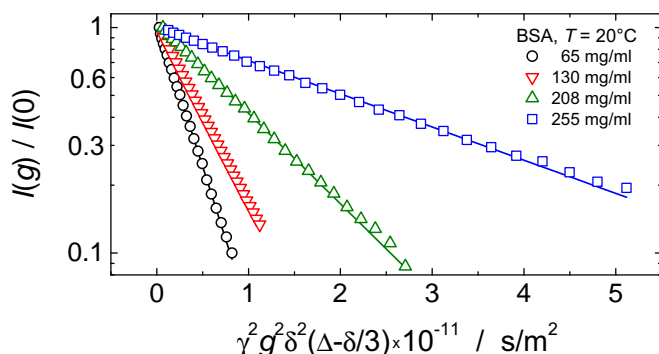
In the experiments performed here, translational diffusion is accessed on a time scale of  $\Delta \sim 20 - 70$  ms. With protein translational diffusion coefficients in the order of  $10^{-10} \text{ m}^2/\text{s}$  down to  $10^{-12} \text{ m}^2/\text{s}$ , the root mean-square displacement  $\langle \mathbf{r}^2 \rangle^{1/2}$  is still in the range of micrometers, i.e. significantly larger than the protein's size or the dimension of the macromolecular cage. Consequently, protein translational diffusion as detected by PFG NMR clearly belongs to the long-time regime. Note that in all experiments performed here, the gradient pulse points along the  $z$ -axis, hence  $\mathbf{g} \cdot \mathbf{r} = g_z z$ , with an effective gradient duration (cf. also eq. (3.18)) of  $\delta = 1$  ms.

The pulsed gradient induces eddy currents in the metal structure of the gradient equipment, giving rise to field transients. These uncontrolled magnetic fields distort the NMR signal obtained. The magnitude of eddy currents can be decreased by a suitable shape of the gradient pulses, avoiding a strong rise or quick lowering of the gradient burst (e.g. via sinusoidal or trapezoidal gradient shapes). Additionally, a pre-emphasis of the gradient pulse [Majors et al., 1990] is employed to counterbalance the eddy currents occurring during and after the gradient burst as much as possible. Nevertheless, such treatment cannot provide a perfect compensation, especially in presence of rather long-lasting ring-down effects after switching off the gradient pulse.<sup>7</sup> For this reason, each gradient pulse is followed by a delay of 1 ms during which residual eddy currents decay. See also the longitudinal eddy-current delay (LED) scheme by Gibbs and Johnson Jr. [1991].

The accuracy of diffusion coefficients determined by PFG NMR can be further improved by avoiding the impact of background gradients; these arise, in particular, at regions of distinct magnetic susceptibilities, e.g. at the water–air interface. Their impact on the measurement result can be greatly reduced by using bipolar gradient pulses [Cotts et al., 1989]. Here, the single gradient pulse is split into two separate gradient pulses of opposed direction (sign), combined with a  $(\pi)$ -pulse in-between; see Fig. 3.13d. In this way, the spin ensemble is sensitive to both gradient pulses added together, whereas background gradients have the same direction (sign) and are averaged out. For all measurements of

<sup>7</sup>Remember that (high-field) NMR relies on a homogeneity of the static magnetic field  $\mathbf{B}_0$  in the range of ppm and below.

### 3. Experimental methods



**Figure 3.14.:** Exemplary PFG NMR diffusion decays for BSA at different concentrations (symbols). Note the mono-exponential nature of the decays. The slope reflects the translational diffusion coefficient that is obtained by the use of eq. (3.20) (solid lines).

translational diffusion in *single*-protein solutions, the bipolar pulse scheme was used. A drawback of bipolar gradient pulses is that they are longer than the monopolar gradient pulse, which is mainly caused by the ring-down delays incorporated after each gradient pulse. The larger time span during which the spin ensemble is subject to transverse relaxation goes along with a reduced signal intensity and hence a decreased signal-to-noise ratio. For measurements on protein *mixtures*, including an isotope-filter to only select one out of the two proteins dissolved, the reduced signal strength restricted the measurements<sup>8,¶</sup> to the use of monopolar gradients only. The difference in absolute diffusion coefficients determined by the mono- or bipolar STE, however, is for our setup and samples less than 3% [Kirschke, 2013]. Furthermore, to avoid convection effects, the maximum sample height inside the NMR tube was limited to about 8 mm. The impact of convection on the measuring result can also be compensated to first order using the Double Stimulated Echo sequence [Jerschow and Müller, 1997], which basically is a duplicated version of the STE scheme. Again, such a sequence is problematic when applied to highly concentrated proteins of which  $T_2$  is naturally short.

For data treatment, the integrated signal of the aliphatic region (cf. Fig. 3.1) was evaluated, providing the PFG NMR diffusion decay; see Fig. 3.14. Amide protein protons, in contrast, are subject to chemical exchange effects, which could distort the diffusion decay. This aspect has been addressed via additional measurements presented in the appendix, section S.3. From this study, it follows that a concentration of about 250 g/l gives a reasonable limit for BSA and LYZ till which bias effects are safely absent. The same holds for  $\alpha$ B-crystallin at a concentration of around 180 g/l. These numbers give the highest concentrations used in the studies of crowded protein Brownian motion presented in this thesis. Finally, for more details on PFG NMR including technical aspects, see also the reviews by Price [1997, 1998], Johnson Jr. [1999], and Antalek [2002].

<sup>8</sup>Measurements on protein mixtures to great extent performed by MAIK ROTHE within his Master’s thesis [Rothe, 2015].

<sup>¶</sup>Supervision by MATTHIAS ROOS and KAY SAALWÄCHTER.

### 3.5. Isotope filtering for the measure of protein mixtures

In protein mixtures, given the multitude of resonances of each protein combined with structural similarities of the proteins, the spectra of different proteins strongly overlap. As such, when solely relying on a simple, one-dimensional spectrum, distinction of the NMR signal of one protein species from that of another is (nearly) impossible. For data interpretation beyond the overall slow-down of the protein diffusivity with increasing protein concentration, however, distinction between different proteins is required.

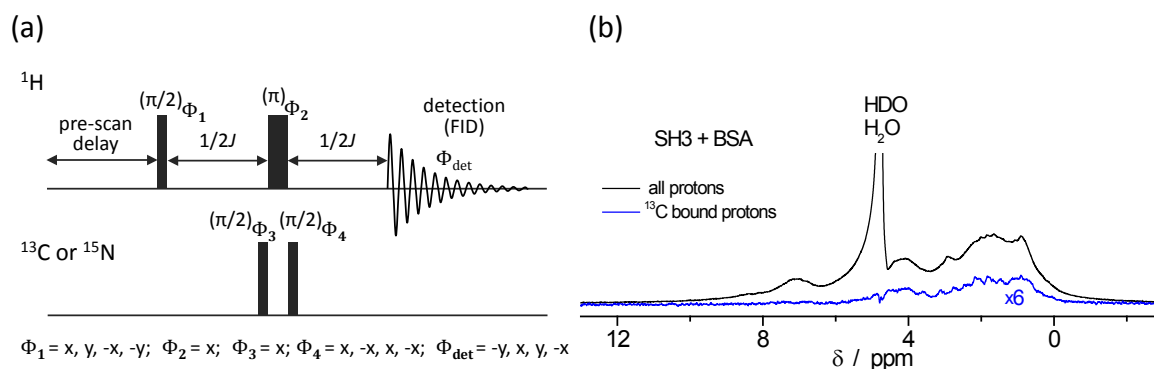
Often, distinct diffusion coefficients permit signal decomposition, providing the basis for diffusion ordered nuclear magnetic resonance spectroscopy (DOSY) spectra. Those experiments still rely on the STE pulse scheme, yet analyzed in a way that provides the spectrum of each component (abscissa) with respect to the accordant diffusion coefficient (ordinate). DOSY spectra are particularly helpful for low-molecular weight liquids. In the study of protein diffusivity in protein mixtures, by contrast, a distinction of the diffusion coefficients *directly* based upon the detected signal would be highly desirable, in particular as the diffusion coefficients of the two proteins may be close to each other. For this purpose, NMR isotope filtering was incorporated, i.e., only the signal of an isotope-labeled ( $^{13}\text{C}$  or  $^{15}\text{N}$ ) protein is detected (or deselected) out of the mixture with unlabeled compounds. Basically, any NMR experiment relying on a signal from a hetero-nucleus ( $^{13}\text{C}$ ,  $^{15}\text{N}$ ) may serve for this purpose, as well as the STE pulse scheme may also be directly applied to such nuclei. However, to achieve suitable encoding/decoding sensitivity, the latter type of experiments would require much stronger (or longer) gradient pulses to compensate for the lower magnetogyric ratio. Thus, measuring protein protons is highly recommended, combined with the also intrinsically stronger NMR signal of protein protons as for other nuclei.

In practice, the pulse scheme proposed by Tillett et al. [1999] is applied, which relies on indirect spin-spin couplings mediated by chemical bonds (so-called J-coupling or *indirect* dipole-dipole coupling). This pulse scheme exploits the idea that the (spin-1/2) hetero-nucleus can influence exclusively those proton spins that are in close chemical environment<sup>9</sup> to this spin. If one protein species is  $^{13}\text{C}$  or  $^{15}\text{N}$  enriched but the other(s) is (are) not, this effect allows the detection of signals from *only* the isotopically enriched protein. *Vice versa*, appropriate phase cycling can also selectively detect those NMR signals that originate from protons that are *not* bound to the spin-carrying hetero-nucleus.

The accordant pulse scheme is shown in Fig. 3.15a and was implemented and tested on our PFG NMR diffusion setup by SEBASTIAN KIRSCHKE within his Bachelor's thesis<sup>¶</sup> [Kirschke, 2013]. To avoid the impact of chemical exchange between protein amide protons and water protons, the pulse scheme was applied on  $^{13}\text{C}$  rather than on  $^{15}\text{N}$ . In this way,

<sup>9</sup>In practice, J-couplings are restricted to not more than three chemical bonds between the two spins. Usually, one relies on single-bound J-couplings, i.e.,  $^{13}\text{C}$ - $^1\text{H}$  or  $^{15}\text{N}$ - $^1\text{H}$ .

### 3. Experimental methods



**Figure 3.15.:** NMR isotope filtering: basic pulse sequence (a) and exemplary application (b) to a protein mixture composed of a  $^{13}\text{C}/^{15}\text{N}$  enriched SH3 domain and unlabeled BSA (mass ratio 1:2, with a total protein concentration of 65 g/l,  $T = 18^\circ\text{C}$ ). Here, the isotope filter was applied to  $^{13}\text{C}$ , thus reflecting the aliphatic region of SH3. Note the absence of water and amide protons in the filtered spectrum (blue) as compared to the spectrum recorded straight after a  $(\pi/2)$ -pulse, in which all sample protons contribute to the spectrum (black). Data recorded by MAIK ROTHE within his Master's thesis.

the isotope-filtered spectrum is strongly dominated by aliphatic protons, whereas amide (and water) protons are largely suppressed, cf. Fig. 3.15b.

In PFG NMR diffusion experiments, the isotope filter can be combined with the encoding/decoding period of the STE pulse sequence [Tillett et al., 1999; Nesmelova et al., 2004]. For a more detailed introduction to the isotope filter, see the Bachelor's thesis by Kirschke [2013].

### 3.6. Viscosimetry

Viscosity was primarily measured using the capillary viscometer mVROC\* (Viscometer/Rheometer-On-a-Chip; RheoSense, San Ramon, CA) that requires sample volumes that are as small as 20  $\mu\text{l}$ . Its working principle relies on knowledge of the shear stress  $\sigma$  and shear rate  $\dot{\gamma}$  of the liquid passing a rectangular slit channel. Using NEWTON's law of viscosity,

$$\sigma = \eta \dot{\gamma}_{(\text{app})} \quad , \quad (3.21)$$

one can estimate the viscosity  $\eta$ . The (apparent) shear rate is determined here by the volumetric flow rate  $\Phi$  and the given geometry of the slit,  $\dot{\gamma}_{\text{app}} = 6\Phi/(wh^2)$ ,<sup>||</sup> where  $w$  and  $h$  are the width and depth of the slit channel. The flow of liquid through a (cylindrical)

\*Instrumentation was purchased in the context of this project to enable measurements using only small sample volumes. Funding is acknowledged to the transregional collaborative research center SFB-TRR 102. Restricted sample volumes were of particular relevance for the recombinantly expressed  $\alpha\text{B}$ -crystallin, provided by SUSANNE LINK (Martin Luther University Halle-Wittenberg, Biophysics) in our laboratory.

<sup>||</sup>Information provided by RheoSense



pipe (or capillary) is accompanied by a drop of pressure  $p$  over a distance of  $L$  along the capillary,

$$\Phi = \frac{\pi a^4}{8\eta} \frac{|\Delta p|}{L} . \quad (3.22)$$

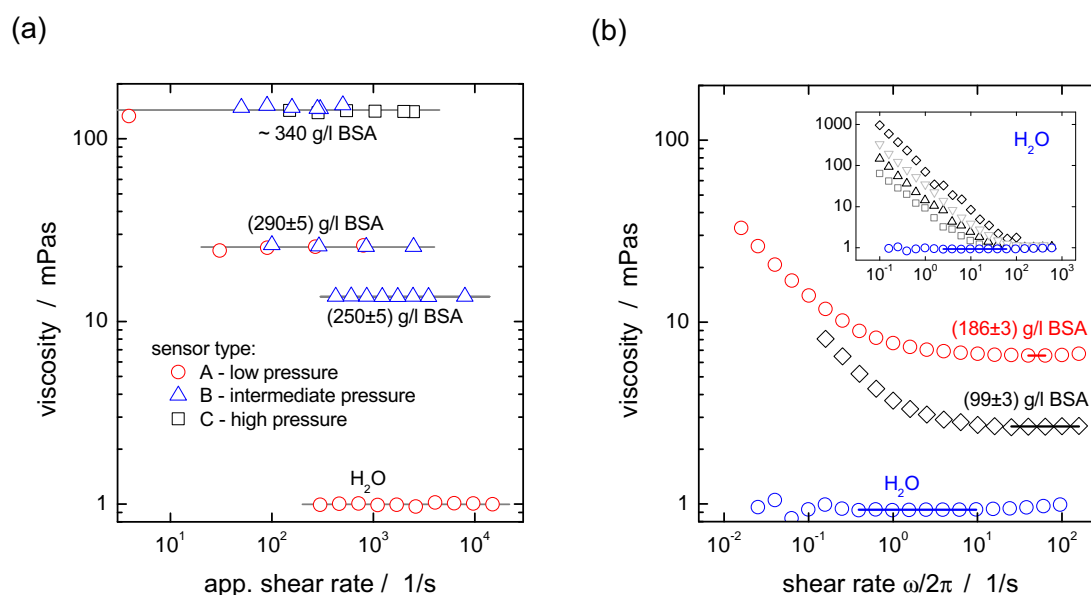
Eq. (3.22) is known as the HAGEN-POISEUILLE equation, with  $a$  denoting the capillary's inner radius. Although the sample channel is a rectangular one, the physical basis stays the same. From the pressure gradient along the slit inside the mVROC, the shear stress  $\sigma$  is determined by making use of  $\frac{\partial p(L)}{\partial L} \propto \sigma$ , with the proportionality factor being again determined by the slit geometry.<sup>||</sup>

Newtonian fluids such as water show no dependence of the viscosity on the shear rate. Such fluids usually consist of (small) isotropic molecules, with no alignment effects of the particles under shear flow. Anisotropic molecules may (but need not to) show non-Newtonian behavior; a reason for which the shear-rate dependence of the viscosity of the highly concentrated protein solutions was examined in a preparatory study. As seen from Fig. 3.16a a Newtonian behavior is observed over the full range of shear rates accessible by the mVROC even for protein concentrations as high as  $\sim 340$  g/l. For shear rates higher than those accessed here, shear-thinning cannot be excluded: flow-induced alignment effects of the usually anisotropic particles can promote a reduced viscosity of the structurally perturbed solution. In the here presented studies, the standard shear rate was chosen to be  $1000 \text{ s}^{-1}$  or  $2000 \text{ s}^{-1}$ : the pressure  $p(\Phi)$  should be low enough to be harmless for the pressure sensors, but high enough for optimal sensitivity. Using the most sensitive sensor (type "A" in Fig. 3.16a) and a shear rate of  $\sim 10^3 \text{ s}^{-1}$  viscosity measurements can be performed at concentrations of more than 300 g/l, which covers the full range of protein concentrations studied within this project.

An absence of a shear-rate dependence may be different for cone-plate or plate-plate rheometers as well as for rheometers using concentric cylinders, the measuring result of the latter setup is shown in Fig. 3.16b. In these measurements, the viscosity is calculated from the torque needed to rotate the inner cylinder (or the upper cone/plate) relative to the outer cylinder (or the lower plate). At low frequencies, these measurements are sensitive to liquid/air interface effects [Johnston and Ewoldt, 2013]. This effect is demonstrated by the impact of air bubbles on the *apparent* viscosity of water (see inset in Fig. 3.16b). For the protein solutions, note the unusual concentration dependence of the apparent shear thinning effect if this were related to the behavior of the protein solution: the amplitude of the shear rate dependence is stronger at around 100 g/l than at the higher concentration of around 190 g/l.

Several research articles [Inoue and Matsumoto, 1996; Ikeda and Nishinari, 2000, 2001a,b; Brownsey et al., 2003] report on shear-thinning effects of protein solutions for shear rates in-between  $10^{-2}$  and  $10^2 \text{ s}^{-1}$  – a time scale that matches that one in Fig. 3.16b. In these articles, the experimentally observed shear rate dependence was argued to originate from

### 3. Experimental methods

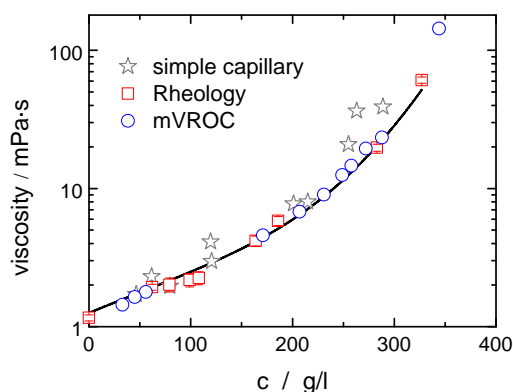


**Figure 3.16.:** Shear-rate dependence of the viscosity of BSA solutions as obtained by the mVROC (a) and by the use of a rotational rheometer using concentric cylinders\*\* (b). The inset in Fig. (b) gives rheology measurements of water in the presence (black/gray data points) or absence (blue data points) of air bubbles artificially created using a pipette. The lines in each figure give the taken value of the viscosity.

long-range (electrostatic) repulsive forces that induce lattice-like structures of the colloidal arrangement in solution. The molecular packing was considered to be distorted by further adding BSA [Brownsey et al., 2003] to explain the unusual concentration dependence depicted in Fig. 3.16b. Different to these articles, emphasis on the gas-liquid interface effect has been given by Sharma et al. [2011], stressing the importance of interfacial viscoelasticity that is linked to surface-active components of the solution. After all, Johnston and Ewoldt [2013] showed that measurements using rotational rheometers are *generally* subject to a rate-independent torque that is needed to compensate surface tension and liquid-air contact line effects. With increasing angular velocity, rotational rheometers become less sensitive to the contribution of this constant torque on the measurement result. This situation generally leads to an *apparent* shear thinning behavior of the specimen.

Using steady-shear experiments with concentric cylinders\*\*, but evaluating only shear rates that are high enough to be not sensitive to interface effects, the values obtained agree well with those from the mVROC, see Fig. 3.17. Nevertheless, viscosity determination by a conventional rheometer was impracticable for the  $\alpha$ B-crystallin solutions due to the large amount of sample needed (several ml). Also note that in measurements using a rheometer, a film of silicon oil had to be added to the surface of the sample to prevent sample drying effects (see also Brownsey et al. [2003]). For the mVROC, due to *no* contact of the liquid with air, such sample treatment is not required.

\*\*Instrumentation of the group of Experimental Polymer Physics, Martin Luther University Halle-Wittenberg



**Figure 3.17.:** Concentration dependence of the viscosity of BSA solutions as measured by the mVROC, conventional rheology and a simple capillary. To easily compare with the NMR samples, D<sub>2</sub>O was used as the solvent.

Before the availability of the mVROC in our laboratory and before the rheometer measurements were performed, the first test measurements were performed using a simple capillary, with the viscosity determined by the use of the HAGEN-POISEUILLE equation. For details, see the Supporting Information, section S.4. Combined with data on translational diffusion, the preliminary viscosity data already indicated the value of a more precise study of the viscosity concentration dependence (the precision of this approach is reflected in Fig. 3.17, star symbols), and initiated the viscosity studies using professional equipment.

### 3.7. Concentration determination

Determining the protein concentration usually relies on photoabsorption using the equation by LAMBERT and BEER,

$$\ln\left(\frac{I_0}{I_1}\right) = \epsilon c d \quad . \quad (3.23)$$

Here,  $I_0$  and  $I_1$  denote the intensity of the incoming and transmitted light beam, respectively,  $d$  is the path length through the specimen, and  $\epsilon$  is the (molar) attenuation coefficient of the attenuating particles of concentration  $c$ . This technique is limited to low concentrations. Thus, a high dilution of the concentrated protein solution is required, with dilution factors as high as 1000. Highly concentrated protein solutions are quite viscous, causing pipetting to be less precise. Reliably and reproducibly determined protein concentrations require repetitive measurements and a well-controlled dilution series. Only the concentrations of the  $\alpha$ B-crystallin solutions was determined in this way.<sup>10</sup>

Unlike the  $\alpha$ B-crystallin samples, no buffer or salt was added to the BSA/LYZ/SH3 solutions. Avoiding high dilution factors, the BSA/LYZ/SH3 protein concentration can thus be estimated by *directly* determining the mass of protein dissolved per volume. For this purpose, the concentrated protein solution of known volume was lyophilized and weighed. Note that a fraction of the (highly concentrated and thus viscous) protein solution

<sup>10</sup>Performed by Susanne Link.

### 3. *Experimental methods*

usually remains inside the pipette's tip, which was accounted for during concentration determination. Every mass was checked at least three times, with a reproducibility of usually 0.1 to 0.3 mg. The reproducibility of the value obtained for  $c$  turned out be better than 3 g/l for concentrations up to about 200 g/l, and  $\pm 5$  g/l for concentrations in-between 200 g/l and 300 g/l.

# 4 The impact of macromolecular crowding on protein diffusivity

### 4.1. NMR-detected Brownian dynamics of $\alpha$ B-crystallin over a wide range of concentrations

Crowding conditions are prototypical for the eye lens. A major constituent of the mammalian eye lens is  $\alpha$ -crystallin – a large protein complex that is composed of two subunits,  $\alpha$ A- and  $\alpha$ B-crystallin, with a molar ratio of about 3:1. The occurrence of  $\alpha$ A-crystallin is rather specific for the eye lens;  $\alpha$ B-crystallin is widely spread across the body [Bhat and Nagineni, 1989; Dubin et al., 1989] and is involved in serious neurological diseases such as Alzheimer’s disease [Lowe et al., 1992], Parkinson’s disease [Iwaki et al., 1992], Creutzfeldt-Jakob disease [Iwaki et al., 1992; Renkawek et al., 1992], and Alexander’s disease [Iwaki et al., 1989].  $\alpha$ -crystallin behaves chaperone-like in that it non-reversibly binds to (partially) denatured proteins. In this way,  $\alpha$ -crystallin inhibits the development of larger protein aggregates that, particularly in the eye lens, would cause or increase light scattering.  $\alpha$ -crystallin has thus been implicated to be of central importance for preventing cataracts; see also Derham and Harding [1999], Horwitz [2000, 2003] or Bloemendal et al. [2004]. Also note that crystallization of  $\alpha$ -crystallin is difficult to achieve as it would be accompanied by cataract formation, and is thus avoided by nature. In fact, efforts to crystallize this protein have not been successfully yet. The name “crystallin” originates from the first isolation of these proteins from the so-called “crystalline lens” by Berzelius in 1830 [Augusteyn, 1998]; the naming has persisted to the present day despite it being a bit misleading.

In the following, the investigation of Brownian dynamics in concentrated, recombinantly expressed  $\alpha$ B-crystallin solutions will be presented. Similar to natural  $\alpha$ -crystallin [Peschek et al., 2009],  $\alpha$ B-crystallin [Braun et al., 2011] forms large spherical assemblies with rotational correlation times of about 1  $\mu$ s. In particular, NMR  $R_{1\rho}$  relaxation measurements will be used to determine these long rotational correlation times.

*Author contributions:* K.S., J.B. and A.K. designed research; S.L. prepared samples; A.K. and M.R. performed measurements; M.R. analyzed the data; M.R., J.B., A.K. and K.S. interpreted the findings; A.K., M.R. and K.S. wrote the paper with refinements by all coauthors.

The following article [M. Roos et al. NMR-detected Brownian dynamics of  $\alpha$ B-crystallin over a wide range of concentrations. *Biophys. J.* 108: 98-106, 2015] is reprinted with the permission by Elsevier that the thesis including the article can be posted publicly by the awarding institution. The link back to the ScienceDirect publication is <http://www.sciencedirect.com/science/article/pii/S0006349514030707> .  
No changes were made.

## Article

# NMR-Detected Brownian Dynamics of $\alpha$ B-Crystallin over a Wide Range of Concentrations

Matthias Roos,<sup>1</sup> Susanne Link,<sup>1</sup> Jochen Balbach,<sup>1</sup> Alexey Krushelnitsky,<sup>1,\*</sup> and Kay Saalwächter<sup>1,\*</sup><sup>1</sup>Institut für Physik, Martin-Luther-Universität Halle-Wittenberg, Halle (Saale), Germany

**ABSTRACT** Knowledge about the global translational and rotational motion of proteins under crowded conditions is highly relevant for understanding the function of proteins in vivo. This holds in particular for human  $\alpha$ B-crystallin, which is strongly crowded in vivo and inter alia responsible for preventing cataracts. Quantitative information on translational and rotational diffusion is not readily available, and we here demonstrate an approach that combines pulsed-field-gradient NMR for translational diffusion and proton  $T_{1\rho}/T_2$  relaxation-time measurements for rotational diffusion, thus overcoming obstacles encountered in previous studies. The relaxation times measured at variable temperature provide a quantitative measure of the correlation function of protein tumbling, which cannot be approximated by a single exponential, because two components are needed for a minimal and adequate description of the data. We find that at high protein concentrations, rotational diffusion is decoupled from translational diffusion, the latter following the macroscopic viscosity change almost quantitatively, resembling the behavior of spherical colloids. Analysis of data reported in the literature shows that well-packed globular proteins follow a scaling relation between the hydrodynamic radius and the molar mass,  $R_h \sim M^{1/d}$ , with a fractal dimension of  $d \sim 2.5$  rather than 3. Despite its oligomeric nature,  $R_h$  of  $\alpha$ B-crystallin as derived from both NMR methods is found to be fully consistent with this relation.

## INTRODUCTION

In a living cell, proteins exist and function in a rather concentrated solution of a wide range of different solutes. In comparison with dilute conditions, such crowding can significantly alter the protein behavior (1,2). The most important parameters in describing protein overall Brownian motion are the translational and rotational diffusion coefficients. Obviously, crowding increases the viscosity of the solution and slows down protein diffusion. However, this differs from a simple increase of the solution viscosity by adding, for instance, glycerol (3). Intermolecular protein interactions and their influence on Brownian diffusion are rather complicated in nature, which cannot be effectively described by increased viscosity alone. The complex changes of the protein dynamics at high protein concentrations, and the key factors determining these changes, are largely unclear at present. Experimental data on this topic are still rather sparse. Some results even contradict each other; for example, fluorescence data demonstrate that upon increasing protein concentration, translational diffusion is slowed down to a larger extent than rotational diffusion (4), whereas NMR experiments yield the opposite conclusion (3).

The interior of the vertebrate eye lens is a typical example of a crowded protein solution. Here, a highly concentrated mixture of short-range ordered (5)  $\alpha$ -,  $\beta$ -, and  $\gamma$ -crystallins

provides a high refractive index and lens transparency without protein metabolism (6,7). The main constituent of this protein mixture is  $\alpha$ -crystallin, which comprises ~35% (w/w) of the lens crystallins (8).  $\alpha$ -crystallin consists of two homologous proteins,  $\alpha$ A- and  $\alpha$ B-crystallin, which have a monomer molecular mass of ~20 kDa each. They form oligomeric associations with a molecular mass distribution from 500 to >1000 kDa and an average mass of ~800 kDa (9). Besides maintaining the high refractive index,  $\alpha$ -crystallin acts as a molecular chaperone, preventing protein aggregation that causes cataracts to form (10,11).

Although  $\alpha$ B-crystallin has been studied quite intensively over the last decades (for reviews, see Narberhaus (12), Horwitz (13), Augusteyn (14), and Andley (15)), its dynamics, especially at high concentrations, has not been investigated in much detail. Delaye et al. concluded that  $\alpha$ -crystallin acts as a good model system for colloids with an effective hard-sphere radius that is not dependent on concentration (16,17), with translational self-diffusion coefficients (SDCs) that closely follow the macroscopic viscosity (18). Conversely, another report indicated that  $\alpha$ -crystallin does not form a compact sphere at all (17) but has a dynamic quaternary structure (19).

Here, we present a detailed comparative study of the rotational and translational diffusion of  $\alpha$ B-crystallin as a function of concentration. The translational and rotational dynamics of  $\alpha$ B-crystallin were studied by pulsed-field-gradient (PFG) NMR and proton NMR relaxation-time measurements, respectively. PFG NMR provides an objective and robust measure of the SDC, even at high concentrations

Submitted May 15, 2014, and accepted for publication November 11, 2014.

\*Correspondence: [krushelnitsky@physik.uni-halle.de](mailto:krushelnitsky@physik.uni-halle.de) or [kay.saalwaechter@physik.uni-halle.de](mailto:kay.saalwaechter@physik.uni-halle.de)

Editor: Patrick Loria.

© 2015 by the Biophysical Society  
0006-3495/15/01/0098/9 \$2.00

<http://dx.doi.org/10.1016/j.bpj.2014.11.1858>



where dynamic light-scattering (DLS) experiments are challenged by the appearance of a slow mode and the necessary conversion of a cooperative diffusion coefficient into an SDC (20). Due to its large aggregate size, the resonances of the rigid core of  $\alpha$ B-crystallin are broadened beyond detection (9) and thus cannot be studied by conventional high-resolution NMR techniques. We rely instead on low-resolution  $^1\text{H}$  NMR relaxation-time measurements at various concentrations and temperatures, which in fact allows us to address potential ambiguities related to the common use of the NMR  $T_1/T_2$  relaxation-time ratio as a measure of rotational diffusion. For the relaxation measurements, we rely on an integration of the whole proton spectrum, thus analyzing the integral signal from all protein protons. We present a consistent and quantitative treatment of relaxation data in terms of a bicomponent rotational autocorrelation function. Our analysis reveals a progressive decoupling of translational and rotational motion upon an increase in concentration.

## MATERIALS AND METHODS

### Sample preparation

#### *Protein expression and purification of human $\alpha$ B-crystallin*

The human  $\alpha$ B cDNA (construct kindly provided by Prof. Wilbert Boelens, Nijmegen Centre for Molecular Life Sciences, Gelderland, The Netherlands) was cloned into a modified, His-tag-free pET16b vector and expressed in *Escherichia coli* BL21(DE3). Protein expression and purification were performed as described in Mainz et al. (21), with minor modifications including autoinduction media (ZYM 5052) instead of minimal media (M9), microfluidizer instead of French press, an additional DNA digestion step after cell lysis, and different column materials: in place of Q-Sepharose and Superose 6, TMAE and Superdex 200, respectively, were used. The lyophilized sample was dissolved in 50 mM Na-phosphate buffer, 50 mM NaCl, and 0.002%  $\text{NaN}_3$   $\text{D}_2\text{O}$  buffer, pH 7.6, with prior minimization of labile protons to lower the water signal and so as not to have an impact on the solvent viscosity via isotope effects. For more details, see the [Supporting Material](#). Hen egg white lysozyme was delivered from Sigma-Aldrich (St. Louis, MO). Similar to  $\alpha$ B-crystallin, lysozyme was dissolved in  $\text{D}_2\text{O}$ , lyophilized, and dissolved in  $\text{D}_2\text{O}$  again for maximal removal of residual water protons.

### Viscosity

Steady-shear viscosities were measured at high shear rates ( $1000\text{ s}^{-1}$ / $2000\text{ s}^{-1}$ ) using the microfluid viscometer-rheometer on chip (m-VROC, Rheosense, San Ramon, CA), which determines the sample viscosity by analyzing the pressure gradient inside of a capillary ( $d = 50\ \mu\text{m}$ ). The samples for the viscosity measurements were also prepared using  $\text{D}_2\text{O}$  buffer instead of  $\text{H}_2\text{O}$ .

### NMR experiments

Translational diffusion and  $T_{1\rho}$  measurements were conducted on a Bruker Avance II spectrometer (Billerica, MA) with a  $^1\text{H}$  resonance frequency for protons of 400 MHz equipped with a Diff60 probehead.  $T_{1\rho}$ s were measured at spin-lock frequencies of 20, 40, and 60 kHz; the latter was measured using a resonance offset of the spin-lock field with angle  $\theta$  between the  $B_0$  and

$B_{1e}$  fields fixed to  $42^\circ$ .  $T_2$  measurements were performed on a Bruker Minispec mq20 at 20 MHz  $^1\text{H}$  resonance frequency. The low-resonance frequency for  $T_2$  experiments was chosen to avoid  $T_2$  shortening due to the chemical exchange of protein protons, which may significantly affect  $T_2$  values at high resonance frequencies (22). For the relaxation measurements, in all cases, we employ single short-pulse excitation and a sufficiently large spectral width of 50 kHz, thus assuring that all types of protons (rigid and mobile) in the protein contribute equally to the integral signal. In all cases, the accuracy of the temperature calibration and stabilization was  $\pm 1^\circ\text{C}$ . For more details, see the [Supporting Material](#).

Translational SDCs were obtained from the PFG NMR diffusion decays using the well-known formula (23)

$$A(g) = A(0) \times \exp(-\gamma^2 g^2 D \delta^2 (\Delta - \delta/3)), \quad (1)$$

where  $A(g)$  is the signal intensity,  $g$  is the field gradient strength,  $\gamma$  is the proton gyromagnetic ratio,  $\Delta$  is the diffusion time,  $\delta$  is the duration of the field gradient pulse, and  $D$  is the SDC.

Rotational correlation times were obtained by analyzing NMR relaxation times. These are determined by the spectral density function, which is the Fourier transformation of the rotational autocorrelation function (RACF). The RACF of protein motion in solution is complicated in nature. For its unambiguous determination from experimental data, multiple measurements of relaxation times at different resonance frequencies are required, since each relaxation time reflects molecular dynamics only within the frequency domain around the circular (i.e., multiplied by  $2\pi$ ) resonance frequency. Because of the high molecular mass of the  $\alpha$ B-crystallin oligomer, its Brownian tumbling is very slow, and thus, the often employed  $T_1$  relaxation times are not useful for studying such a slow motion since they provide information on (sub)nanosecond-timescale motions, which is much faster than the  $\alpha$ B-crystallin tumbling. For this reason, we used  $T_{1\rho}$  proton relaxation times, which enable one to shift the sampling frequency of motions down to the 10–100 kHz range. Specifically, we measured the temperature dependences of  $T_{1\rho}$  values at spin-lock frequencies (the analog of the resonance frequency for  $T_1$ ) of 20, 40, and 60 kHz, the latter values being measured using the resonance offset of spin-lock irradiation (see the [Supporting Material](#)). We stress that measuring relaxation times at different temperatures is important for a reliable data analysis, since the slope of the temperature dependence is more informative than the absolute value of the relaxation time measured at one temperature in correctly determining the rotational correlation time. In addition to  $T_{1\rho}$  values, we also measured the proton  $T_2$ , which provides the low-frequency limit of the spectral density function. The relaxation is governed by the homonuclear ( $^1\text{H}$ - $^1\text{H}$ ) dipole-dipole mechanism; all other mechanisms are negligible in this case. The equations for the homonuclear dipolar  $T_1$ ,  $T_2$ , and  $T_{1\rho}$  relaxation times are well known (24). However, since we measured off-resonance  $T_{1\rho}$ , for the data analysis, we should use a general expression defining relaxation times  $T_1$  and  $T_{1\rho}$  at arbitrary off-resonance angle. Such an expression was derived a long time ago by Jones (25). In a more compact form, applying the approximation  $\omega_0 \gg \omega_{1e}$ , this expression reads (26)

$$\frac{1}{T_{1\rho}} = \frac{1}{T_1} + \sin^2\theta \left[ \frac{1}{T_{1\rho}^\Delta} - \frac{3}{4T_1} \right], \quad (2)$$

where

$$\frac{1}{T_1} = \frac{2}{3} K_{\text{HH}} (J(\omega_0) + 4J(2\omega_0)), \quad (3)$$

$$\frac{1}{T_{1\rho}^\Delta} = K_{\text{HH}} \left( \cos^2\theta \times J(\omega_{1e}) + \sin^2\theta \times J(2\omega_{1e}) + \frac{3}{2} J(\omega_0) \right). \quad (4)$$



Here,  $\theta$  is the off-resonance angle (the angle between the static field  $B_0$  and effective spin-lock field  $B_{1c}$ ),  $J(\omega)$  is the spectral density function,  $K_{\text{HH}}$  is the squared effective proton-proton dipolar coupling (second moment), and  $\omega_0/2\pi$  and  $\omega_{1c}/2\pi$  are the resonance and spin-lock frequencies, respectively. At  $\theta = 0^\circ$  and  $90^\circ$ , Eq. 2 converts to the standard expressions for  $T_1$  and  $T_{1\rho}$ , respectively. Relaxation time  $T_2$  corresponds to the case  $\theta = 90^\circ$  and  $\omega_{1c} = 0$ . Note that these equations are valid not only in the fast-motion limit, but for slow motions as well (27). For the case of the relaxation of protein protons, the effective  $K_{\text{HH}}$  can be expressed as

$$K_{\text{HH}} = \frac{9}{20} \times \frac{1}{N} \times \hbar^2 \gamma^4 \sum_{i \neq j} r_{ij}^{-6}, \quad (5)$$

where  $N$  is the number of protons in a protein,  $\hbar$  is the Planck constant,  $\gamma$  is the proton gyromagnetic ratio, and  $r_{ij}$  is the distance between the  $i$ th and  $j$ th protons in the protein. Although summation in Eq. 5 formally extends over all protons in the protein, because of the  $r^{-6}$  dependence, the dominant contribution to the coupling for each proton is attributable to the two to three nearest neighbors, and proton-proton interactions with more distant neighbors are practically negligible.

Since an integral proton signal was detected in the relaxation experiments, the spectral density function  $J(\omega)$  in Eqs. 2–4 is the average spectral density of all protons in the protein:

$$J(\omega) = \frac{1}{N} \sum J_i(\omega). \quad (6)$$

The same is true for the RACF:

$$C(t) = \frac{1}{N} \sum C_i(t). \quad (7)$$

Each individual RACF can be written as a product of the correlation functions of the overall Brownian tumbling and internal local motion (28),

$$\begin{aligned} C_i(t) &= C_t(t) \times C_{li}(t) \\ &= C_t(t) \times \left[ S_{li}^2 + (1 - S_{li}^2) \exp\left(-\frac{t}{\tau_{li}}\right) \right], \end{aligned} \quad (8)$$

where  $C_t(t)$  is the overall tumbling RACF and  $S_{li}^2$  and  $\tau_{li}$  are the order parameter and correlation time of the internal motion for the  $i$ th proton. Since  $C_t(t)$  is the same for all protons, Eq. 7 can be rewritten as

$$C(t) = C_t(t) \times \left[ S_t^2 + \frac{1}{N} \sum (1 - S_{li}^2) \exp\left(-\frac{t}{\tau_{li}}\right) \right], \quad (9)$$

where  $S_t^2 = 1/N \sum S_{li}^2$ . Then, the spectral density function is

$$J(\omega) = S_t^2 J_t(\omega) + \frac{1}{N} \sum \frac{(1 - S_{li}^2) \tau_{li}}{1 + (\omega \tau_{li})^2}, \quad (10)$$

where  $J_t(\omega)$  is the Fourier transform of  $C_t(t)$ . In our analysis, we assume that all values of  $\tau_{li}$  are much smaller than the correlation time of the overall protein tumbling. Then, the second term in Eq. 10 can be neglected. This formalism is very similar to that applied in the analysis of field-cycling  $T_1$  relaxation data of protein protons in  $\text{D}_2\text{O}$  solutions (29).

The key point of our analysis is an assumption of a biexponential overall tumbling RACF  $C_t(t)$ . Even in relatively dilute protein solutions, long-range electrostatic intermolecular protein interactions give rise to a local anisotropy that renders the so-called normal Brownian tumbling somewhat anisotropic (30,31). Thus, the RACF decays not to zero, but to a certain value that we denote as the rotational order parameter,  $S_{\text{rot}}^2$ . Its physical meaning is similar to that of the order parameter of the internal motions (28): both are measures of the anisotropy of rotational motion of the overall tumbling

and internal mobility, respectively. In infinitely dilute solutions,  $S_{\text{rot}}^2 = 0$ , and it increases with increasing concentration, as interprotein interactions become stronger and Brownian tumbling thus becomes more anisotropic. Since proteins diffuse relative to each other, the local anisotropy has a finite lifetime; hence,  $C_t(t)$  finally decays to zero, but on a longer timescale than for normal Brownian tumbling. Therefore,  $C_t(t)$  can be presented as a sum of two components,

$$C_t(t) = (1 - S_{\text{rot}}^2) \exp(-t/\tau_{\text{rot}}) + S_{\text{rot}}^2 \exp(-t/\tau_S), \quad (11)$$

where  $\tau_{\text{rot}}$  is the correlation time of Brownian rotation and  $\tau_S$  is the correlation time of the slow component of  $C_t(t)$ , i.e., the lifetime of local anisotropy. The apparent slow contribution to protein Brownian tumbling has been observed experimentally and computationally in a number of independent works (for a review, see Krushelnitsky (31)).

The corresponding spectral density function reads

$$J_t(\omega) = \frac{(1 - S_{\text{rot}}^2) \tau_{\text{rot}}}{1 + (\omega \tau_{\text{rot}})^2} + \frac{S_{\text{rot}}^2 \tau_S}{1 + (\omega \tau_S)^2}. \quad (12)$$

For typical protein concentrations of high-resolution NMR samples (a few mM),  $S_{\text{rot}}^2$  is very low, less than a few percent (31). Hence,  $T_1$  relaxation times are not sensitive to the slow component. However,  $T_2$ s are quite sensitive to it because of the spectral density function at zero frequency:

$$J_t(0) = (1 - S_{\text{rot}}^2) \tau_{\text{rot}} + S_{\text{rot}}^2 \tau_S. \quad (13)$$

Despite the fact that  $S_{\text{rot}}^2 \ll 1$ , the two terms in Eq. 13 are comparable, since  $\tau_S \gg \tau_{\text{rot}}$ . For this reason, using the  $T_1/T_2$  ratio for determination of the tumbling correlation time,  $\tau_{\text{rot}}$ , can provide imprecise results. The higher the concentration, the less correct is the value of  $\tau_{\text{rot}}$  obtained from the  $T_1/T_2$  ratio. Thus, the assumption of a biexponential form of  $C_t(t)$  is a prerequisite for the correct analysis of the relaxation data at different concentrations.

In fitting the temperature dependences of the relaxation times, we assume an Arrhenius dependence of the correlation times,

$$\tau_{S,\text{rot}} = \tau_{S,\text{rot}}(293\text{K}) \exp\left[\frac{E_{S,\text{rot}}}{R} \left(\frac{1}{T} - \frac{1}{293\text{K}}\right)\right], \quad (14)$$

where  $E_{S,\text{rot}}$  is the activation energy of the  $\tau_S/\tau_{\text{rot}}$  correlation times and  $R$  is the universal gas constant. Thus, the fitting parameters in the analysis were two correlation times, two activation energies, the order parameter  $S_{\text{rot}}^2$  (separate sets for each concentration), and the product of the rigid-lattice second moment,  $K_{\text{HH}}$  (which for rigid globular proteins has an approximate value of  $\sim 1.3 \times 10^{10} \text{ s}^{-2}$  (32)), and the order parameter  $S_t^2$ :

$$K_{\text{HH}}^{\text{av}} = S_t^2 K_{\text{HH}}, \quad (15)$$

where  $K_{\text{HH}}^{\text{av}}$  is the motionally averaged second moment of the protein protons; we assume it to be the same for all concentrations. The overall number of the fitting parameters for all four concentrations was 21: five parameters for each concentration (see above) and one parameter ( $K_{\text{HH}}^{\text{av}}$ ) shared between all concentrations. For the fitting, we used Eqs. 2–4, replacing  $J(\omega)$  and  $K_{\text{HH}}$  by  $J_t(\omega)$  and  $K_{\text{HH}}^{\text{av}}$ , respectively. A similar approach was taken previously by Bertini et al. (29), stressing the use of  $S_t^2$  as a qualitative indicator of internal rigidity. For simplicity, we assume  $S_{\text{rot}}^2$  to be temperature-independent and we neglect the distribution of sizes. Strictly speaking, this is not absolutely true, yet it has only a minor effect on the analysis, as demonstrated in Fig. S8.  $K_{\text{HH}}^{\text{av}}$  is also assumed to be temperature-independent. Within the temperature range of our experiments, the temperature dependence of  $K_{\text{HH}}^{\text{av}}$  is rather weak (32,33). Simply assuming a reasonable temperature dependence, however, has practically no influence on the results (see Table S2).

The fitting procedure was based upon a minimization of the root mean-square deviation,

$$\text{RMSD} = \sqrt{\frac{1}{N} \sum_{i=1}^N \left( \frac{T_{\text{sim}} - T_{\text{exp}}}{T_{\text{exp}}} \right)^2}, \quad (16)$$

where  $T_{\text{sim}}$  and  $T_{\text{exp}}$  are the simulated (according to the current set of the fitting parameters) and experimental relaxation times ( $T_{1\rho}$  or  $T_2$ ), respectively, and  $N$  is the number of all relaxation times measured at all temperatures and concentrations. For the minimization procedure, we used the Metropolis algorithm.

## RESULTS

### Translational diffusion and viscosity

Typical examples of the PFG NMR intensity decays (signal intensity versus strength of the pulsed field gradient) are shown in Fig. 1. In the representation  $\log(I)$  versus  $g^2$ , a deviation of the intensity decay from a straight line reflects an SDC dispersion, i.e., a distribution of molecular mass. Fig. 1 likewise demonstrates that this distribution is somewhat wider for a higher concentration; the normalized (relative) RMSDs of the diffusion coefficients from their average, as estimated from a bicomponent decomposition of the decays, are 0.2, 0.2, 0.25, and 0.35 for the concentrations 35, 80, 113, and 185 mg/mL, respectively, in qualitative agreement with DLS data (20). In this analysis we did not quantify the

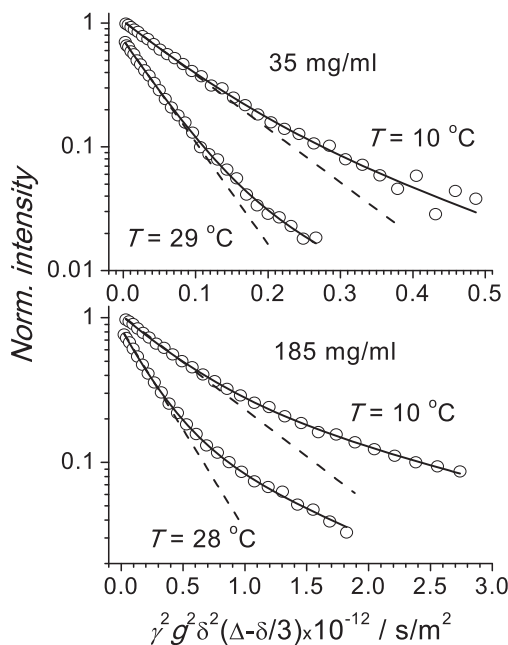


FIGURE 1 Typical examples of diffusion-dependent PFG NMR decays of  $\alpha$ B-crystallin at two different concentrations and two different temperatures. The experimental error corresponds to the size of the symbols in the initial part of the decays. Solid red lines are bicomponent fits of the decays and dashed lines denote the initial slope of the decays corresponding to the mean SDC.

distribution, instead defining the mean SDC, which corresponds to the initial slope of the decay. Practically, we fitted the decay with a sum of two components as a minimal but sufficient model and then calculated the mean SDC as

$$\langle D \rangle = (P_1 \times D_1 + P_2 \times D_2) / (P_1 + P_2), \quad (17)$$

where  $P_{1,2}$  and  $D_{1,2}$  are the intensities and SDCs, respectively, of the two components. The specific values of  $D_1$  and  $D_2$  depend on the weighting factor; these values taken separately have no physical meaning, yet the average diffusion coefficient is well-defined and reliable. Note that the subunit exchange between  $\alpha$ -crystallin oligomers occurs on a timescale of minutes (34); thus, the observed SDC is not the exchange-averaged value of SDCs of oligomers and mono(dimers). In fact, the amount of  $\alpha$ -crystallin mono(dimeric) subunits in solution is very low; otherwise, we would see a corresponding fast component in the PFG intensity decays.

Fig. 2 *a* presents the temperature dependences of mean SDCs at four different concentrations of  $\alpha$ B-crystallin in an Arrhenius representation. It is useful to analyze these data in comparison with lysozyme ( $M = 14.3$  kDa). At acidic pH, lysozyme forms no dimers or oligomers and retains its rigid native structure over a wide range of concentrations and temperatures (35). This comparison shows that the slope of the SDC temperature dependences (i.e., the activation energy of translational diffusion) for  $\alpha$ B-crystallin at all concentrations is quite similar to that of lysozyme. This indicates that at all concentrations, the mean molecular mass of the  $\alpha$ B-crystallin assemblies is independent of temperature. Fig. 2 *b* presents the macroscopic viscosity, along with data for pure  $D_2O$ . In the Arrhenius plot, given the

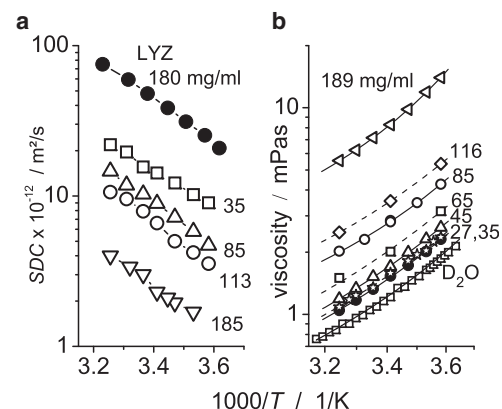


FIGURE 2 Translational self-diffusion coefficients and viscosity of  $\alpha$ B-crystallin. In both graphs, the size of the symbols corresponds to the experimental uncertainty. (a) Temperature dependences of the mean SDCs at four different concentrations of  $\alpha$ B-crystallin (open symbols). For comparison, SDCs of a lysozyme solution (concentration 180 mg/mL, pH 3.5) are shown (solid symbols). (b) Temperature dependence of the viscosity at seven concentrations of  $\alpha$ B-crystallin (symbols), displaying a Vogel-Fulcher relationship (solid lines). The solid line for the  $D_2O$  viscosity represents literature data as recalculated (48) from the viscosity of water (49).

improved data quality of the viscosity measurements compared to those of translational diffusion, a slight curvature of the viscosity can be seen, indicating that the glassy dynamics related to the solvent can be described by a Vogel-Fulcher relationship (36). A slightly stronger deviation from the Arrhenius behavior is observed for the highest concentrations, pointing to the increasing relevance of studying the glassy dynamics of confined/bound water.

At the smallest concentration (35 mg/mL), the protein SDC concentration dependence is weak (37), and thus, intermolecular protein interactions have almost no influence on the SDC. This allows us to estimate the size of  $\alpha$ B-crystallin using the Stokes-Einstein relationship and the experimental viscosity data. This gives a temperature-independent value (see Fig. S7) of  $R_H = (95 \pm 3) \text{ \AA}$ , which exactly matches the value obtained from the DLS experiments by Licinio et al. (20). Note that those authors were studying  $\alpha$ -crystallin from calf lenses, which are oligomers composed of a mixture of  $\alpha$ A- and  $\alpha$ B-crystallins.

### Rotational diffusion

Fig. 3 depicts the relaxation times and fitting curves for different concentrations of  $\alpha$ B-crystallin. Fitting a single-component correlation function for the protein Brownian rotation (i.e., assuming  $S_{\text{rot}}^2 = 0$ ) results in a pronounced mismatch, confirming the invalidity of the one-component model. The fitting results are summarized in Table 1.

The absolute value of the rotational correlation time for the lowest concentration investigated in this study, 35 mg/mL, is  $0.9 \mu\text{s}$ . Applying the Stokes-Einstein-Debye law with the experimentally determined viscosity of this sample, these values correspond to an  $\alpha$ B-crystallin radius

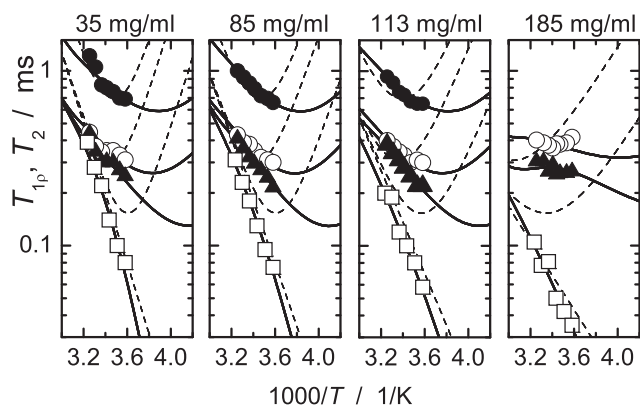


FIGURE 3 Proton  $T_{1p}$  and  $T_2$  for  $\alpha$ B-crystallin solutions at different concentrations. The experimental error corresponds to the size of the symbols. Information is provided for relaxation times  $T_2$  (open squares),  $T_{1p}$  at the spin-lock frequency, 20 kHz (solid triangles),  $T_{1p}$  at 40 kHz (open circles), and off-resonance  $T_{1p}$  at 60 kHz (solid circles) (the latter parameter was not measured for 185 mg/mL). Solid lines show the best fits for double-exponential correlation functions, and dashed lines correspond to the best fit assuming  $S_{\text{rot}}^2 = 0$ , i.e., a single-exponential correlation function.

of  $\sim 82 \text{ \AA}$ . This value is somewhat less than  $95 \text{ \AA}$ , as obtained from the translational diffusion data (see above). This may indicate that the Stokes-Einstein-Debye law does not hold for the rotational diffusion, and that the macroscopic viscosity should not be used to determine rotational correlation time (see below). If, instead of the solution viscosity, the viscosity of a pure solvent ( $D_2O$ ) is used in the Stokes-Einstein-Debye equation, then the calculated  $\alpha$ B-crystallin radius reaches  $89 \text{ \AA}$ . Given the overall experimental accuracy, the discrepancy between  $89 \text{ \AA}$  and  $95 \text{ \AA}$  can be considered as negligible. Note that the activation energy of the Brownian tumbling  $E_{\text{rot}}$  corresponds quite well to that of the viscous flow of pure water, which is  $\sim 19 \text{ kJ/mol}$  (38). The decrease of  $E_{\text{rot}}$  to  $10 \text{ kJ/mol}$  at a concentration of  $185 \text{ mg/mL}$  is obviously an apparent effect associated with the increased distribution of molecular masses and the probably more complex form of the  $C_i(t)$ .

## DISCUSSION

### The impact of crowding: rotational diffusion is less hindered than translational diffusion

Since the rotational diffusion is described by the two-component overall tumbling RACF  $C_i(t)$ , we define, as in the case of translational diffusion, a mean rotational diffusion rate  $\langle R_{\text{rot}} \rangle$  equal to the initial slope of the rotational correlation function:

$$\langle R_{\text{rot}} \rangle = \frac{1 - S_{\text{rot}}^2}{\tau_{\text{rot}}} + \frac{S_{\text{rot}}^2}{\tau_S}. \quad (18)$$

Since  $S_{\text{rot}}^2$  and  $\tau_S$  are poorly defined at low concentrations there is a certain ambiguity in defining the rotational diffusion rate at low concentrations. However, since the second term in Eq. 18 is much smaller than the first, this ambiguity is obviously negligible.

Fig. 4 presents the central result of this work, the comparative retardation of translational and rotational diffusion relative to macroscopic viscosity with increasing concentration. It is seen that the trend of the translational diffusion nicely corresponds to that of viscosity, which confirms previous findings by Licinio and Delaye (18). Thus, the Stokes-Einstein law appears to be valid even at high concentrations. This in turn shows that the mean size of  $\alpha$ B-crystallin under our conditions does not depend on concentration.

On the other hand, these results clearly demonstrate a significant difference between translational and rotational diffusion of  $\alpha$ B-crystallin at high concentrations, far beyond all the assumptions and uncertainties of the data analysis. We also stress that fitting the relaxation data with a fixed ratio of the correlation times,  $\tau_{\text{rot}}$ , at different concentrations after the known increase in viscosity (and thus the slow-down of translational diffusion) results in a strong mismatch with the experimental data (see Fig. S10).

**TABLE 1** Dynamic parameters obtained from the data fitting

$c/\text{mg/mL}$	$\tau_{\text{rot}}/\mu\text{s}$ at $20^\circ\text{C}$	$S_{\text{rot}}^2$	$\tau_{\text{S}}/\mu\text{s}$ at $20^\circ\text{C}$	$S_{\text{rot}}^2 \tau_{\text{S}}/\mu\text{s}$ at $20^\circ\text{C}$	$E_{\text{rot}}/\text{kJ/mol}$	$E_{\text{S}}/\text{kJ/mol}$
35	$0.90 \pm 0.02$	$<0.03$	$>30$	$0.64 \pm 0.02$	$16 \pm 1$	$66 \pm 2$
85	$0.96 \pm 0.02$	$<0.03$	$>30$	$0.83 \pm 0.02$	$18 \pm 1$	$51 \pm 2$
113	$1.03 \pm 0.03$	$<0.03$	$>80$	$1.36 \pm 0.03$	$17 \pm 1$	$40 \pm 2$
185	$1.04 \pm 0.03$	$0.22 \pm 0.02$	$17 \pm 1$	$3.70 \pm 0.06$	$10 \pm 1$	$30 \pm 1$

$K_{\text{HH}}^{\text{av}}$  (Eq. 15) was found to be  $(4.2 \pm 0.2) \times 10^9 \text{ s}^{-2}$  by a shared fit of all data sets. Since  $S_{\text{rot}}^2$  is very small, the parameters  $S_{\text{rot}}^2$  and  $\tau_{\text{S}}$  cannot be determined separately at low concentrations; only the product  $S_{\text{rot}}^2 \tau_{\text{S}}$  could be reliably obtained from the fitting. For details, see Krushelnitsky (31).

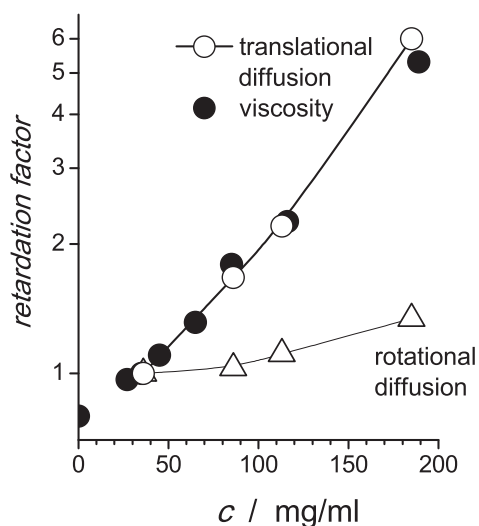
Interestingly, a similar conclusion regarding the limited applicability of the Stokes-Einstein-Debye law for high  $\alpha$ B-crystallin concentrations can be deduced from recently published data on Brownian tumbling of this protein obtained by field-cycling relaxometry of the water protons (39). At a single concentration of 100 mg/mL,  $T = 25^\circ\text{C}$  in 80%  $\text{H}_2\text{O}$  and 20% glycerol solvent, the correlation time,  $\tau_{\text{rot}}$ , of  $\alpha$ B-crystallin was found to be 1.4  $\mu\text{s}$ . Using our viscosity data for this protein concentration and temperature, and taking into account the correction factor for viscosity between  $\text{D}_2\text{O}$  and 80%  $\text{H}_2\text{O}$  and 20% glycerol solvents, we estimated the apparent radius of  $\alpha$ B-crystallin to be 68  $\text{\AA}$ . This value is obviously too small, in accordance with our finding that rotational diffusion is less hindered than expected by the increase of viscosity.

Similar concentration dependences of rotational and translational protein diffusion have been reported previously (4), but this is the first time, to our knowledge, that such a large quantitative difference has been observed in a protein system. The effect of less hindered rotations compared to

the translational self-diffusion goes far beyond a pure viscosity effect resulting from the difference in local microviscosity around the protein and the bulk viscosity. Increasing the bulk viscosity with ethylene glycol by a factor of 6 (corresponding to a retardation factor of 6 in Fig. 4) results in retardation of translation diffusion and rotational diffusion by factors of 5.5 and 4, respectively, for a small globular protein (40). The less hindered rotation at high protein concentrations can be easily understood in terms of the cage effect, which is well known for the case of spherical colloids (41,42). For translational diffusion, each probe molecule needs to escape a cage formed by the surrounding particles ( $\alpha$ -relaxation) and thus has to interact with its neighboring proteins, which represent obstacles to translational motion. For rotational diffusion, proteins may rotate rather freely within a cage ( $\beta$ -relaxation); hence, intermolecular protein interactions can be expected to have an appreciably smaller effect on it. Note that the effect of decoupling between translational and rotational diffusion has been observed not only in experimental studies, but also in numerical simulations of protein diffusion at high concentrations (43). However, we refrain from extrapolating the findings of this work to other proteins and experimental conditions. The acquired data are obviously not sufficient to make generalizations, and more experimental work is required to further advance our knowledge in this area.

### Fractal structure: $\alpha$ B-crystallin behaves like a normal globular protein

Further information on  $\alpha$ B-crystallin properties can be obtained by comparing the absolute values of the diffusion constants with those of other proteins. Although such data have been published, at least for translational diffusion (see, e.g., the work of Delaye and colleagues (16,18,20)), the comparison has apparently not yet been made. The SDC is inversely proportional to the linear size of the Brownian particle, whereas the rotational correlation time is proportional to its volume. To minimize the influence of intermolecular protein interactions, we compared the diffusion parameters only for the dilute  $\alpha$ B-crystallin solution. If the average protein density is the same for proteins of different molecular mass,  $M$ , one might expect that  $\text{SDC} \sim M^{-0.33}$  and  $\tau_{\text{rot}} \sim M$ , but this is not the case. Computer analysis of a large number of 3D protein structures



**FIGURE 4** Retardation of the translational (*open circles*) and rotational diffusion (*open triangles*) as a function of  $\alpha$ B-crystallin concentration as compared to the normalized macroscopic viscosity (*solid circles*). The retardation factor was defined as the ratio of the translational (rotational) diffusion rate to the value at 35 mg/mL, taken as a reference for the higher concentrations. The viscosity was normalized in the same way. The size of the symbols reflects the experimental error, and the solid line simply guides the eye.

(44,45) demonstrates that  $V \sim R^d$ , where  $V$  is the van der Waals (or solvent-accessible) volume of the protein directly proportional to the molecular mass, which is connected to the linear size,  $R$  (more specifically, the radius of gyration) of the protein molecule by a scaling exponent,  $d$ , of  $\sim 2.5$ . Such a relation between size and volume reflects a fractal nature of the protein packing, which has generated increased interest over the last few years (see the review by Banerji and Ghos (46)). The fractal dimension  $d < 3$  indicates that the protein density decreases with increasing  $M$  (47). If  $V \sim M$ , then the SDC is  $\sim M^{-1/d}$  and  $\tau_{\text{rot}} \sim M^{3/d}$ .

Fig. 5 *a* shows a collection of data from the literature on SDCs as a function of  $M$  for many proteins and includes the  $\alpha$ B-crystallin results from this work. Two important points must be mentioned. First, the slope of the SDC versus  $M$  dependence matches reasonably well the value of  $d$  obtained by Liang and Dill (44). To our knowledge, this is the first experimental confirmation of this fractal dependence based on diffusion data, reporting on the hydrodynamic radius  $R_h$ . Second, the  $\alpha$ B-crystallin SDC is located close to this line, which indicates that it has no specific features, as compared

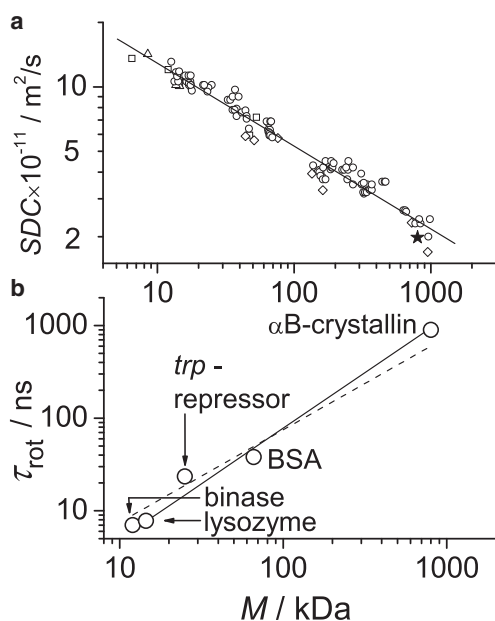


FIGURE 5 (a) SDCs for different proteins as a function of  $M$  at 20°C. Literature data are either taken directly from Tyn and Gusek (50) and Ilyina et al. (51) (open circles and triangles) or recalculated from the hydrodynamic radii reported in Wilkins et al. (52) and Armstrong et al. (53) (open squares and diamonds). Proteins too anisotropic in shape or that are intrinsically disordered were not taken into account. The  $\alpha$ B-crystallin SDC as obtained in this study is indicated by the solid star. The  $\alpha$ B-crystallin SDC was measured at a concentration of 10 mg/mL and multiplied by 1.25 to account for the viscosity difference between  $\text{H}_2\text{O}$  (literature data of  $\text{H}_2\text{O}$  solutions) and  $\text{D}_2\text{O}$  (this work). The solid line is a best fit to the data, with a slope (power-law exponent) of  $0.39 \pm 0.03$ . (b) Rotational correlation time  $\tau_{\text{rot}}$  at 20°C for five different proteins as a function of molecular weight. The solid line presents the dependence  $\tau_{\text{rot}} \sim M^{1.2}$ . For comparison, the dashed line shows the dependence  $\tau_{\text{rot}} \sim M$ .

to other globular proteins, and is just as compact as might be expected based on the fractal scaling law,  $R_h \sim M^{1/d}$ , and its high molecular mass.

A similar dependence can likewise be plotted for rotational diffusion. Many studies have been published on protein dynamics in solution over the last 20 to 30 years. However, in most of these, the rotational correlation time was determined from the NMR  $T_1/T_2$  relaxation-time ratio assuming only a single-component RACF or, at best, a more complex form of it accounting for the anisotropic shape of the protein. We again stress that this is quantitatively not correct (see above). The amplitude of the slow component of the RACF depends on many parameters (concentration, ionic strength, pH, and electrostatic properties of a protein) and hence is different for different experiments. This induces a spread of  $\tau_{\text{rot}}$  values that makes it difficult to reliably define the power-law exponent of  $M$  (see Fig. S11). Therefore, for comparison, we took the data of only four proteins, binase, lysozyme, *trp*-repressor, and bovine serum albumin, as described in Krushelnitsky (31). In that study,  $\tau_{\text{rot}}$  was determined according to the same protocol as in this work, so the correlation times can be compared directly.

Fig. 5 *b* presents  $\tau_{\text{rot}}$  as a function of  $M$  for five proteins, including  $\alpha$ B-crystallin. Despite the poor statistics, it can be clearly seen that rotational diffusion also confirms the findings of Liang and Dill (44). The largest deviation from the solid line in Fig. 5 *b* is observed for *trp*-repressor. This can be explained by the fact that *trp*-repressor is a symmetric dimer with two long (12 residues each) unstructured chains exposed to the solvent. Thus, the *trp*-repressor is not a completely rigid protein, and an apparently increased  $\tau_{\text{rot}}$  is easily understood. Note that the rotational diffusion of  $\alpha$ B-crystallin again reveals no evident specificity in comparison with other globular proteins.

## CONCLUSIONS

In this study, we have provided an accurate determination of the translational and rotational diffusion of  $\alpha$ B-crystallin over a wide range of concentrations. Our data allowed us to draw three important conclusions. First, our main finding was that upon increasing the protein concentration, the translational diffusion of  $\alpha$ B-crystallin nicely followed the trend measured for the inverse solution viscosity, whereas the rotational diffusion was found to be affected by the concentration increase to a much smaller extent. This could be explained on the basis of the cage effect typical for spherical colloids. The temperature dependence of all observables was found to be largely governed by the flow activation energy of pure water, with deviations visible only at the largest concentrations. Second, despite its large size and oligomeric structure,  $\alpha$ B-crystallin in dilute solution behaves like a normal rigid globular protein, showing no specificity in Brownian dynamics compared to other, even much smaller,

proteins. Third, both the translational and rotational diffusion data (reporting on the hydrodynamic radius,  $R_h$ ) confirm the fractal scaling law,  $V \sim M \sim R^d$ , with  $d \sim 2.5$  instead of  $d \sim 3$  for a variety of protein structures of different size,  $R$ . This finding is in agreement with previous statistical analyses of protein packing density.

The methodological approach presented here, in particular addressing the autocorrelation function of the overall protein tumbling by a bimodal analysis of NMR relaxation times measured at different frequencies and temperatures, provides an efficient and reliable tool for studying the effect of crowding on Brownian dynamics. Application of this approach to the  $\alpha$ B-crystallin study enabled a qualitative step forward in the description of protein mobility at high concentrations. We expect that the use of this approach for other proteins and protein mixtures will help in constructing a detailed and consistent general picture of protein dynamics under crowding conditions.

## SUPPORTING MATERIAL

Supporting Materials and Methods, eleven figures, and three tables are available at [http://www.biophysj.org/biophysj/supplemental/S0006-3495\(14\)03070-7](http://www.biophysj.org/biophysj/supplemental/S0006-3495(14)03070-7).

## ACKNOWLEDGMENTS

We thank Wilbert Boelens for the plasmid for  $\alpha$ B crystallin and Qi Zhang for the expression clone.

Funding for this work was provided by the Deutsche Forschungsgemeinschaft (DFG) in the framework of the collaborative research center SFB-TRR 102 (project A08). We also acknowledge significant investments in our NMR facility by the European Regional Development Fund (ERDF) of the European Union.

## REFERENCES

- Zimmerman, S. B., and A. P. Minton. 1993. Macromolecular crowding: biochemical, biophysical, and physiological consequences. *Annu. Rev. Biophys. Biomol. Struct.* 22:27–65.
- Ellis, R. J. 2001. Macromolecular crowding: obvious but underappreciated. *Trends Biochem. Sci.* 26:597–604.
- Wang, Y., C. Li, and G. J. Pielak. 2010. Effects of proteins on protein diffusion. *J. Am. Chem. Soc.* 132:9392–9397.
- Zorrilla, S., M. A. Hink, ..., M. P. Lillo. 2007. Translational and rotational motions of proteins in a protein crowded environment. *Biophys. Chem.* 125:298–305.
- Delaye, M., and A. Tardieu. 1983. Short-range order of crystallin proteins accounts for eye lens transparency. *Nature.* 302:415–417.
- Bloemendal, H. 1977. The vertebrate eye lens. *Science.* 197:127–138.
- Bloemendal, H., W. de Jong, ..., A. Tardieu. 2004. Ageing and vision: structure, stability and function of lens crystallins. *Prog. Biophys. Mol. Biol.* 86:407–485.
- Pierscionek, B., and R. C. Augusteyn. 1988. Protein distribution patterns in concentric layers from single bovine lenses: changes with development and ageing. *Curr. Eye Res.* 7:11–23.
- Carver, J. A. 1999. Probing the structure and interactions of crystallin proteins by NMR spectroscopy. *Prog. Retin. Eye Res.* 18:431–462.
- Horwitz, J. 1992. Alpha-crystallin can function as a molecular chaperone. *Proc. Natl. Acad. Sci. USA.* 89:10449–10453.
- Derham, B. K., and J. J. Harding. 1999.  $\alpha$ -Crystallin as a molecular chaperone. *Prog. Retin. Eye Res.* 18:463–509.
- Narberhaus, F. 2002.  $\alpha$ -Crystallin-type heat shock proteins: socializing minichaperones in the context of a multichaperone network. *Microbiol. Mol. Biol. Rev.* 66:64–93.
- Horwitz, J. 2003.  $\alpha$ -Crystallin. *Exp. Eye Res.* 76:145–153.
- Augusteyn, R. C. 2004.  $\alpha$ -Crystallin: a review of its structure and function. *Clin. Exp. Optom.* 87:356–366.
- Andley, U. P. 2007. Crystallins in the eye: function and pathology. *Prog. Retin. Eye Res.* 26:78–98.
- Delaye, M., and A. Gromiec. 1983. Mutual diffusion of crystallin proteins at finite concentrations: a light-scattering study. *Biopolymers.* 22:1203–1221.
- Tardieu, A., D. Laporte, and M. Delaye. 1987. Colloidal dispersions of  $\alpha$ -crystallin proteins. 1. Small-angle x-ray analysis of the dispersion structure. *J. Phys. (Paris).* 48:1207–1215.
- Licinio, P., and M. Delaye. 1988. Mutual and self-diffusion in concentrated  $\alpha$ -crystallin protein dispersion. A dynamic light-scattering study. *J. Phys. (Paris).* 49:975–981.
- Haley, D. A., J. Horwitz, and P. L. Stewart. 1998. The small heat-shock protein,  $\alpha$ B-crystallin, has a variable quaternary structure. *J. Mol. Biol.* 277:27–35.
- Licinio, P., M. Delaye, ..., L. Leger. 1987. Colloidal dispersions of  $\alpha$ -crystallin proteins. 2. Dynamics: a maximum-entropy analysis of photon-correlation spectroscopy data. *J. Phys. (Paris).* 48:1217–1223.
- Mainz, A., S. Jehle, ..., B. Reif. 2009. Large protein complexes with extreme rotational correlation times investigated in solution by magic-angle-spinning NMR spectroscopy. *J. Am. Chem. Soc.* 131:15968–15969.
- Luz, Z., and S. Meiboom. 1963. Nuclear magnetic resonance study of protolysis of trimethylammonium ion in aqueous solution: order of reaction with respect to solvent. *J. Chem. Phys.* 39:366–370.
- Stejskal, E. O., and J. E. Tanner. 1965. Spin diffusion measurements: spin echoes in the presence of a time-dependent field gradient. *J. Chem. Phys.* 42:288–292.
- Kimmich, R., and E. Anordo. 2004. Field-cycling NMR relaxometry. *Prog. NMR Spectrosc.* 44:257–320.
- Jones, G. P. 1966. Spin-lattice relaxation in rotating frame: weak-collision case. *Phys. Rev.* 148:332–335.
- Krushelnitsky, A., and D. Reichert. 2004. Response of lysozyme internal dynamics to hydration probed by C-13 and H-1 solid-state NMR relaxation. *Appl. Magn. Reson.* 27:501–518.
- Goldman, M. 2001. Formal theory of spin-lattice relaxation. *J. Magn. Reson.* 149:160–187.
- Lipari, G., and A. Szabo. 1982. Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 1. Theory and range of validity. *J. Am. Chem. Soc.* 104:4546–4559.
- Bertini, I., Y. K. Gupta, ..., H. Schwalbe. 2005. NMR spectroscopic detection of protein protons and longitudinal relaxation rates between 0.01 and 50 MHz. *Angew. Chem. Int. Ed.* 44:2223–2225.
- Krushelnitsky, A. G., and V. D. Fedotov. 1993. Overall and internal protein dynamics in solution studied by the nonselective proton relaxation. *J. Biomol. Struct. Dyn.* 11:121–141.
- Krushelnitsky, A. 2006. Intermolecular electrostatic interactions and Brownian tumbling in protein solutions. *Phys. Chem. Chem. Phys.* 8:2117–2128.
- Krushelnitsky, A. G., V. D. Fedotov, ..., J. Straka. 1996. Dynamic structure of proteins in solid state.  $^1\text{H}$  and  $^{13}\text{C}$  NMR relaxation study. *J. Biomol. Struct. Dyn.* 14:211–224.
- Bleas, D. J., and S. S. Danyluk. 1968. Proton wide-line nuclear magnetic resonance spectra of hydrated proteins. *Biochim. Biophys. Acta.* 154:17–27.

34. Bova, M. P., L.-L. Ding, ..., B. K.-K. Fung. 1997. Subunit exchange of  $\alpha$ A-crystallin. *J. Biol. Chem.* 272:29511–29517.
35. Sedgwick, H., K. Kroy, ..., W. C. Poon. 2005. Non-equilibrium behavior of sticky colloidal particles: beads, clusters and gels. *Eur Phys J E Soft Matter.* 16:77–80.
36. Strobl, G. 2007. *The Physics of Polymers*, 3rd ed. Springer, Berlin-Heidelberg.
37. Nesmelova, I. V., V. D. Skirda, and V. D. Fedotov. 2002. Generalized concentration dependence of globular protein self-diffusion coefficients in aqueous solutions. *Biopolymers.* 63:132–140.
38. Horne, R. A., R. A. Courant, ..., F. F. Margosian. 1965. The activation energy of viscous flow of pure water and sea water in the temperature region of maximum density. *J. Phys. Chem.* 69:3988–3991.
39. Ravera, E., G. Parigi, ..., C. Luchinat. 2013. Experimental determination of microsecond reorientation correlation times in protein solutions. *J. Phys. Chem. B.* 117:3548–3553.
40. Zeeb, M., M. H. Jacob, ..., J. Balbach. 2003.  $^{15}\text{N}$  relaxation study of the cold shock protein CspB at various solvent viscosities. *Biomol. NMR.* 27:221–234.
41. Doliwa, B., and A. Heuer. 1998. Cage effect, local anisotropies, and dynamic heterogeneities at the glass transition: a computer study of hard spheres. *Phys. Rev. Lett.* 80:4915–4918.
42. Pusey, P. N. 2008. Colloidal glasses. *J. Phys. Condens. Matter.* 20:494202.
43. Mereghetti, P., and R. C. Wade. 2012. Atomic detail Brownian dynamics simulations of concentrated protein solutions with a mean field treatment of hydrodynamic interactions. *J. Phys. Chem. B.* 116:8523–8533.
44. Liang, J., and K. A. Dill. 2001. Are proteins well-packed? *Biophys. J.* 81:751–766.
45. Moret, M. A., M. C. Santana, ..., G. F. Zebende. 2006. Protein chain packing and percolation threshold. *Physica A.* 361:250–254.
46. Banerji, A., and I. Ghosh. 2011. Fractal symmetry of protein interior: what have we learned? *Cell. Mol. Life Sci.* 68:2711–2737.
47. Fischer, H., I. Polikarpov, and A. F. Craievich. 2004. Average protein density is a molecular-weight-dependent function. *Protein Sci.* 13:2825–2828.
48. Harris, K. R., and L. A. Woolf. 2004. Temperature and volume dependence of the viscosity of water and heavy water at low temperatures. *J. Chem. Eng. Data.* 49:1064–1069.
49. Kestin, J., M. Sokolov, and W. A. Wakeham. 1978. Viscosity of liquid water in range  $-8^\circ\text{C}$  to  $150^\circ\text{C}$ . *J. Phys. Chem. Ref. Data.* 7:941–948.
50. Tyn, M. T., and T. W. Gusek. 1990. Prediction of diffusion coefficients of proteins. *Biotechnol. Bioeng.* 35:327–338.
51. Ilyina, E., V. Roongta, ..., K. H. Mayo. 1997. A pulsed-field gradient NMR study of bovine pancreatic trypsin inhibitor self-association. *Biochemistry.* 36:3383–3388.
52. Wilkins, D. K., S. B. Grimshaw, ..., L. J. Smith. 1999. Hydrodynamic radii of native and denatured proteins measured by pulse field gradient NMR techniques. *Biochemistry.* 38:16424–16431.
53. Armstrong, J. K., R. B. Wenby, ..., T. C. Fisher. 2004. The hydrodynamic radii of macromolecules and their effect on red blood cell aggregation. *Biophys. J.* 87:4259–4270.





## 4.2. The “long tail” of the protein tumbling correlation function: observation by $^1\text{H}$ NMR relaxometry in a wide frequency and concentration range

In the previous article it was shown that rotational diffusion of  $\alpha\text{B}$ -crystallin is strongly decoupled from long-time translational diffusion in crowded conditions and that spin relaxation rates cannot be properly fit using a single exponential RACF. The results stress the importance of accounting for the non-exponential behavior of the RACF. To be noted, an intermediate range order structure has been depicted for LYZ solutions, with the lifetime of such inter-molecular ordering being longer than, but still comparable to, the time required by a LYZ monomer to diffuse over a distance of its own size [Liu et al., 2011]. This finding strongly corroborates the concept of rotational diffusion being affected by intermolecular alignment effects, and is in line with the concept of electro-kinetic steering.

Sensitivity to the “slow component” of rotational diffusion was till now only provided via  $R_2$  relaxation rates. Transverse relaxation rates relate to the zero-frequency limit of the spectral density,  $J(0) = \int_0^\infty C_r(t) dt$ , but cannot address the actual frequency dependence of the “slow tail” of rotational diffusion. Thus, the behavior of the RACF at times in-between the time scale of usual Brownian tumbling and the zero-frequency limit has remained unresolved thus far.

The following study aims to elucidate the non-exponential behavior of the RACF. For this purpose, FC NMR on LYZ and BSA solutions were combined with  $R_{1\rho}$  and  $R_2$  measurements on the same samples. LYZ and BSA are both much smaller than the large  $\alpha\text{B}$ -crystallin complexes and consequentially have smaller  $\tau_{\text{rot}}$  values. Then,  $R_1$  measurements at low Larmor frequencies as well as  $R_{1\rho,2}$  data are sensitive to the RACF at time scales beyond the usual Brownian tumbling and the long-time decay of the RACF can be evaluated. The data presented in this paper will be used in paper #3 to compare rotational with translational diffusion rates.

*Author contributions:* A.K. and K.S. designed research; M.R., M.H., and A.K. performed NMR measurements; M.R. analyzed the NMR data; S.L. provided the size-exclusion chromatography and blue-native PAGE data; M.O. discussed X-ray scattering findings; M.R., M.O., E.R., J.B., A.K. and K.S. interpreted the findings; A.K., M.R. and K.S. wrote the paper with refinements by all coauthors.

The following article [Roos et al., J. Biomol. NMR, 63: 403-415, 2015] has been published under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>) that permits unrestricted use, distribution, and reproduction in any medium under specification of the authors (see the article) and the source (J. Biomol. NMR, DOI: 10.1007/s10858-015-0001-1). The link back to the article on the publisher's website is <http://dx.doi.org/10.1007/s10858-015-0001-1> .  
No changes were made.

# The “long tail” of the protein tumbling correlation function: observation by $^1\text{H}$ NMR relaxometry in a wide frequency and concentration range

Matthias Roos<sup>1</sup> · Marius Hofmann<sup>2</sup> · Susanne Link<sup>1</sup> · Maria Ott<sup>1</sup> · Jochen Balbach<sup>1</sup> · Ernst Rössler<sup>2</sup> · Kay Saalwächter<sup>1</sup> · Alexey Krushelnitsky<sup>1</sup>

Received: 5 August 2015 / Accepted: 11 November 2015 / Published online: 18 November 2015  
© The Author(s) 2015. This article is published with open access at Springerlink.com

**Abstract** Inter-protein interactions in solution affect the auto-correlation function of Brownian tumbling not only in terms of a simple increase of the correlation time, they also lead to the appearance of a weak slow component (“long tail”) of the correlation function due to a slowly changing local anisotropy of the microenvironment. The conventional protocol of correlation time estimation from the relaxation rate ratio  $R_1/R_2$  assumes a single-component tumbling correlation function, and thus can provide incorrect results as soon as the “long tail” is of relevance. This effect, however, has been underestimated in many instances. In this work we present a detailed systematic study of the tumbling correlation function of two proteins, lysozyme and bovine serum albumin, at different concentrations and temperatures using proton field-cycling relaxometry combined with  $R_{1\rho}$  and  $R_2$  measurements. Unlike high-field NMR relaxation methods, these techniques enable a detailed study of dynamics on a time scale longer than the normal protein tumbling correlation time and, thus, a reliable estimate of the parameters of the “long tail”. In this work we analyze the concentration dependence of the intensity and correlation time of the slow

component and perform simulations of high-field  $^{15}\text{N}$  NMR relaxation data demonstrating the importance of taking the “long tail” in the analysis into account.

**Keywords** Inter-protein interactions · Brownian tumbling · Field-cycling · Relaxation · Correlation function

## Introduction

Overall Brownian tumbling of proteins in solution is an important issue in many biophysical and biochemical studies, and may provide information on the size and shape of the protein under investigation, as well as on inter-molecular interactions. Starting from the pioneering works by Kay et al. (1989) and Clore et al. (1990), a long series of papers has been published that deal with high-resolution NMR relaxation studies of internal dynamics of proteins in solution. Almost all these studies utilized the well known model-free approach (Lipari and Szabo 1982a, b) for relaxation times analysis. According to the simplest form of this approach, the normalized second-order reorientational correlation function reads

$$C(t) = \exp(-t/\tau_{\text{rot}}) [S_{\text{int}}^2 + (1 - S_{\text{int}}^2) \exp(-t/\tau_{\text{int}})], \quad (1)$$

where  $\tau_{\text{rot}}$  is the correlation time of the overall protein Brownian rotation (tumbling), and  $S_{\text{int}}^2$  and  $\tau_{\text{int}}$  are the order parameter and the correlation time of the internal motion, respectively. More sophisticated protocols take into account different components of the internal mobility, the non-spherical shape of the protein molecule leading to anisotropic overall motion and the contribution of chemical exchange to the relaxation rate  $R_2$ , as surveyed in a number of reviews (Daragan and Mayo 1997; Korzhnev et al. 2001; Palmer 2001; Boehr et al. 2006; Kleckner and Foster 2011;

**Electronic supplementary material** The online version of this article (doi:10.1007/s10858-015-0001-1) contains supplementary material, which is available to authorized users.

✉ Kay Saalwächter  
kay.saalwaechter@physik.uni-halle.de

✉ Alexey Krushelnitsky  
krushelnitsky@physik.uni-halle.de

<sup>1</sup> Institut für Physik, Martin-Luther-Universität Halle-Wittenberg, Betty-Heimann-Str. 7, 06120 Halle, Germany

<sup>2</sup> Universität Bayreuth, Lehrstuhl Experimentalphysik II, Universitätsstr. 30, 95440 Bayreuth, Germany

Ishima 2012; Saito 2014). The main target of those studies was providing site-specific information on internal motions. However, as seen from Eq. (1), correct estimation of the internal dynamics parameters ( $S_{\text{int}}^2$  and  $\tau_{\text{int}}$ ) is impossible without an exact determination of the molecule's overall rotational correlation time  $\tau_{\text{rot}}$ . To a first approximation  $\tau_{\text{rot}}$  can be obtained from the ratio of the spin–lattice and spin–spin relaxation rates,  $R_1/R_2$ , measured for the most rigid residues of a protein undergoing (almost) no slow internal mobility (Kay et al. 1989; Clore et al. 1990). Then,  $\tau_{\text{rot}}$  is usually refined during the global fit of all the data. This approach assumes a free rotation of the molecule to be described in terms of a single exponential, or at most a few exponentials describing the non-spherical shape of a protein molecule. In our opinion, this treatment can lead to imprecise results, in particular at high protein concentrations.

In a series of preceding papers (Krushelnitsky and Fedotov 1993; Ermolina et al. 1993; Ermakova et al. 2002; Krushelnitsky 2006; Roos et al. 2015) it has been shown that inter-protein long-range electrostatic interactions not just increase  $\tau_{\text{rot}}$ , but cause the appearance of a weak slowly decaying component of the tumbling correlation function (“long tail”). At protein concentrations of a few mM (as typical for protein NMR samples) the inter-protein distances are comparable with protein's size, and the energy of mutual electrostatic steering is comparable with the thermal energy  $kT$  (Ermolina et al. 1993). The micro-surrounding around each protein induces a local anisotropy of the “normal” Brownian tumbling. The lifetime of this local anisotropic configuration of proteins is controlled by micro-environmental fluctuations primarily mediated by the translational motion of proteins in respect to each other, rendering this lifetime considerably longer than  $\tau_{\text{rot}}$ . Thus, the correlation function (1) can be better approximated by (Krushelnitsky 2006)

$$C(t) = \exp(-t/\tau_S) [S_{\text{rot}}^2 + (1 - S_{\text{rot}}^2) \exp(-t/\tau_{\text{rot}})] \cdot [S_{\text{int}}^2 + (1 - S_{\text{int}}^2) \exp(-t/\tau_{\text{int}})], \quad (2)$$

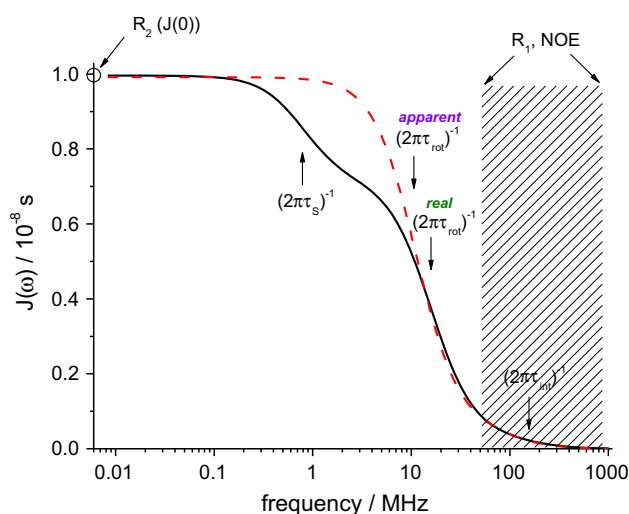
where  $S_{\text{rot}}^2$  is the order parameter of the local Brownian rotation anisotropy and  $\tau_S$  is the slow correlation time characterizing the lifetime of this anisotropy. Note that on a long time scale, protein rotation remains fully isotropic, and  $S_{\text{rot}}^2$  characterizes the Brownian rotation anisotropy solely on a time scale of  $\tau_{\text{rot}}$ . As long as the protein concentration is not very high,  $S_{\text{rot}}^2$  is usually very small, i.e. values around 1 % or even less. At first glance, this practically negligible component should have no significant effect on the relaxation rates. However, one has to keep in mind that the relaxation rates are proportional to frequency-dependent values of the spectral density  $J(\omega)$ , which is the Fourier transform of  $C(t)$ . The effect of the

“long tail” on relaxation rates is demonstrated in Fig. 1, where it is depicted that high-field relaxation measurements are not affected by the slow component. In contrast,  $R_2$  is proportional to the spectral density function at zero frequency  $J(0)$ , or equivalently, to the time integral over the correlation function  $\int_0^\infty C(t) dt$ :

$$J(0) \sim (1 - S_{\text{rot}}^2) \tau_{\text{rot}} + S_{\text{rot}}^2 \tau_S. \quad (3)$$

Since  $\tau_{\text{rot}} \ll \tau_S$ , the two terms in Eq. (3) may be comparable to each other in spite of  $S_{\text{rot}}^2 \ll 1$ . Thus,  $R_2$  is appreciably affected by the “long tail” of the tumbling correlation function, such that the standard protocol of high-field NMR relaxation rates analysis may produce imprecise results when applied to protein solutions that are not highly diluted, see Fig. 1.

The “long tail” can hardly be recognized in the analysis of a conventional set of high-field relaxation parameters ( $R_1$ ,  $R_2$  and  $NOE$  measured at several resonance frequencies, usually from 500 to 900 MHz for protons), meaning that these data can always be well fitted assuming the standard approach, as we demonstrate below. In fact, the slow



**Fig. 1** Representation of the spectral density function as directly sampled by relaxation parameters for the example of  $^{15}\text{N}$ ( $^{-1}\text{H}$ ) relaxation. Three dispersions are relevant, corresponding to three modes of motion, see Eq. (2); the inflection points of the dispersions corresponding to the condition  $\omega\tau = 1$  are marked by arrows. The hatched area marks the frequency range sampled by  $R_1$ 's and  $NOE$ 's measured at the  $^1\text{H}$  resonance frequencies from 500 to 800 MHz.  $R_2$  provides the value of the low-frequency limit of the spectral density. The frequencies in-between are not accessible by high-field relaxation measurements. The “long tail” dispersion is located right in this gap. The dashed line indicates the “apparent” spectral density as obtained from the relaxation data neglecting the impact of the “long tail”. The “real” and the “apparent” spectral densities were simulated according to the dynamic parameters presented for the “mobile” residue, Fig. 7 at  $S_{\text{rot}}^2\tau_S = 4$  ns (see details in the final part of the paper)

component can hardly be seen in the analysis since these measurements do not sample the low frequency range of  $J(\omega)$ , as clearly shown in Fig. 1. When, however, much lower frequencies in high resolution  $^{15}\text{N}$  relaxation experiments are sampled using a field-shuttling system, the existence of the additional contribution to  $R_2$  rates becomes evident for practically all residues in ubiquitin, where it likewise was admitted that this effect can hardly be explained by chemical exchange (Charlier et al. 2013). It is very likely that this contribution to  $R_2$ 's also results from the “long tail”. It is worth noting that the relaxation measurements at resonance frequencies of several MHz (as was also done in our previous papers mentioned above) can detect the slow component, but cannot provide the parameters  $S_{\text{rot}}^2$  and  $\tau_S$ , separately: from Eq. (3) it follows that only the product  $S_{\text{rot}}^2 \tau_S$  can be safely obtained from the analysis of these data.

Although the existence of the “long tail” is by now well documented, detailed data characterizing its behavior at different conditions are very sparse and incomplete due to the methodological challenges described above. To obtain exact quantitative information on the “long tail”, one has to measure the relaxation times at low resonance frequencies ( $\sim 1$  MHz and below), which is only possible using field-cycling (FC) NMR relaxometry (Koenig and Brown 1990; Kimmich and Anoardo 2004; Kruk et al. 2012). However, almost all FC experiments in protein solutions by now dealt with the solvent ( $\text{H}_2\text{O}$  or  $\text{D}_2\text{O}$ ) instead of the protein signal, as FC NMR features a low sensitivity. The analysis of the water spin–lattice relaxation rate dispersions  $R_1(\omega)$  can indeed resolve the “long tail” of the protein tumbling correlation function (Krushelnitsky and Fedotov 1993; Krushelnitsky 2006), yet these estimates can be affected by the finite lifetime of water molecules within the protein structure (Denisov and Halle 1996). If the water lifetime overlaps with  $\tau_S$ , the parameters of the slow component cannot be determined accurately.

Measuring the protein signal in the FC experiments is more challenging, and we are aware of very few applications of this kind so far which benefitted from the increased sensitivity of modern instruments (Bertini et al. 2005; Luchinat and Parigi 2007). The key point of these important studies was to demonstrate that the integral proton signal can be used to extract an average order parameter that serves as a faithful global measure of internal protein flexibility. The data shown do exhibit indications of the “long tail” (Bertini et al. 2005), but the related gradual increase of  $R_1$  at low frequencies was discussed in terms of protein aggregation. As discussed earlier (Krushelnitsky 2006), protein aggregation is an unlikely explanation for the appearance of the slow component in the tumbling correlation function. The results shown below will provide additional evidence that protein aggregation can hardly

explain the slow component, at least not in the general case.

In this work we conduct a systematic study of the “long tail” of the tumbling correlation function using  $^1\text{H}$  FC NMR relaxometry of protein protons in  $\text{D}_2\text{O}$  solutions combined with standard  $R_2$  and  $R_{1\rho}$  measurements that complement the FC NMR data at low frequencies. The main questions targeted in this study are: how do the parameters of the “long tail” behave upon varying the protein concentration?; how critical is the neglect of the “long tail” in the high-field relaxation data analysis?; how small should the protein concentration be for safely neglecting the “long tail”? We try to answer these questions by means of relaxation times analysis, simulations, and considering independent literature data.

## Materials and methods

### NMR experiments

The frequency dependence (dispersion) of the spin–lattice relaxation rate  $R_1(\omega)$  was measured with a commercially available STELAR FC 2000 relaxometer located at the University of Bayreuth. It allows for proton frequencies of  $2\pi \cdot 10 \text{ kHz} \leq \omega = \gamma_H B_0 \leq 2\pi \cdot 20 \text{ MHz}$ , with  $\gamma_H$  denoting the gyromagnetic ratio and  $B_0$  the magnetic field, respectively. The latter is generated by an electromagnet and thus variable. The typical time necessary for switching and stabilizing the coil current is around 2 ms. Basics of electronic FC NMR are discussed extensively in the literature (Kimmich and Anoardo 2004). The temperature can be varied within  $-120^\circ\text{C} \leq T \leq 180^\circ\text{C}$  while for the present contribution only a small interval of  $4^\circ\text{C} \leq T \leq 48^\circ\text{C}$  could be covered avoiding protein freezing and denaturation of the sample. In order to decompose the magnetization decay into protein and residual water (HDO,  $\text{H}_2\text{O}$ ) contributions the polarization time was always 3 s, enough to provide sufficient water signal. The analysis of the multi-component magnetization decays will be discussed below. Although the relaxometer enables measurements at the frequencies down to 10 kHz, in our FC experiments the lowest frequency was only 100 kHz. At lower frequencies the relaxation rates could only be determined with some uncertainty, especially at high protein concentrations, see below. Details on this issue are presented in the “Appendix”. This limitation, however, was not significant since the resulting frequency gap is complemented by the  $R_{1\rho}$  measurements.

$R_2$  and  $R_1$  at 20 MHz were measured using a BRUKER MINISPEC mq20. The low resonance frequency for  $R_2$  experiments allows avoiding the contribution of chemical

exchange of protein protons to  $R_2$  (Luz and Meiboom 1963; Hills et al. 1989).  $R_{1\rho}$  measurements were performed on a BRUKER AVANCE II spectrometer with a magnetic field corresponding to 400 MHz proton Larmor frequency. Since we used spin-lock fields of 20 and 40 kHz, the contribution of chemical exchange to the  $R_{1\rho}$  relaxation rates is likewise negligible.  $R_2$ ,  $R_1$  at 20 MHz and  $R_{1\rho}$  were measured in Halle on the same samples as in Bayreuth.

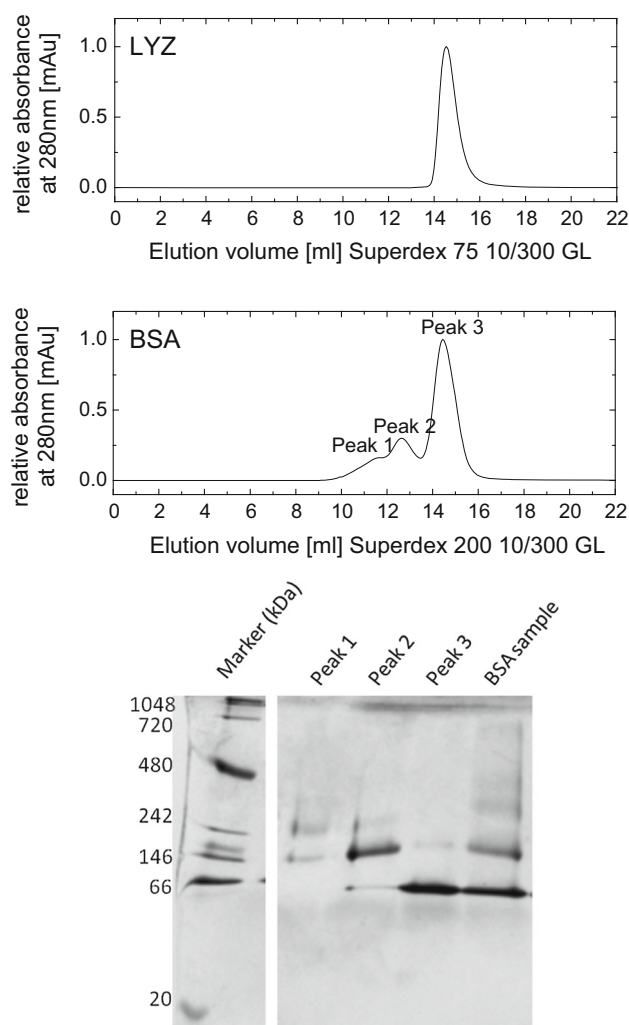
Prior to all measurements the protein- $D_2O$  solutions of different concentrations were filled into thoroughly cleaned NMR glass tubes. We abstained from degassing the samples to avoid  $D_2O$  evaporation and altered concentrations. For all experiments, we applied single short-pulse excitation with a sufficiently large spectral width of at least 50 kHz to ensure that all protein protons contribute equally to the integral signal. The accuracy of the temperature calibration and stabilization was in all cases better than  $\pm 1^\circ C$ .

### Sample preparation

Lysozyme (LYZ),  $M_w = 14,300$ , from chicken egg white and fatty acid free bovine serum albumin (BSA),  $M_w = 66,500$ , were delivered from Sigma-Aldrich (product numbers 62970 and A7030, respectively) as lyophilized powders. Both proteins were dissolved in  $D_2O$  for few hours and lyophilized again to maximally remove labile protons that would increase the residual water signal. Afterwards, the protein was dissolved in  $D_2O$  again without adding salt or buffer. The pH obtained was pH 3.8 for LYZ and pH 7.0 for BSA and is well distinguished from the isoelectric point (pH 11.35 and pH 4.7, respectively). No significant changes in the pH (more than 0.1–0.2) were observed upon varying the protein concentration.

Size-exclusion chromatography at a flow rate of 0.4 ml/min and blue native polyacrylamide electrophoresis (BN-PAGE) reveal the monodispersity of the LYZ sample and the polydispersity of the BSA sample (Fig. 2). The BSA polydispersity corresponds well to previous observations (Squire et al. 1968; Atmeh et al. 2007). To avoid unspecific interaction of the proteins with the column material, the elution buffer is adjusted to 50 mM NaCl in the case of BSA and 50 mM Na-phosphate buffer, 50 mM NaCl at pH 7.5 in the case of LYZ. The BN-PAGE of BSA was performed based on the method of Schagger et al. (1994) using a native unstained protein marker obtained from life technologies.

In our experiments we observed that at temperatures above 28–30 °C BSA solutions reveal a slow (on the time scale of several hours) increase of the oligomers portion. For this reason, for all BSA samples we limited the temperature range of the experiments to  $4^\circ C \leq T \leq 26^\circ C$ .

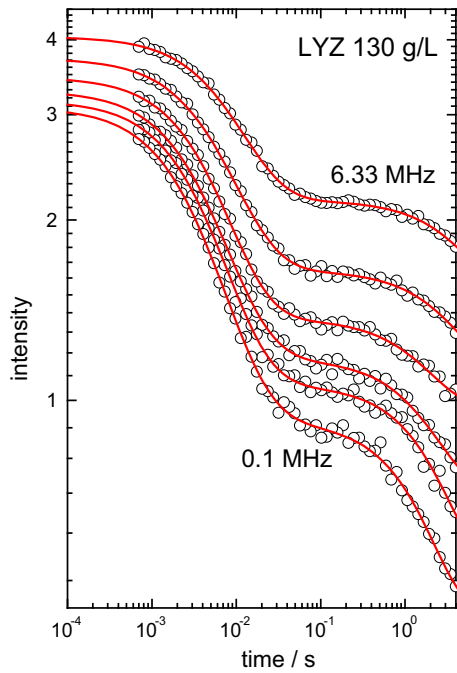


**Fig. 2** Size-exclusion chromatogram of LYZ and BSA (top) and blue native (BN) PAGE of BSA performed on the elution volume of different peaks of the size-exclusion chromatography (bottom)

## Results and discussion

### Analysis of the relaxation decays

Figures 3 and 4 present examples of the FC-NMR relaxation decays of the same sample at different magnetic field strengths. The relaxation decays consist of two components: the fast and slow decaying components belong to protein and residual water protons, respectively (Krushelnitsky and Fedotov 1993). Figure 3 shows raw data and Fig. 4 demonstrates the procedure of the water component subtraction and the form of the pure protein component at two different resonance frequencies [similar plots for the transversal magnetization decays are shown in ref. (Roos et al. 2015)]. Because of the fast exchange between hydrated and bulk water molecules (Denisov and Halle 1996), the water component is always single-exponential,



**Fig. 3** FC-NMR relaxation decays of 130 mg/ml LYZ at selected relaxation field strengths (from top to bottom): 6.33, 3.77, 2.25, 1.34, 0.795, 0.1 MHz. Solid lines are the log-normal fits combined with a single-exponential decay of the residual water protons (Eq. 4)

whereas the protein component is not (Krushelnitsky 2006; Luchinat and Parigi 2007; Roos et al. 2015).

The raw decays were fitted using a sum of the water and the protein components, the latter featuring a log-normal distribution of the relaxation rates:

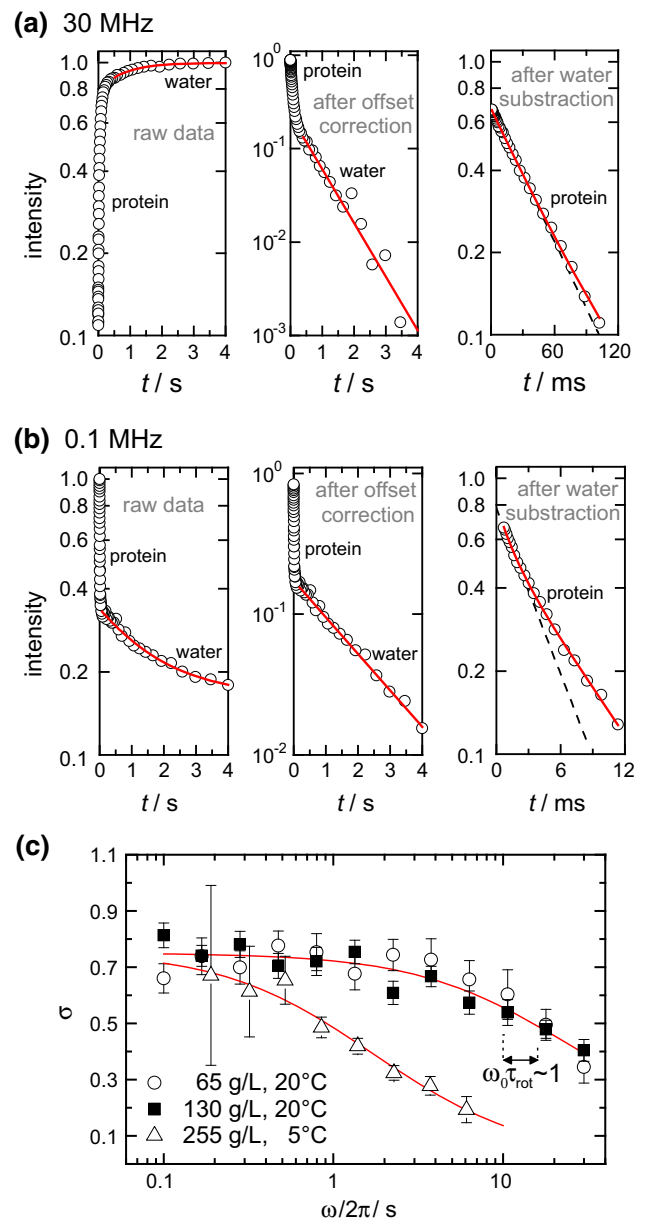
$$I(t) = (I_0 - I_{eq})f_{RLX}(t) + I_{eq}$$

$$f_{RLX}(t) = (1 - p_W) \int_0^\infty p(R_1) \exp(-R_1 t) dR_1 + p_W \exp(-R_W t)$$

$$p(R_1) = \frac{1}{\sqrt{2\pi\sigma^2 R_1}} \exp\left[-\frac{(\ln[R_1] - \ln[R_{median}])^2}{2\sigma^2}\right], \tag{4}$$

where  $I_0$  and  $I_{eq}$  correspond to the initial and equilibrium intensity, respectively,  $p_W$  and  $R_W$  are the relative amplitude and relaxation rate of residual water protons, respectively.  $R_{median}$  and  $\sigma$  correspond to the most populated relaxation rate ( $p(R_{median})$  has a maximal value) and the distribution width parameter ( $0 < \sigma < \infty$ ; 0 corresponds to the delta-function, infinity means infinitely wide distribution) of the protein component, respectively. The integral in Eq. (4) was calculated numerically during the fitting. The mean (arithmetic average) of the protein relaxation rates

$$\langle R_1 \rangle = \int_0^\infty p(R_1) R_1 dR_1 = R_{median} \cdot \exp\left(\frac{\sigma^2}{2}\right) \tag{5}$$



**Fig. 4** Relaxation decays in a 130 mg/ml lysozyme sample at (a) 30 MHz and (b) 0.1 MHz and separate fitting result (solid lines) of the water and the protein signal. Left (a, b) Intensities versus relaxation delay as directly obtained in the field-cycling experiment. Middle (a, b): After subtracting the value of the equilibrium magnetization, the mono-exponential decay of the water protons can be clearly seen. Right (a, b): Protein signal as observed after subtracting the water signal. Dotted lines indicate the initial slope of the decays, i.e. the mean relaxation rate. (c) Distribution width parameter of LYZ  $R_1$ s for three concentrations as a function of the proton resonance frequency as obtained by a log-normal fit, see Eq. 4. Vertical dotted lines indicate the frequencies obeying the condition  $\omega_0\tau = 1$  for the protein concentrations 65 and 130 mg/ml. Red solid lines are polynomial fits to guide an eye

equals the initial slope of the relaxation decay and is proportional to the mean spectral density function  $\langle J(\omega) \rangle$  (that is the spectral density function averaged over all protein

protons), as explained in the Supporting Information of ref. (Krushelnitsky et al. 2014). For the subsequent analysis, we use the mean relaxation rate values defined by Eq. (5) for all types of experiments, that is  $R_1$ ,  $R_{1\rho}$  and  $R_2$ . The dispersion profiles of the water component were not analyzed because of the inferior signal-to-noise ratio. The water component intensity in all samples was several times less than that of the protein; in addition, the long-time plateau limit of the relaxation decays was not always measured because of the long delays. Increasing the quality of these data would require significant increase of the measuring time, which was practically not affordable.

Defining the mean relaxation rates using the log-normal distribution and previously used (Roos et al. 2015) bi-exponential decomposition of the protein component provides in most cases essentially the same results, see Electronic Supporting Material (ESM), Fig. S2. However, the log-normal distribution enables a quantitative characterization of the non-exponentiality of the protein proton relaxation component. Figure 4c shows the distribution width parameter  $\sigma$  as a function of the resonance frequency. For 65 and 130 mg/ml, this dependence corresponds well to the theoretical prediction (Kalk and Berendsen 1976): if  $\omega\tau \gg 1$  ( $\tau$  is the correlation time of the protein Brownian tumbling), the rate of the spin exchange between protons in a protein is much faster than the relaxation rate of individual protons, rendering the integral relaxation decay singly exponential. Yet, even at relatively high frequencies ( $\omega > 10$  MHz) the inequality  $\omega\tau \gg 1$  does not hold strictly, and thus the decay is still somewhat non-exponential. On the opposite, at  $\omega\tau \ll 1$  spin exchange is slow in comparison to spin relaxation. Thus, the intrinsic distribution width of longitudinal relaxation times is observed. At 255 mg/ml at low temperatures, in turn, the anisotropy of Brownian tumbling (“long tail”) is quite pronounced, and the effective correlation time of Brownian tumbling becomes much longer. Spin exchange becomes much more efficient, so that the distribution width  $\sigma$  is reduced to lower values even at lower frequencies. A similar behavior of the shape of the relaxation decay as a function of frequency in FC experiments was reported earlier by Luchinat and Parigi (2007).

We stress that neglecting the distribution of longitudinal relaxation times and fitting such decay with a single exponent may provide quantitatively incorrect results. It is worth mentioning that, irrespective of whether spin exchange is fast or slow, the arithmetic average relaxation rate (as defined by the initial slope) is not affected by spin exchange (Kalk and Berendsen 1976). To avoid possible misunderstandings, we stress that the non-exponential form of relaxation decays bears absolutely no relation to a potentially non-exponential correlation function of motion.

Figure 5 presents exemplary  $R_1$  dispersions for LYZ and BSA along with the fitting curves (see below for the fitting

procedure) for different concentrations (the full set of the dispersion data are presented in ESM, Figs. S3–S5). The low-frequency FC relaxation data for BSA solutions are unfortunately not suitable for the analysis, which has two reasons. The first one is the fact that the switching time in FC-experiments (switching from polarization field down to the relaxation field and then back to the detection field) is comparable to the relaxation time. In the “Appendix” we show that this induces no error for a single-exponential decay, but may lead to erroneous results in the case of multi-exponential relaxation decays. The presence of oligomers in the BSA solution (see “Materials and methods” section) causes an additional source of distribution of the  $R_1$  relaxation rates, which leads to a much larger deviation from a mono-exponential behavior of the BSA relaxation decays at low frequencies in comparison with LYZ.

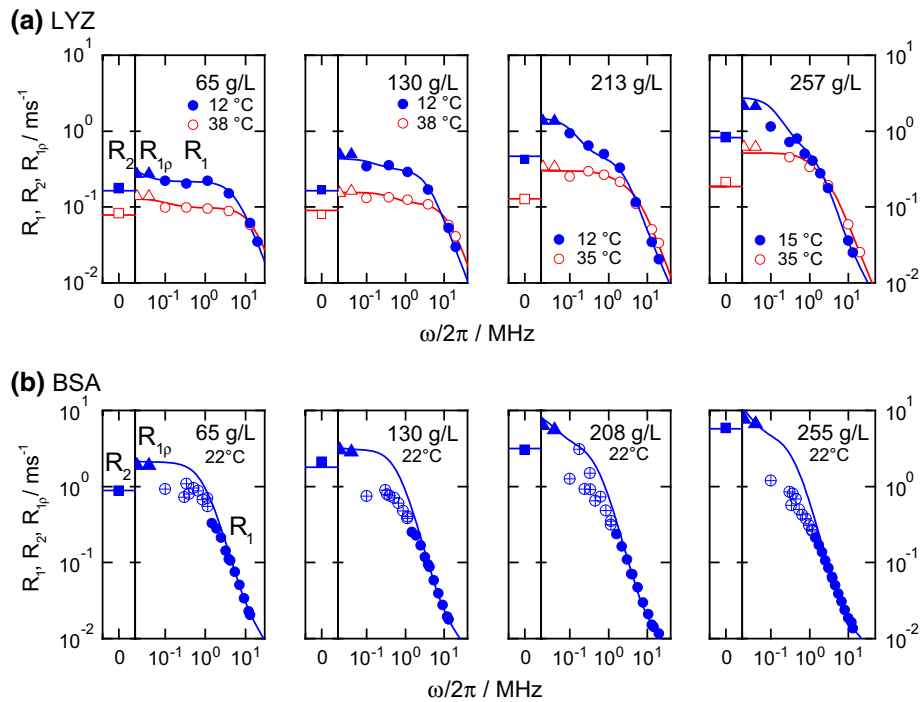
The second reason is the shortest possible relaxation delay in the FC-experiments which was 0.7 ms (“dead time”). Because of the large  $\tau_{\text{rot}}$ , the relaxation time of BSA protons at low frequencies is quite short and it is comparable to this “dead time”. The remaining long-time decay appears to be close to singly exponential, but the mean rate (initial slope) is severely underestimated, as demonstrated by Fig. S1. As a criterion whether the respective measurement is reliable or not, the initial slope as provided from the log-normal fit (Eq. 5) was compared to the initial slope as obtained by the bi-exponential decomposition, see ESM, Fig. S2. If and only if these two were the same, the data were considered as being reliable for the analysis. For LYZ, in turn, the spread of relaxation times within one decay curve is not only less pronounced, but also the FC relaxation times are longer, rendering the issue of ill-defined initial slopes less problematic. In fact, for LYZ at all frequencies measured, the initial slope defined from bi-exponential decomposition matches within the experimental error the initial slope of the log-normal fits (Fig. S2).

Another way of presenting the data, namely, temperature dependencies of relaxation rates at several fixed frequencies, is shown in Fig. S5. If  $R_1$  is plotted against inverse temperature  $1/T$ , the slope of the data already provides information whether  $\omega\tau_{\text{rot}} < 1$  (positive slope) or  $\omega\tau_{\text{rot}} > 1$  (negative slope). For this reason, the temperature dependencies of the relaxation times measured at various frequencies enable much more precise fitting and more confident identification of different motional processes in comparison with frequency dependencies only.

### Analysis of the relaxation times

The  $R_2$ ,  $R_{1\rho}$  and  $R_1$  relaxation times of protein protons are dominated by the homonuclear dipole–dipole relaxation mechanism,





**Fig. 5** Dispersion profiles of (a) LYZ and (b) BSA at different concentrations. For direct visual comparison of  $R_{1\rho}$  (triangles) and  $R_1$  (circles),  $R_{1\rho}$  data were multiplied by 10/3 (see ESM, Eqs. S1–S5).  $R_2$ 's (squares) were measured at 20 MHz and are shown in a separate column of each plot. Solid lines provide the best fit result. Uncertain data points (BSA at  $\omega_0/2\pi < 3$  MHz, see text) were not taken into account for fitting (crossed symbols). For BSA, a detailed frequency

dependence was recorded only at 22 °C; at other temperatures the data for only few frequencies were measured. The here shown fit to the BSA data assumes monomers only. A fit result involving oligomers is presented in the ESM, Fig. S4. For both proteins, the data shown here were fitted together with all the available data shown in ESM, Figs. S3 and S5

$$R_1 = \frac{2}{3} K_{HH} (J(\omega_0) + 4J(2\omega_0)) \tag{6}$$

$$R_{1\rho} = \frac{1}{3} K_{HH} (3J(2\omega_1) + 5J(\omega_0) + 2J(2\omega_0)) \tag{7}$$

$$R_2 = \frac{1}{3} K_{HH} (3J(0) + 5J(\omega_0) + 2J(2\omega_0)), \tag{8}$$

where  $K_{HH}$  is the mean dipole–dipole coupling of protons in the protein,  $J(\omega)$  is the spectral density function, and  $\omega_0/2\pi$  and  $\omega_1/2\pi$  are the proton resonance and spin-lock frequencies, respectively. Assuming that  $\tau_{int} \ll \tau_{rot} \ll \tau_S$ , the spectral density function can be derived from Eq. (2) as

$$J(\omega) = (1 - S_{int}^2) \frac{\tau_{int}}{1 + (\omega\tau_{int})^2} + S_{int}^2 (1 - S_{rot}^2) \frac{\tau_{rot}}{1 + (\omega\tau_{rot})^2} + S_{int}^2 S_{rot}^2 \frac{\tau_S}{1 + (\omega\tau_S)^2}. \tag{9}$$

Since we analyzed relaxation times measured at different temperatures, we assumed an Arrhenius temperature dependence for the motional correlation times:

$$\tau_{S,rot,int} = \tau_{S,rot,int}(293K) \exp \left[ \frac{E_{S,rot,int}}{R} \left( \frac{1}{T} - \frac{1}{293K} \right) \right]. \tag{10}$$

Note that in reality the temperature dependencies of the correlation times may deviate from the Arrhenius law, however, within the relatively narrow temperature range of our experiments this approximation works quite well. Overall, the set of fitting parameters included the correlation times  $\tau_{int}$ ,  $\tau_{rot}$  and  $\tau_S$  at 20 °C, the activation energies  $E_{int}$ ,  $E_{rot}$  and  $E_S$ , the order parameters  $S_{int}^2$  and  $S_{rot}^2$ , and the mean proton coupling constant  $K_{HH}$ . Although the existence of the “long tail” was established in our previous papers, we demonstrate here again that neglecting it in the fitting model leads to systematic inconsistency between the experimental and fitting values of the relaxation rates, see ESM, Fig. S6.

The number of the fitting parameters is quite large, yet the number of experimental data was much larger (overall for all temperatures and concentrations, LYZ: 361 data points, BSA: 262 data points) and, thus, the fitting was quite stable and provides reasonable fitting uncertainties. The values of the fitting parameters were obtained from the simultaneous (global) Monte-Carlo fit of the set of all relaxation times measured at all temperatures and concentrations. The fitting aimed to minimize the root mean square deviation

$$RMSD = \sqrt{\frac{1}{N} \sum_{i=1}^N \left(1 - \frac{R_{\text{exp}}}{R_{\text{sim}}}\right)^2}, \quad (11)$$

where  $R_{\text{exp}}$  and  $R_{\text{sim}}$  are the simulated (according to the current set of the fitting parameters) and experimental relaxation times, respectively, and  $N$  is the number of all relaxation times. While fitting, the parameters of the internal motions ( $S_{\text{int}}^2$ ,  $E_{\text{int}}$  and  $\tau_{\text{int}}$ ) as well as  $K_{\text{HH}}$  were assumed to be the same for all concentrations (that is why we fitted the data at all concentrations at a time); all other parameters were assumed to be concentration-dependent. Tables S1 and S2 of ESM contain all the fitting parameters for the two proteins at all concentrations investigated.

A special remark on internal motions is in order. The internal motion parameters are not of central interest in this study, yet they need to be included in the analysis since relaxation times at high frequencies (around 10 MHz and higher) have an appreciable contribution from the first term in Eq. (9), and neglecting it may lead to an incorrect estimation of other parameters. We describe internal motions by a single correlation time  $\tau_{\text{int}}$ , however, we actually assume two components of the internal motions since the mean proton coupling constant  $K_{\text{HH}}$  (Eqs. 6–8) is kept as a freely adjustable fitting parameter. Hence, fast internal motions contribute to the  $K_{\text{HH}}$  motional averaging, so that the parameters  $S_{\text{int}}^2$  and  $\tau_{\text{int}}$  only reflect the slow mode of internal dynamics. Indeed, the rigid lattice value of  $K_{\text{HH}}$  for globular proteins is around  $1.3 \times 10^{10} \text{ s}^{-2}$ . Fast methyl proton rotation reduces  $K_{\text{HH}}$  down to  $\sim 0.85 \times 10^{10} \text{ s}^{-2}$  (Krushelnitsky et al. 1996), whereas the  $K_{\text{HH}}$  fitting value (see Tables S1 and S2) is around  $0.6 \times 10^{10} \text{ s}^{-2}$ . For a more accurate evaluation of the internal motion parameters  $S_{\text{int}}^2$  and  $\tau_{\text{int}}$  (and thus more precise estimation of the overall tumbling parameters), we included previously published data on  $^1\text{H}$   $R_1$  temperature dependencies in BSA and LYZ solutions measured at 11, 27 and 90 MHz (Krushelnitsky 2006) ( $R_2$  data from this work were not included). The relaxation decays in ref. (Krushelnitsky 2006) were analyzed according to the same protocol as in this study, thus the relaxation times can be compared directly. These  $R_1$  data were added to the set of relaxation times at the lowest concentration (65 mg/ml).

The fact that the apparent internal correlation times were in all cases a factor of at least 5–25 shorter than the global tumbling times (see Tables S1 and S2) demonstrates that slower internal motions of some residues, reaching the timescale of global rotation and thus potentially distorting the tumbling correlation function at its  $\tau_{\text{rot}}$ -related onset, are probably sparse. If they were abundant, we would expect the separation to be less clear.

For LYZ, the fitting provides a rotational correlation time  $\tau_{\text{rot}}(T = 20 \text{ }^\circ\text{C}) = (10.5 \pm 0.2) \text{ ns}$  at the lowest

concentration measured in this study. We also estimated  $\tau_{\text{rot}}$  using the Stokes–Einstein–Debye law and the values of the viscosity of the protein solution ( $1.538 \pm 0.010 \text{ mPa s}$ , measured by the micro-viscometer mVROC, Rheosense, San Ramon, CA) and lysozyme’s hydrodynamic radius of 1.9 nm (Parmar and Muschol 2009). The result appeared to be  $(10.9 \pm 0.1) \text{ ns}$ , which is in a good accordance with our analysis.

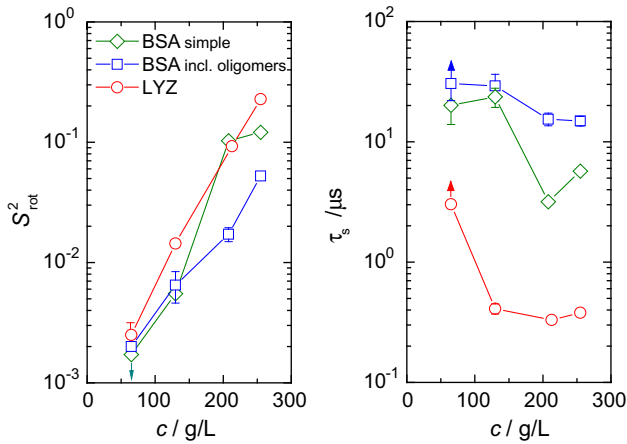
The analysis of the BSA data is less unambiguous since BSA solutions contain a significant portion of oligomers, see “Materials and methods” section. We applied a more complicated form of the rotational correlation function that contains an additional component attributed to oligomers, see ESM, Eq. S8. However, this more sophisticated analysis does not change significantly the behavior of the “long tail” parameters. Note, however, that  $\tau_{\text{rot}}$  for BSA monomers obtained from this fitting (40 ns) matches well the independent literature value (Ferrer et al. 2001) confirming the adequacy of this analysis in spite of the increased number of fitting parameters.

### Concentration dependence of the “long tail” parameters

In our previous studies (Krushelnitsky and Fedotov 1993; Krushelnitsky 2006; Roos et al. 2015) we could not estimate the parameters of the slow component,  $S_{\text{rot}}^2$  and  $\tau_{\text{S}}$ , separately for the reasons described in the Introduction. Combining FC NMR and routine  $R_2$  and  $R_{1\rho}$  measurements, we now can overcome this issue, and the concentration dependence of  $S_{\text{rot}}^2$  and  $\tau_{\text{S}}$  can be resolved. Figure 6 presents the concentration dependence of these two parameters for both proteins. At lowest concentration, however, we still cannot determine reliably the parameters  $S_{\text{rot}}^2$  and  $\tau_{\text{S}}$  separately, since  $\tau_{\text{S}}$  appears to be too long and the dispersion step corresponding to the “long tail” (Fig. 1) falls into a “dead zone” between  $R_2$  and the  $R_{1\rho}$ ’s. Consequently, for this concentration, Fig. 6 shows merely high and low limits for  $S_{\text{rot}}^2$  and  $\tau_{\text{S}}$ , respectively, indicated by arrows. At higher concentrations,  $\tau_{\text{S}}$  is shorter (see below), and both “long tail” parameters could be obtained from the fitting independently.

The behavior of  $S_{\text{rot}}^2$  appears reasonable: with increasing concentration, the average distance between proteins shortens, resulting in stronger inter-protein interactions. This increases the anisotropy of local Brownian rotation and, hence, the order parameter  $S_{\text{rot}}^2$  becomes larger.

In contrast,  $\tau_{\text{S}}$  decreases with increasing concentration, which seems unexpected. In our view, this tendency can be explained as follows. The lifetime of the local anisotropy relies on the time span needed to change the local environment of the particle (i.e. the protein). With decreasing



**Fig. 6** Order parameter  $S_{rot}^2$  and correlation time of the slow component  $\tau_s$  ( $T = 20\text{ }^\circ\text{C}$ ) of rotational diffusion as a function of protein concentration of LYZ (circles) and BSA (squares and diamonds). For BSA, two sets of data are shown corresponding to the analyses using the simple and the more complicated form of the correlation function that includes oligomers (see ESM, Table S2). At the lowest concentration (65 g/L), the upper and lower boundaries of  $S_{rot}^2$  and  $\tau_s$ , respectively, are shown (indicated by arrows). The minimum (65 g/L) and maximum (260 g/L) concentrations correspond to 4.5 and 18.2 mM for LYZ and 1 and 3.9 mM for BSA, respectively

inter-protein distances (i.e. increasing concentration), fluctuations in the particle alignment have an increased impact on the actual configuration of the particle’s local environment. The lifetime of a certain particle configuration is in particular limited by translational diffusion on the length scale of neighboring particles. Therefore,  $\tau_s$  becomes shorter with increasing protein concentration, at least until a certain concentration close to jamming conditions is reached. Such a trend can also be seen in collective diffusion coefficients measured by dynamic light scattering (Heinen et al. 2012); yet the detailed concentration dependence looks different due to the different natures and time scales of the two processes.

At the same time, we have to admit that this simple explanation cannot explain the  $\tau_s$  concentration dependence quantitatively, especially the significant decrease of  $\tau_s$  between the first and the second concentration points for LYZ. In part, such behavior may be caused by experimental uncertainty since only few relaxation rates report on  $\tau_s$  and  $S_{rot}^2$  at low concentrations, where  $S_{rot}^2$  is very small. More importantly, any simplified consideration of the slow tail of rotational diffusion does not take into account multi-body interactions, as well as real sizes and shapes of the proteins and their charge distributions. All these aspects render the physics very complex, and simplified models can be expected to account for the qualitative behavior only. We believe that the quantitative description of the “long tail” parameters is only possible using Brownian

dynamics simulations that take into account the real shape and the sophisticated electrostatic structure of the protein of consideration (McGuffee and Elcock 2006).

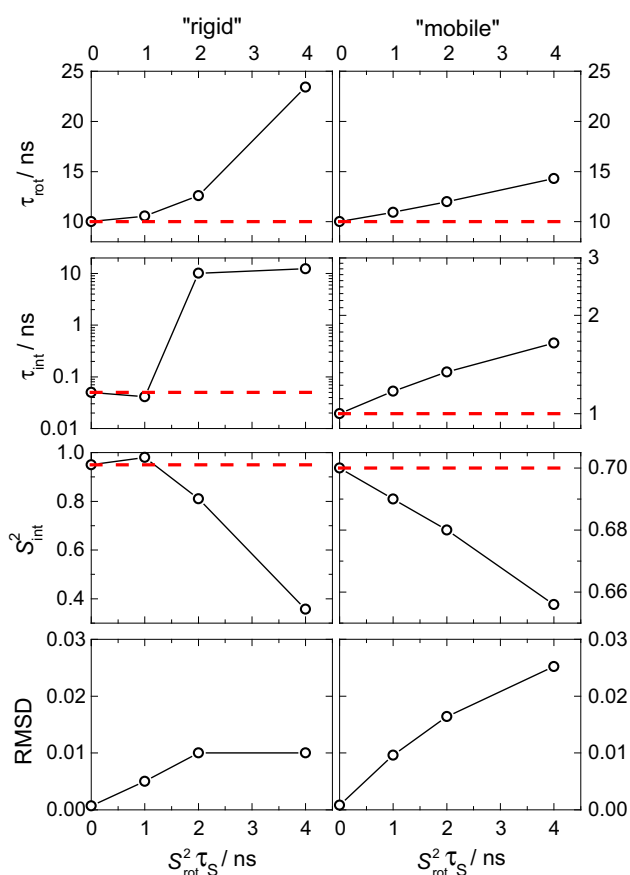
The observed concentration dependence of  $\tau_s$  supports the previous conclusion on the protein aggregation as an unlikely source of the “long tail” (Krushelnitsky 2006). A small portion of large aggregates would also lead to appearance of the slow component; this, however, would be physically unrealistic: large aggregates without an appreciable amount of small oligomers do not conform to Ostwald’s dilution law. The concentration dependence of  $\tau_s$  for both proteins indicates that the size of potential oligomers must decrease with increasing concentration, which is neither physical nor conceivable (see for comparison the behavior of the correlation time of the oligomer component in the BSA data analysis, Table S2, part B). Moreover, if large stable aggregates were present in protein solution, one would observe two components in the magnetization decays with distinctly different relaxation times corresponding to monomers and aggregates, which is not the case. Thus, one may conclude that the protein aggregation does not contribute, at least significantly, to the “long tail”.

### Influence of neglecting the “long tail” while fitting high-field NMR relaxation data

For high-field NMR relaxation studies of protein dynamics the important question is: what is the effect of neglecting the “long tail” in the routine model-free analysis of the relaxation data? To check this aspect, we simulated and then fitted the  $^{15}\text{N}$   $R_1$ ,  $R_2$  and  $NOE$  data at three proton resonance frequencies—500, 600 and 800 MHz. Overall, we analyzed a set of nine experimental (simulated) parameters. The relaxation parameters  $R_1$ ,  $R_2$  and  $NOE$  were simulated assuming the correlation function to be in the form of Eq. (2), see ESM, Fig. S7. In all cases we assumed  $\tau_{rot} = 10\text{ ns}$ , and simulated the data for two cases: a “rigid” residue ( $S_{int}^2 = 0.95$ ,  $\tau_{int} = 50\text{ ps}$ ) and a “mobile” one ( $S_{int}^2 = 0.7$ ,  $\tau_{int} = 1\text{ ns}$ ). These values of the internal motion parameters are typical for residues in the secondary structure elements and (partially) unstructured domains, respectively. For both cases, we simulated the relaxation data assuming four different amplitudes of the slow component:  $S_{rot}^2 = 0, 0.005, 0.01$ , and  $0.02$ , where  $\tau_s$  was fixed to 200 ns in all cases (As mentioned above, only the product  $S_{rot}^2 \cdot \tau_s$  matters for the analysis). It should be mentioned that the  $S_{rot}^2$  values derived from the  $^1\text{H}$  non-selective relaxometry cannot be directly transferred into the analysis of  $^{15}\text{N}$  relaxation data since different N–H vectors may experience different rotational anisotropies depending on their orientation within the protein structure. Still, the

order of magnitude of  $S_{\text{rot}}^2$  for  $^1\text{H}$  and  $^{15}\text{N}$  relaxation should be the same.

Figure 7 shows the values of the parameters  $\tau_{\text{rot}}$ ,  $\tau_{\text{int}}$  and  $S_{\text{int}}^2$  obtained from fitting (with fixed  $S_{\text{rot}}^2 = 0$ ) the simulated relaxation data generated with various  $S_{\text{rot}}^2$  values (see above). See Fig. 1 for an illustration of the input (real) and the fitted (apparent) spectral densities for the “mobile” case with  $S_{\text{rot}}^2 = 0.02$ . It can be clearly seen that neglecting the “long tail” can lead to appreciably mismatching fitting parameters. Although the root mean square deviation of the fitting result increases with increasing  $S_{\text{rot}}^2$ , it, however, still remains quite small and reaches just a few percent, which corresponds to the typical experimental error of high-field NMR relaxation measurements. Again, this outcome demonstrates that the “long tail” *cannot* be seen in this type of analysis. If one aims at the precise NMR relaxation analysis of the protein dynamics in solution, one has either to combine high-field and low-field experiments, or to measure the concentration dependencies of the relaxation



**Fig. 7** The values of the dynamic parameters  $\tau_{\text{rot}}$ ,  $\tau_{\text{int}}$  and  $S_{\text{int}}^2$  as a function of the product  $S_{\text{rot}}^2 \tau_S$  as obtained from fitting the simulated relaxation data assuming  $S_{\text{rot}}^2 = 0$ . The red dotted lines indicate the correct values of the parameters used in the simulations. The two bottom plots show the root mean square deviation (RMSD) of the fitting result

parameters in order to extrapolate them to zero concentration, where the “long tail” vanishes. This approach, however, would make the study obviously much more laborious.

It would be of interest to know the concentration from which on the inter-protein interactions can be assumed to be negligible. NMR is definitely not the best method for experimental detection of the “long tail” at very low concentrations since it suffers from the inherently low sensitivity in comparison with many other techniques. More sensitive methods for direct observation of inter-protein interactions are neutron and X-ray small-angle scattering (SAS) experiments (Roosen-Runge et al. 2011). Structure factors as obtained from SAS demonstrate that proteins cannot generally be assumed to act as non-interacting hard spheres. The inherent order of proteins in solutions is strongly affected by the strength of the repulsive electrostatic interactions, and consequently depends on the pH and ionic strength of the solvent (Velev et al. 1998) that influence the net charge and electrostatic screening, respectively.

Essentially, NMR and SAS methods observe the same phenomenon—long-range inter-protein electrostatic interactions. It is important, however, that the information provided by these methods is not similar but complementary: SAS and NMR report on spatial and orientational ordering of proteins, respectively, caused by electrostatic interactions. Thus, studies of protein solutions combining the capabilities of both methods will likely enable a deeper understanding of the nature of inter-protein interactions. At the moment we would like to attract reader’s attention to the studies (Stradner et al. 2004; Shukla et al. 2008; Heinen et al. 2012) in which SAS data were recorded in protein solutions in a wide concentration range. These results show that the spatial ordering (i.e. repulsive inter-protein electrostatic interactions) becomes negligible only at protein concentrations at about a few mg/ml. At such low concentrations, liquid-state NMR relaxation measurements are hardly possible without a cryoprobe.

## Conclusions

Mutual protein electrostatic steering is an essential effect in protein solution even at relatively low concentrations, which may appreciably affect the NMR relaxation rates, primarily  $R_2$ . Neglecting it in the analysis of NMR relaxation data may lead to imprecise results, in particular at high concentrations. The parameters of the “long tail” of the rotational correlation function reveal different concentration dependencies: the intensity of the slow component (order parameter of the Brownian tumbling) expectedly increases with increasing concentration.

However, its correlation time, i.e. the lifetime of the local anisotropy, decreases. This behavior can be understood when taking into account the shorter inter-protein distances at higher concentrations. Reduced inter-particle distances lead to an increased sensitivity to environmental fluctuations, which in turn decreases the life-time of the actual particle configuration. Literature data on X-ray and neutron small-angle scattering demonstrate the spatial ordering of proteins in solution, which complements the NMR results on the orientational ordering. As a more sensitive technique, small-angle scattering experiments demonstrate that protein spatial ordering is present at concentrations even below 1 mM, i.e. at the edge of the concentration range suitable for the NMR relaxation experiments.

**Acknowledgments** We thank the Deutsche Forschungsgemeinschaft (DFG, SFB-TRR 102 Project A8) for funding this work.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

**Appendix: Influence of the limited ramp time in FC-experiments on the shape of relaxation decay**

In field-cycling experiments, the magnetic field is ramped down from the polarization field (here:  $\omega_0^{(pol)}/2\pi = 20$  MHz) towards the adjustable relaxation field, and is then ramped up again to the detection field (here:  $\omega_0^{(acq)}/2\pi = 16$  MHz). The ramp time of the FC relaxation measurements was around 2 ms, and is thus comparable to the protein’s mean relaxation rate at low magnetic fields. The following section addresses the consequences for the multi-exponential relaxation decays as observed in this study.

In presence of a time dependent magnetic field, a *single-component* relaxation decay can be written as

$$\frac{I(t_N)}{I_0} = \lim_{\Delta t \rightarrow 0} \prod_{i=1}^N \exp[-R_1(\omega(t_i))\Delta t], \quad t_N = N\Delta t, \quad (12)$$

from which one can obtain a time-dependent mean relaxation rate  $\bar{R}_1(t)$ :

$$\frac{I(t)}{I_0} = \exp[-\bar{R}_1(t) t], \quad \bar{R}_1(t) = \frac{1}{t} \int_0^t R_1(\omega(t')) dt'. \quad (13)$$

Hence, the intensity after ramping down the magnetic field simply reads

$$I^{(rd)} = \left( I^{(pol)} - I^{(rlx)} \right) \exp\left[-\bar{R}_1^{(rd)} \tau^{(rd)}\right] + I^{(rlx)}, \quad (14)$$

where  $I^{(pol)}$  and  $I^{(rlx)}$  denote the equilibrium intensity at the polarization and relaxation field, respectively,  $\tau^{(rd)}$  is the field-switching time, and  $\bar{R}_1^{(rd)}$  is the mean relaxation rate during the magnetization ramp as defined by Eq. (13). After the magnetization ramp, the intensity is now subject to relaxation during the adjustable relaxation delay  $t^{(rlx)}$ , and takes place at the wanted relaxation field of frequency  $\omega_0^{(rlx)}$ ,

$$I(t^{(rlx)}) = \left( I^{(rd)} - I^{(rlx)} \right) \exp\left[-R_1(\omega_0^{(rlx)}) t\right] + I^{(rlx)}. \quad (15)$$

Finally, ramping up the magnetic field for data acquisition during the field-switching time  $\tau^{(ru)}$  causes the intensity to relax towards its new equilibrium magnetization  $I^{(acq)}$ , so that the signal finally detected reads

$$\begin{aligned} I_{acq}(t^{(rlx)}) &= \left( I^{(acq)} - I(t^{(rlx)}) \right) \left( 1 - \exp\left[-\bar{R}_1^{(ru)} \tau^{(ru)}\right] \right) + I(t^{(rlx)}) \\ &= a \exp\left[-R_1(\omega_0^{(rlx)}) t^{(rlx)}\right] + b, \end{aligned} \quad (16)$$

where  $\bar{R}_1^{(ru)}$  is the mean relaxation rate during increasing the magnetic field, and  $a$  and  $b$  are constants defined as

$$a = \left( I^{(pol)} - I^{(rlx)} \right) \exp\left[-\left(\bar{R}_1^{(rd)} \tau^{(rd)} + \bar{R}_1^{(ru)} \tau^{(ru)}\right)\right] \quad (17)$$

$$b = I^{(acq)} \left( 1 - \exp\left[-\bar{R}_1^{(ru)} \tau\right] \right) + I^{(rlx)} \exp\left[-\bar{R}_1^{(ru)} \tau\right]. \quad (18)$$

The above equations demonstrate that if the protein’s relaxation decay is single-exponential, the limited ramp time of the magnetic field would only influence the initial and final intensity of the decay observed, but not the relaxation rate.

In case of a multi-exponential decay, in turn, the amplitude  $a$  and the equilibrium value  $b$  may be different for the individual components of the decay. A change in the offset  $b$  from one proton site to another is of no relevance for correctly estimating the mean relaxation rate  $\langle R_1(\omega_0^{(rlx)}) \rangle = \sum_i a_i R_{1,i}(\omega_0^{(rlx)})$ , as the plateau value of the relaxation decay is independently fitted, yet modified amplitudes  $a_i$  do affect a correct estimate of  $\langle R_1(\omega_0^{(rlx)}) \rangle$ .

The faster the relaxation during the magnetization ramp, the smaller is the contribution of this component to the experimentally observed decay. This issue is only of relevance at measurements at low relaxation fields, as only those magnetization ramps involve fast relaxation occurring at low magnetic fields. The potential decrease of the amplitude of quickly relaxing components may hence

cause an underestimation of the relaxation rate at low frequencies.

The presence of BSA oligomers in solution causes an additional spread of the relaxation rates, rendering the issue of multi-component decays problematic at low relaxation fields. Thus, measurements in BSA solutions at relaxation fields below 3 MHz  $^1\text{H}$  Larmor frequency were not taken into account, also due to the reasons discussed in ESM, Fig. S2.

The mean relaxation rate of LYZ, however, could be safely estimated even at relaxation fields as low as 0.1 MHz: fitting the relaxation decays provides stable results (Fig. S2) with a distribution width of relaxation times reaching a plateau value at low frequencies (Fig. 4), thus indicating that the *shape* of the FC-NMR relaxation decay is the same at different magnetic fields. At low magnetic fields, where relaxation is fast as compared to spin exchange, the distribution width  $\sigma$  reflects the intrinsic spread of relaxation times of the protein protons, and hence has to be independent of the actual magnetic field applied. If, however, relaxation during the magnetization ramps reduced the contribution of quickly relaxing proton sites to the overall signal, the experimentally observed distribution of relaxation times would become narrower at low magnetic fields. Such a trend cannot be seen in the distribution width (Fig. 4), thus indicating that the magnetization ramps have no significant effect on the mean relaxation rate in LYZ solutions.

## References

- Atmeh RF, Arafa IM, Al-Khateeb M (2007) Albumin aggregates: hydrodynamic shape and physico-chemical properties. *Jordan J Chem* 2:169–182
- Bertini I, Gupta YK, Luchinat C, Parigi G, Schorb C, Schwalbe H (2005) NMR spectroscopic detection of protein protons and longitudinal relaxation rates between 0.01 and 50 MHz. *Angew Chem Int Ed* 44:2223–2225
- Boehr DD, Dyson HJ, Wright PE (2006) An NMR perspective on enzyme dynamics. *Chem Rev* 106:3055–3079
- Charlier C, Khan SN, Marquardsen T, Pelupessy P, Reiss V, Sakellariou D, Bodenhausen G, Engelke F, Ferrage F (2013) Nanosecond time scale motions in proteins revealed by high-resolution NMR relaxometry. *J Am Chem Soc* 135:18665–18672
- Clore GM, Driscoll PC, Wingfield PT, Gronenborn AM (1990) Analysis of the backbone dynamics of interleukin-1 $\beta$  using two-dimensional inverse detected heteronuclear  $^{15}\text{N}$ - $^1\text{H}$  NMR spectroscopy. *Biochemistry* 29:7387–7401
- Daragan VA, Mayo KH (1997) Motional model analyses of protein and peptide dynamics using C-13 and N-15 NMR relaxation. *Prog Nucl Magn Reson Spectrosc* 31:63–105
- Denisov VP, Halle B (1996) Protein hydration dynamics in aqueous solution. *Faraday Discuss* 103:227–244
- Ermakova EA, Krushelnitsky AG, Fedotov VD (2002) Brownian dynamics simulation of electrostatically interacting particles. *Mol Phys* 100:2849–2855
- Ermolina IV, Krushelnitsky AG, Ivoylov IN, Feldman YD, Fedotov VD (1993) Investigation of molecular motions and interprotein interactions by NMR and TDDS. *Appl Magn Reson* 5:265–283
- Ferrer ML, Duchowicz R, Carrasco B, de la Torre GJ, Acuña AU (2001) The conformation of serum albumin in solution: a combined phosphorescence depolarization-hydrodynamic modeling study. *Biophys J* 80:2422–2430
- Heinen M, Zanini F, Roosen-Runge F, Fedunova D, Zhang F, Hennig M, Seydel T, Schweins R, Sztucki M, Antalík M, Schreiber F, Nägele G (2012) Viscosity and diffusion: crowding and salt effects in protein solutions. *Soft Matter* 8:1404–1419
- Hills BP, Takacs SF, Belton PS (1989) The effects of proteins on the proton NMR transverse relaxation times of water. I. Native bovine serum albumin. *Mol Phys* 67:903–918
- Ishima R (2012) Recent developments in N-15 NMR relaxation studies that probe protein backbone dynamics. *Top Curr Chem* 326:99–122
- Kalk A, Berendsen HJC (1976) Proton magnetic relaxation and spin diffusion in proteins. *J Magn Reson* 24:343–366
- Kay LE, Torchia DA, Bax A (1989) Backbone dynamics of proteins as studied by nitrogen-15 inverse detected heteronuclear NMR spectroscopy: application to staphylococcal nuclease. *Biochemistry* 28:8972–8979
- Kimmich R, Anorado E (2004) Field-cycling NMR relaxometry. *Prog Nucl Magn Reson Spectrosc* 44:257–320
- Kleckner IR, Foster MP (2011) An introduction to NMR-based approaches for measuring protein dynamics. *Biochim Biophys Acta* 1814:942–968
- Koenig SH, Brown RD (1990) Field-cycling relaxometry of protein solutions and tissue: implications for MRI. *Prog Nucl Magn Reson Spectrosc* 22:487–567
- Korzhev DM, Billeter M, Arseniev AS, Orekhov VY (2001) NMR studies of Brownian tumbling and internal motions in proteins. *Prog Nucl Magn Reson Spectrosc* 38:197–266
- Kruk D, Herrmann A, Rössler EA (2012) Field-cycling NMR relaxometry of viscous liquids and polymers. *Prog Nucl Magn Reson Spectrosc* 63:33–64
- Krushelnitsky A (2006) Intermolecular electrostatic interactions and Brownian tumbling in protein solutions. *Phys Chem Chem Phys* 8:2117–2128
- Krushelnitsky AG, Fedotov VD (1993) Overall and internal protein dynamics in solution studied by the nonselective proton relaxation. *J Biomol Struct Dyn* 11:121–141
- Krushelnitsky AG, Fedotov VD, Spevacek J, Straka J (1996) Dynamic structure of proteins in solid state. H-1 and C-13 NMR relaxation study. *J Biomol Struct Dyn* 14:211–224
- Krushelnitsky A, Zinkevich T, Reif B, Saalwächter K (2014) Slow motions in microcrystalline proteins as observed by MAS-dependent N-15 rotating-frame NMR relaxation. *J Magn Reson* 248:8–12
- Lipari G, Szabo A (1982a) Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 1. Theory and range of validity. *J Am Chem Soc* 104:4546–4559
- Lipari G, Szabo A (1982b) Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 2. Analysis of experimental results. *J Am Chem Soc* 104:4559–4570
- Luchinat C, Parigi G (2007) Collective relaxation of protein protons at very low magnetic field: a new window on protein dynamics and aggregation. *J Am Chem Soc* 129:1055–1064
- Luz Z, Meiboom S (1963) Nuclear magnetic resonance study of protolysis of trimethylammonium ion in aqueous solution: order of reaction with respect to solvent. *J Chem Phys* 39:366–370
- McGuffee SR, Elcock AH (2006) Atomically detailed simulations of concentrated protein solutions: the effect of salt, pH, point

- mutations, and protein concentration in simulations of 1000-molecule systems. *J Am Chem Soc* 128:12098–12110
- Palmer A (2001) NMR probes of molecular dynamics: overview and comparison with other techniques. *Ann Rev Biophys Biomol Struct* 30:129–155
- Parmar AS, Muschol M (2009) Hydration and hydrodynamic interactions of lysozyme: effects of chaotropic versus kosmotropic ions. *Biophys J* 97:590–598
- Roos M, Link S, Balbach J, Krushelnitsky A, Saalwächter K (2015) NMR-detected Brownian dynamics of  $\alpha$ B-crystallin over a wide range of concentrations. *Biophys J* 108:98–106
- Roosen-Runge F, Henniga M, Zhang F, Jacobs RMJ, Sztucki M, Schober H, Seydel T, Schreiber F (2011) Protein self-diffusion in crowded solutions. *Proc Natl Acad Sci USA* 108:11815–11820
- Saito H (2014) Dynamic pictures of proteins by NMR. *Ann Rep NMR Spectrosc* 83:1–66
- Schägger H, Cramer WA, von Jagow G (1994) Analysis of molecular masses and oligomeric states of protein complexes by blue native electrophoresis and isolation of membrane protein complexes by two-dimensional native electrophoresis. *Anal Biochem* 217:220–230
- Shukla A, Mylonas E, di Cola E, Finet S, Timmins P, Narayanan T, Svergun DI (2008) Absence of equilibrium cluster phase in concentrated protein solutions. *Proc Natl Acad Sci USA* 105:5075–5080
- Squire PG, Moser P, O’Konski CT (1968) The hydrodynamic properties of bovine serum albumin monomer and dimer. *Biochemistry* 7:4261–4272
- Stradner A, Sedgwick H, Cardinaux F, Poon WCK, Egelhaaf SU, Schurtenberger P (2004) Equilibrium cluster formation in concentrated protein solutions and colloids. *Nature* 432:492–495
- Velev OD, Kaler EW, Lenhoff AM (1998) Protein interactions in solution characterized by light and neutron scattering: comparison of lysozyme and chymotrypsinogen. *Biophys J* 75:2682–2697





### 4.3. Coupling and decoupling of rotational and translational diffusion of proteins under crowding conditions

After having focused on methodological consequences of crowding for NMR spin relaxation analysis of proteins, with direct evidence that the RACF deviates from a single exponential, the detailed impact of crowding on the Brownian diffusion rates will be addressed now. A strong decoupling of rotational and long-time translational diffusion has been shown for  $\alpha$ B-crystallin, with indications that rotational diffusion is rather sensitive to the solvent viscosity than to the solution viscosity. The behavior of  $\alpha$ B-crystallin can by no means be generalized as being representative for all proteins, especially as  $\alpha$ B-crystallin is a spherically symmetric assembly while many other proteins are not. Pronounced protein-protein and protein-solvent-protein interactions *may* give rise to a viscous drag of rotational diffusion that follows a similar concentration dependence as the macroscopic solution viscosity. Opposed to that, the concept that rotational diffusion reports on a solvent-related microviscosity is widely spread in *in vivo* protein science [Luby-Phelps, 2000].

To address this issue, the rotational correlation times reported in the previous study will be combined with viscosity data and both long and short-time translational diffusion coefficients received from PFG NMR and neutron scattering literature data [Roosen-Runge et al., 2011], respectively. Additional support comes from independent data by polarized FCS that stress the accuracy of the previously determined rotational (and newly determined translational) diffusion coefficients from NMR. The approved data on  $\alpha$ B-crystallin and BSA will be compared to the results on the LYZ solutions. SAXS measurements on all three protein solutions reflect pronounced differences in the inter-protein interactions among the samples, and are in line with the findings on the protein Brownian motion.

From a combined analysis of the data, a clear, comprehensive physical picture on the Brownian motion of concentrated protein solutions emerges. Intermolecular charge-charge interactions will be identified to significantly retard rotational diffusion, accompanied by a large impact of electro-kinetic effects on the concentration dependence of rotational diffusion. The findings presented in the following are strongly corroborated by the study of Liu et al. [2011].<sup>1</sup> In their work, a shared concentration dependence of short-time and long-time translational diffusion of LYZ monomers is depicted, with the evidence for a dynamic molecular ordering among neighboring molecules.

*Author contributions:* K.S and A.K. designed research; M.R., M.H. and A.K performed NMR measurements; M.R. analyzed the NMR data and did the viscosity measurements; M.O. performed and analyzed the SAXS and FCS experiments; S.L. provided the  $\alpha$ B-crystallin samples and the SDS PAGE data (see also SI); M.R., M.O., E.R., J.B., A.K and K.S. interpreted the findings; M.R. wrote the first draft of the paper; K.S., M.R. and A.K. wrote the paper in final form with further refinements by all coauthors.

---

<sup>1</sup>Unfortunately, we became aware of this article only after publication of the following paper.

The following article [M. Roos et al. Coupling and decoupling of rotational and translational diffusion of proteins under crowding conditions. *J. Am. Chem. Soc.* 138: 10365-10372, 2016] is reprinted with permission. Copyright 2016 American Chemical Society. The link back to the article on the publisher's website is <http://dx.doi.org/10.1021/jacs.6b06615> .  
No changes were made.

# Coupling and Decoupling of Rotational and Translational Diffusion of Proteins under Crowding Conditions

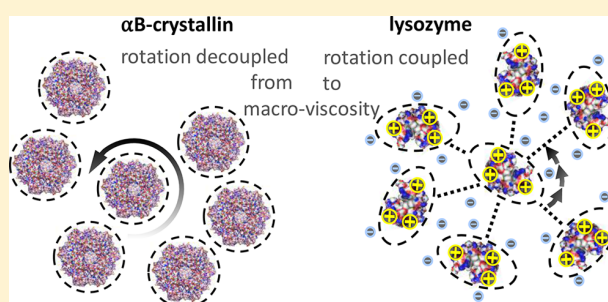
Matthias Roos,<sup>†,§</sup> Maria Ott,<sup>†,§</sup> Marius Hofmann,<sup>‡</sup> Susanne Link,<sup>†</sup> Ernst Rössler,<sup>‡</sup> Jochen Balbach,<sup>†</sup> Alexey Krushelnitsky,<sup>\*,†</sup> and Kay Saalwächter<sup>\*,†</sup>

<sup>†</sup>Institut für Physik, Martin-Luther-Universität Halle-Wittenberg, 06099 Halle (Saale), Germany

<sup>‡</sup>Experimentalphysik II, Universität Bayreuth, 95440 Bayreuth, Germany

**S** Supporting Information

**ABSTRACT:** Molecular motion of biopolymers *in vivo* is known to be strongly influenced by the high concentration of organic matter inside cells, usually referred to as crowding conditions. To elucidate the effect of intermolecular interactions on Brownian motion of proteins, we performed <sup>1</sup>H pulsed-field gradient NMR and fluorescence correlation spectroscopy (FCS) experiments combined with small-angle X-ray scattering (SAXS) and viscosity measurements for three proteins,  $\alpha$ B-crystallin ( $\alpha$ Bc), bovine serum albumin, and hen egg-white lysozyme (HEWL) in aqueous solution. Our results demonstrate that long-time translational diffusion quantitatively follows the expected increase of macro-viscosity upon increasing the protein concentration in all cases, while rotational diffusion as assessed by polarized FCS and previous multi-frequency <sup>1</sup>H NMR relaxometry experiments reveals protein-specific behavior spanning the full range between the limiting cases of full decoupling from ( $\alpha$ Bc) and full coupling to (HEWL) the macro-viscosity. SAXS was used to study the interactions between the proteins in solution, whereby it is shown that the three cases cover the range between a weakly interacting hard-sphere system ( $\alpha$ Bc) and screened Coulomb repulsion combined with short-range attraction (HEWL). Our results, as well as insights from the recent literature, suggest that the unusual rotational–translational coupling may be due to anisotropic interactions originating from hydrodynamic shape effects combined with high charge and possibly a patchy charge distribution.



## INTRODUCTION

Inside cells, macromolecules occupy 20–40% of the cytoplasmic volume,<sup>1,2</sup> providing an environment in which the mean distance between neighboring particles is similar to their size. Highly concentrated solutions of proteins and other organic molecules mimicking the cell interior are usually referred to as crowded environment.<sup>3–5</sup> Crowding affects many aspects of cellular function and organization,<sup>6,7</sup> including biochemical reactions, enhanced protein refolding rates, and the stabilization or destabilization of the compact folded states.<sup>8–11</sup> In particular, crowding sensitively affects *in vivo* molecular motion,<sup>12</sup> where protein Brownian dynamics is rather complex due to the usually non-spherical shape of the globule and its complex non-symmetric electrostatic interactions.

The Brownian dynamics of concentrated particle suspensions can be quantitatively described by mean-field models that depend only on the overall volume fraction  $\phi$  of the dispersed particles. This leads to the generalized Stokes–Einstein (GSE) and generalized Stokes–Einstein–Debye (GSED) relationships for the long-time translational diffusion coefficient  $D^L$  and the rotational correlation time  $\tau_r$ , respectively:

$$D^L(\phi) = \frac{k_B T}{6\pi\eta(\phi)R_H} \quad (1)$$

$$\tau_r(\phi) = \frac{4\pi\eta(\phi)R_H^3}{3k_B T} \quad (2)$$

where  $R_H$  and  $k_B T$  denote the hydrodynamic radius and the thermal energy, respectively. For the generalized forms, the solvent viscosity  $\eta_0$  is merely replaced by the macroscopic dispersion viscosity  $\eta(\phi)$ ; for a critical discussion see ref 13. Such a treatment implies that a macromolecular solute is surrounded by an effective, continuous medium—a situation that is, at first glance, violated under crowding conditions. However, it is well established that the GSE relation for translational diffusion, eq 1, holds for concentrated hard-sphere (HS)<sup>14–16</sup> and even soft colloid systems,<sup>17</sup> but does not necessarily hold for charge-stabilized colloids.<sup>13,15</sup> Crowded proteins represent, in general, a case in-between these limiting situations.

Received: June 27, 2016

Published: July 19, 2016

The validity of the GSED relationship, eq 2, has not yet unambiguously been assessed for proteins. In the presence of neighboring particles, rotational diffusion depends on the particle shape<sup>18</sup> and on intermolecular electrostatic interactions<sup>19</sup> that are, concerning proteins, usually of rather complex nature due to a non-symmetric charge distribution within the protein. Despite the high interest in crowding effects for understanding *in vivo* behavior of proteins, the effect on Brownian dynamics remains little studied and controversial. Notably, recent combined studies of translational and rotational diffusion of proteins contradict each other with regards to the effect of crowding: it remains unclear whether rotational diffusion becomes less retarded<sup>20</sup> than translational diffusion of the protein, or *vice versa*.<sup>21</sup> Thus, assessing the potential applicability of both the GSE and GSED relationships is of high relevance to ultimately link microscopic observables with biological function.

Recently we have undertaken a detailed nuclear magnetic resonance (NMR) study of the Brownian dynamics of the eye-lens protein  $\alpha$ B-crystallin ( $\alpha$ Bc) over a wide range of concentrations.<sup>22</sup> We found a pronounced decoupling between translational and rotational diffusion: while the slow-down of translational diffusion upon increasing the protein concentration perfectly matched the increase in macro-viscosity, rotational diffusion was almost unaffected. This effect can be explained in terms of a “caging” of the probed molecule by surrounding ones,<sup>23–25</sup> and is generally linked to the phenomenon of the colloidal glass transition.<sup>26–28</sup> Indeed, a HS-like glass transition in eye-lens  $\alpha$ -crystallin solutions was shown recently.<sup>29</sup> Notably, the stable  $\alpha$ Bc assembly has a rather symmetric, quasi-spherical shape<sup>30</sup> as it consists of several symmetrically arranged monomers,<sup>30</sup> such that it resembles an almost perfect hard-sphere particle,<sup>31</sup> while other proteins may not. Thus, the behavior of  $\alpha$ Bc can hardly be taken as universal.

In the present work, we extend our studies by two other proteins, bovine serum albumin (BSA) and hen egg white lysozyme (HEWL). We show that the coupling or decoupling of rotational and long-time translational diffusion under crowding conditions is protein-specific and appears to be related to the specific type of interactions between neighboring proteins.

Protein molecular mobility is characterized here by both NMR spectroscopic measurements of translational and rotational<sup>22,32</sup> diffusion and independent measurements of the same quantities by polarized fluorescence correlation spectroscopy (FCS). These data are complemented by measurements of the macroscopic viscosity and the intermolecular interactions by capillary rheology and small-angle X-ray scattering (SAXS), respectively. Short-time translational diffusion coefficients from neutron-scattering literature data<sup>33</sup> are also taken into account. The combined results provide a comprehensive picture on the Brownian dynamics of proteins under “self-crowding” conditions.

## MATERIALS AND METHODS

**Samples.** Native  $\alpha$ -crystallin is a spherical assembly of two homologous proteins,  $\alpha$ A-crystallin ( $\alpha$ Ac) and  $\alpha$ Bc, each of a monomer molecular mass of  $\sim$ 20 kDa. The  $\alpha$ -crystallin complex has a molecular mass distribution from 500 to 1000 kDa, with the average mass around 800 kDa. Subunit exchange occurs on the time scale of minutes<sup>34</sup> and is much slower than the time scale of our experiments. Here, we rely on our previous data reporting on pure human  $\alpha$ Bc in buffer solution. In fact, pure  $\alpha$ Bc has very similar properties to the

mixture of  $\alpha$ Ac and  $\alpha$ Bc. For details, also concerning recombinant  $\alpha$ Bc purification, see ref 22.

HEWL from chicken egg white and fatty acid-free BSA were obtained from Sigma-Aldrich (product numbers 62970 and A7030, respectively) as lyophilized powders and dissolved in D<sub>2</sub>O to keep the water NMR signal low. Via lyophilizing and dissolving the protein solution once again in D<sub>2</sub>O, residual water proteins were further reduced. No buffer was added to ensure almost unscreened electrostatic interactions. The pD obtained was pD 3.8 for HEWL and pD 7.0 for BSA (isoelectric points of pH 11.35 and pH 4.7, respectively). No significant pD changes (more than 0.1–0.2) were observed upon varying the protein concentration. Under these conditions, HEWL (14.4 kDa) is a strongly charged monomeric protein<sup>32</sup> soluble up to high concentrations. BSA consists of monomers (66.4 kDa) and about 50% long-time stable oligomers of different sizes.<sup>32,35</sup>

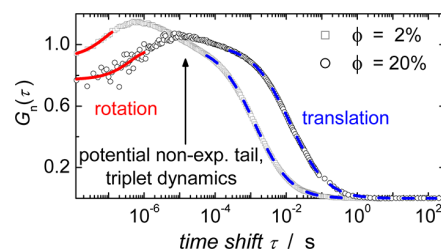
**Experiments.** Translational diffusion coefficients were determined using a Bruker Avance II spectrometer with a <sup>1</sup>H resonance frequency of 400 MHz, using a Diff60 probehead. Pulsed field gradient (PFG) NMR diffusion decays were obtained by use of the stimulated echo technique applying bipolar gradient pulses, and were fitted by

$$A(g) = A(0) \exp(-\gamma^2 g^2 D^L \delta^2 (\Delta - \delta/3)) \quad (3)$$

in which  $A(g)$  is the (integral) signal intensity in dependence of the gradient strength  $g$ , and  $\gamma$  is the <sup>1</sup>H gyromagnetic ratio.  $\delta$  and  $\Delta$  denote the fixed gradient pulse duration and diffusion time, respectively. Exemplary PFG NMR diffusion decays for  $\alpha$ Bc, BSA, and HEWL are shown in ref 22 and the Supporting Information (SI1); data for BSA and HEWL were measured within this work. Translational protein diffusion as characterized by PFG NMR relies on diffusion times of a few tens of milliseconds, thus providing translational displacements in the  $\mu$ m range. Hence, PFG NMR probes protein translational diffusion in the long-time limit.

NMR data for rotational diffusion rely on longitudinal relaxation rates ( $R_1$ ) measured on a field-cycling instrument and/or rotating-frame ( $R_{1\rho}$ ) and transverse ( $R_2$ ) relaxation rate measurements of the integral <sup>1</sup>H signal. The derived rotational correlation times are taken from our previous publications; see refs 22 and 32.

Rotational correlation times  $\tau_r$  and translational diffusion times  $\tau_D$  were also determined by polarized FCS with alternating orthogonal, linearly polarized excitation. We used a home-built setup similar to the one described in ref 36; see SI2 for details on the setup, sample preparation, and data processing. Polarized FCS probes rotations of the transition dipole moment and relies on the use of linearly polarized excitation of protein-bound dyes and separate detection of the two orthogonal emission components on a single-molecule basis. The two signals of the fluorescence components that are collinear to the excitation pulses are then cross-correlated, yielding a time correlation function; see Figure 1. Its initial rise (exponential in nature) encodes  $\tau_r$ , while its long-time decay encodes  $\tau_D$ , the time needed by the molecule to leave again the detection volume. For the latter, due to the well-known issues with focal volume calibration, we refrained from converting it into absolute values for  $D^L$ .



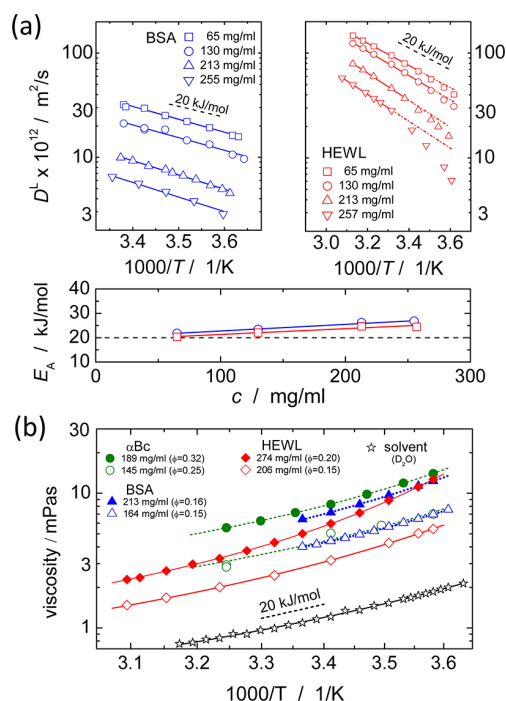
**Figure 1.** Normalized cross-correlation functions  $G_c(\tau)$  from polarized FCS with alternating orthogonal, linearly polarized excitation for BSA at two concentrations, including fits to obtain  $\tau_r$  (solid red lines) and  $\tau_D$  (dashed blue lines).

Steady-shear viscosities were measured at a shear rate of 1000 or 2000  $s^{-1}$  using the microfluid viscometer m-VROC (Rheosense Inc., CA). Upon decreasing the shear rate to 100  $s^{-1}$ , the signal-to-noise ratio decreased, yet no effect on the viscosity measured was found. Despite the rather high shear rates, the measurements still provide the macro-viscosity.

SAXS measurements were performed at 20 °C using an X-ray generator of rotating anode type with Cu target from Rigaku, operated at 2.4 kW, and a 2D gas detector. Generally, the SAXS signal  $I(q)$  can be written as a product of the form factor  $F(q)$  and the structure factor  $S(q)$ ;  $I(q) = F(q)S(q)$ . The form factor contains information regarding the 3D shape of the scattering particles and was determined at low concentrations (0.5–2 vol%) and electrostatic screening conditions. The structure factor does not depend on the shape of the particles but contains the inter-particle interactions, and is determined at different protein concentrations of BSA, HEWL, and  $\alpha$ Bc by  $S(q) = I(q)/I_0(q) \cdot c_0/c$ . For details, see S13.

## RESULTS

Figure 2 presents the temperature dependence of long-time translational diffusion (a) and viscosity (b). For both cases and all proteins, we observe activation energies ( $E_A$ ) close to 20 kJ/mol that increase only slightly with concentration, indicating that both quantities are largely governed by the solvent viscosity (water). This important result also indicates that *transient* or crowding-induced binding among the proteins is of little relevance, as such an effect would lead to a significantly increased apparent  $E_A$  for translational diffusion.<sup>37</sup> At higher concentration, however, HEWL exhibits non-Arrhenius (Vogel–Fulcher-like) behavior, reflecting increased inter-particle interactions.



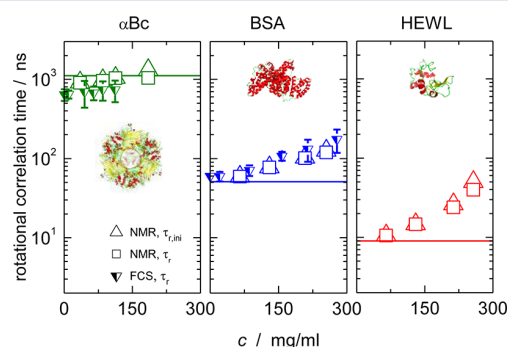
**Figure 2.** Temperature dependence of long-time translational diffusion (a) and viscosity (b) for HEWL, BSA, and  $\alpha$ Bc. Black dashed lines indicate the slope corresponding to  $E_A = 20$  kJ/mol. For HEWL translational diffusion, only the high-temperature region was used to estimate  $E_A$  as plotted in the lower panel in (a). The  $\alpha$ Bc viscosity data were already published in ref 22.

The data for rotational diffusion from multi-frequency relaxometry published in our previous publications<sup>22,32</sup> show the same trends in  $E_A$  as discussed above for translational diffusion and viscosity. A detailed comparison of all quantities is deferred to the Discussion section. It should just be noted that the molecular tumbling times  $\tau_r$ , as reported in ref 22 and the Supporting Information of ref 32, are subject to a potentially large systematic error when neglecting the, at higher concentrations, increasingly nonexponential tailed character of the tumbling correlation function (TCF) in the NMR relaxation data analysis. We have so far used an ad-hoc phenomenological approach, representing the unknown complex TCF as a superposition of two exponentials,  $C_{2exp}(t)$ , with a minority component featuring a much slower decay time constant  $\tau_s$  and comparably low amplitude  $a_s$ . In order to assess the influence of this “slow tail” on the reported rotational correlation time, and to enable a comparison with the value from FCS, we also consider an apparent correlation time defined as the inverse of the initial slope of the fitted apparent TCF:

$$\tau_{r,ini} = \left\langle \frac{dC_{2exp}(t)}{dt} \right\rangle_{t=0}^{-1} = \left[ \frac{1 - a_s}{\tau_r} + \frac{a_s}{\tau_s} \right]^{-1} \quad (4)$$

It is important to note that  $\tau_r$  from FCS, due to the restriction of the exponential fit to the initial time range of the FCS cross-correlation function  $G_n(\tau)$ , provides the same quantity as given by eq 4.

Figure 3 compares  $\tau_{r,ini}$  with  $\tau_r$  (the apparent primary tumbling time) determined by both NMR and FCS. Despite its large influence on the NMR data analysis, the apparent “slow tail” is thus demonstrated to have a nearly negligible influence on  $\tau_{r,ini}$  in the studied concentration range. Additional uncertainty arises from the non-spherical shape of the protein,<sup>38</sup> but as shown in S14, considering tensorial rotational diffusion has little influence on the fitted absolute value of  $\tau_r$  and its concentration dependence. In view of the potentially large uncertainties related to a complex and not necessarily multi-exponential overall correlation function,<sup>39,40</sup> the coincidence between NMR and FCS data, as well as the agreement with the values estimated on the basis of  $R_H$  from translational diffusion

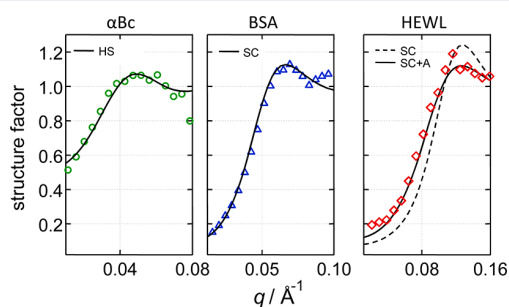


**Figure 3.** Comparison of tumbling times determined by NMR relaxometry ( $\tau_{r,ini}$  and  $\tau_r$ ; the symbol size reflects the experimental uncertainty) and FCS for the three proteins. The solid lines indicate approximate dilute-limit values calculated from eq 2 using  $R_H$  calculated from eq 1, using  $D^L$  from PFG NMR at the lowest concentrations. Within each diagram a visual presentation of the protein (not to scale) is shown based upon Protein Data Bank structures (PDB IDs 2YGD, 4F5S, and 1LYZ).

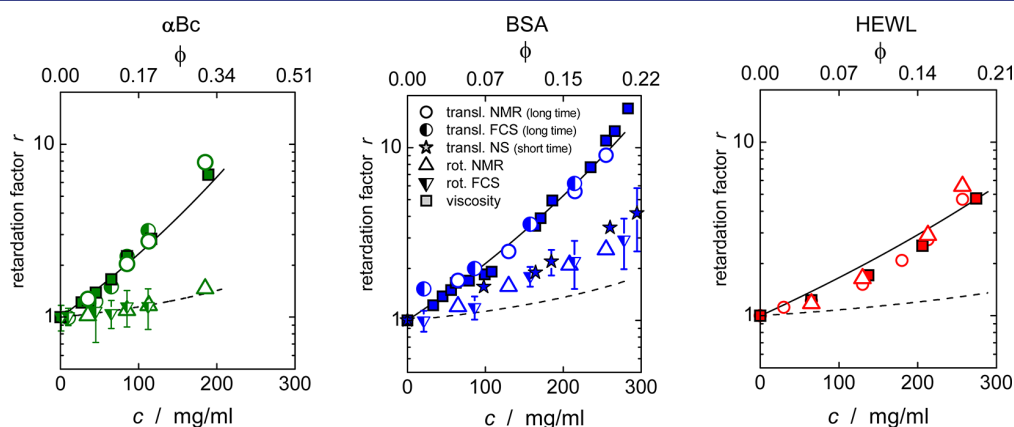
in the most dilute cases, represents a relevant finding. Note that a (possibly non-exponential) long-time tail of the TCF would contribute to the FCS cross-correlation function at intermediate times, but is not reliably detectable due to its low amplitude and additional contributions from triplet dynamics and the onset of translational diffusion (see Figure 1).

The deviation of  $\tau_r$  from NMR and FCS for  $\alpha$ Bc can be explained by the polydispersity of this protein, which is taken into account in different ways in NMR and FCS experiments; see S15. This deviation is relevant mainly for dilute-limit data and does not challenge any conclusions on crowding effects. As to the latter, from the data in the given semi-logarithmic representation, we mainly take that the relative increase of  $\tau_{r,(ini)}$  with concentration differs significantly among the samples, as analyzed further below.

In order to characterize and compare directly the nature of the inter-particle interactions, we have measured SAXS data for the three proteins under study, see Figure 4 and S13. This information is of course subject to the limitation that an isotropic average is obtained. The structure factors exhibit qualitative differences. They suggest that  $\alpha$ Bc assemblies behave



**Figure 4.** Protein interaction strength as assessed by SAXS structure factors for a volume fraction of  $\phi = 10\%$ . The fits to hard-sphere (HS), screened Coulomb (SC), and SC plus short-range attraction (SC+A) models for  $\alpha$ Bc, BSA, and HEWL, respectively, are discussed in the next section.



**Figure 5.** Retardation of long-time translation (circles) and rotation (triangles) as compared to viscosity (squares) in dependence of the protein concentration  $c$ . The retardation factors are the respective measured quantities normalized against their low-concentration limits. For NMR, they are normalized to an extrapolated value (see S16), while for FCS we have experimental results at virtually zero (nM) concentration. If not indicated by error bars, experimental uncertainties are of the order of or smaller than the symbol size. Short-time translational diffusion data observed for BSA by neutron scattering (stars) are reproduced from ref 33. Data of  $\alpha$ Bc and all NMR rotational diffusion data are taken from our previous publications.<sup>22,32</sup> The solid lines correspond to predictions of long-time translational diffusion, eq 5, and are based upon an effective HS volume fraction,  $\phi_{HS} = k\phi$  ( $k = 1, 2.1, \text{ and } 1.5$  for  $\alpha$ Bc, BSA, and HEWL, respectively). The dashed lines are HS predictions for rotational diffusion, eq 6, using the same rescaling factor  $k$  as obtained for translational diffusion.

like hard spheres, while BSA and HEWL are dominated by Coulomb interactions. In the latter case, short-range attractive interactions are to be considered as well. Details on the corresponding analyses, and the concentration effects on the structure factors, will be discussed in the next section.

## DISCUSSION

**Long-Time Translational Diffusion Scales with Macroscopic Viscosity.** For a direct comparison of the concentration dependence of viscosity and translational and rotational diffusion in one and the same plot, we compare inverse reduced diffusion coefficients  $(D^L/D_0^L)^{-1}$ , reduced tumbling or translational diffusion times  $(\tau_{r/D}/\tau_{r/D,0})$ , and normalized viscosities  $(\eta/\eta_0)$  at 20 °C, all referred to as retardation factor  $r$ . In this way, Figure 5 shows  $r$  with increasing extent of crowding (see S16 for absolute values) and presents the central result of this work. Note in particular that the slope of the data is independent of the chosen reference concentration, as the retardation factors are plotted on a semi-logarithmic scale. As seen from Figure 5, for all cases the reduced long-time translational self-diffusion coefficients match the increase of viscosity with concentration, i.e.,  $D^L(\phi)/D_0^L = \eta_0/\eta(\phi)$ . This demonstrates the applicability of the GSE relation, eq 1.

In Figure 5, the steric volume fraction was defined as  $\phi = c\vartheta$ , where  $\vartheta$  is the specific volume of the protein ( $\vartheta = 1.7 \text{ mL/g}$ ,<sup>41</sup>  $0.735 \text{ mL/g}$ ,<sup>33</sup> and  $0.702 \text{ mL/g}$ <sup>42</sup> for  $\alpha$ Bc, BSA, and HEWL, respectively). See S17 for the actual data. The large specific volume for  $\alpha$ Bc (i.e., low density) results from its high molecular weight combined with the deviation of most “globular” proteins from compact-globule scaling:  $V \propto M^{1.2}$  rather than  $V \propto M^3$ .<sup>22,43</sup>

The applicability of the GSE equation as observed in the present three cases is in full accordance with established colloid-science concepts: during translational diffusion in the long-time limit, the particle’s trajectory averages over many different configurations of its local surrounding, which allows for a description of the inhomogeneous environment acting in terms of an effective medium of (zero-shear) viscosity  $\eta(\phi)$ .<sup>13</sup>

In the literature, both accordance<sup>21,44</sup> and mismatch<sup>20,21</sup> between translational diffusion and viscosity were reported. It is important to note that, in studies reporting on a mismatch, tracer and crowding agent were different proteins, or even synthetic polymer crowders (Ficoll, polyethylene glycol) were used. In such studies, translational diffusion measurements solely report on the tracer species, whereas viscosity measurements are strongly dominated by the specific interactions among the crowder molecules due to their much higher volume fraction. For diffusion of the tracer proteins mixed with *other* proteins, transient binding may be important.<sup>37</sup>

Quantitatively, the slow-down of long-time translational diffusion has been addressed via HS models,<sup>45</sup> resulting in

$$\frac{D^L(\phi_{\text{HS}})}{D_0^L} \cong \frac{(1 - \phi_{\text{HS}})^3}{1 + (3/2)\phi_{\text{HS}} + 2\phi_{\text{HS}}^2 + 3\phi_{\text{HS}}^3} \quad (5)$$

Proteins are subject to intermolecular interactions beyond the pure HS potential, in particular through electrostatic interactions. However, the size of the protein can be re-adjusted by use of an *effective* HS radius, corresponding to an effective HS volume fraction  $\phi_{\text{HS}} = k\phi$ ,  $k \geq 1$ . Fits to eq 5 shown in Figure 5 correspond to  $k = 1$ ,  $k = 2.1$ , and  $k = 1.5$  for  $\alpha\text{Bc}$ , BSA, and HEWL, respectively. Note that  $k$  is the only adjustable parameter;  $k = 1$  proves that translational diffusion of  $\alpha\text{Bc}$  follows HS behavior on the basis of the steric volume fraction, indicating only rather weak inter-protein interactions.

**(De)coupling of Rotational Diffusion from Long-Time Translation and Macro-viscosity Is Protein Specific.** For the three proteins investigated, the NMR and FCS results for the apparent rotational tumbling times  $\tau_{\text{r,ini}}$  are now compared to translational diffusion and the macro-viscosity; see again Figure 5. Notably, the (de)coupling of rotational diffusion from translational diffusion and viscosity is evidently protein specific. Such behavior is in line with an increased importance of protein-specific intermolecular interactions and shape effects for rotational dynamics.<sup>18,46,47</sup> Note that, in our experiments, rotation is never observed to be more retarded than translation.

Both fluorescence spectroscopy data<sup>20</sup> and computer simulations<sup>47</sup> have revealed a decoupling between translational and rotational diffusion that is in full accordance with the  $\alpha\text{Bc}$  results.<sup>22</sup> In contrast, an NMR study of hetero-crowding<sup>21</sup> reported a decoupling in the opposite sense, i.e., rotational diffusion becoming more retarded than translational diffusion and viscosity. Colloid theories<sup>24,25</sup> suggest that rotation in concentrated solutions is expected either to be less affected than or to scale with translation, provided that the concentration of dispersed colloids is well below the onset of the colloidal glass transition. Likely, the unexpected finding of ref 21 results from estimating rotational correlation times solely from site-resolved NMR  $T_1/T_2$  ratios<sup>48</sup> at a single resonance frequency, neglecting the non-exponential nature of the TCF. As has been shown recently,<sup>32</sup> such a treatment can lead to an erroneous estimation of the tumbling correlation time, especially at high concentrations. Systematic deviations are, however, hardly detectable by the traditional approach.<sup>32</sup>

Coupling or decoupling of rotational from long-time translational diffusion and the relationship of these two quantities to the macroscopic viscosity can be generally assigned to the presence or absence of correlated motions among neighboring particles. For long-time diffusion, multiple independent encounters with other particles, that may have to rearrange cooperatively at high concentrations,<sup>13</sup> lead to its

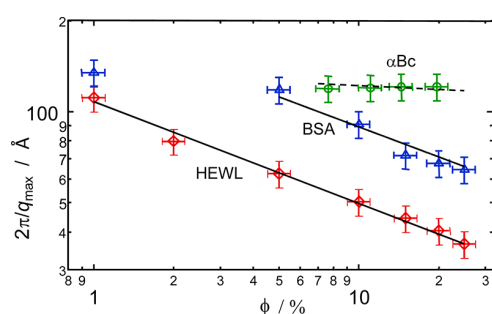
dependence on an average friction corresponding to the macroscopic zero-shear viscosity  $\eta(\phi)$ . In contrast, on the time scale of rotational diffusion (0.01–0.1  $\mu\text{s}$  for HEWL, and  $\sim 1 \mu\text{s}$  for  $\alpha\text{Bc}$ ) the protein's local surrounding neither undergoes substantial reconfiguration, nor do particle collisions appreciably affect the rotational dynamics. In the absence of specific interactions rotational diffusion is almost unhindered.<sup>38</sup> Specifically, even a non-spherical object such as HEWL subject to *only* excluded-volume effects was shown to be able to rotate rather freely within its cage formed by the surrounding particles.<sup>38</sup> Still, local hydrodynamic effects<sup>33</sup> mediated via particle–solvent interactions retard rotational diffusion and account for a measurable but rather weak concentration dependence. In fact, we find that the slow-down of rotational diffusion of  $\alpha\text{Bc}$  with increasing concentration is again quantitatively reproduced by applying a corresponding HS model,<sup>18,49</sup>

$$\frac{\tau_{\text{r}}(\phi_{\text{HS}})}{\tau_{\text{r},0}} \cong [1 - 0.631\phi_{\text{HS}} - 0.762\phi_{\text{HS}}^2]^{-1} \quad (6)$$

without requiring rescaling of the effective HS volume fraction ( $k = 1$ , as for translational diffusion); see Figure 5. Thus, the viscosity experienced by rotation is closer to that of the solvent than to the macroscopic viscosity, as often referred to as micro-viscosity. More precisely, as a short-time quantity rotational diffusion is considered to be sensitive to the viscosity determined in the limit of high shear rates,<sup>18,50</sup> usually denoted as  $\eta_{\infty}(\phi)$ . Hence, a decoupling of rotational from long-time translational diffusion is to be expected as long as  $\eta_{\infty}(\phi) \neq \eta(\phi)$ , as is well established for spherical colloids.<sup>13,18,24,25</sup> For  $\alpha\text{Bc}$  and, to a lesser extent, BSA, we observe such a behavior.

**Rotational vs Short-Time Translational Diffusion.** In contrast to translational diffusion measured by PFG NMR, translational diffusion as detected by neutron scattering is measured on short length scales, corresponding to short observation times of  $0.3 \text{ ns} \leq \tau \leq 5 \text{ ns}$ .<sup>33</sup> The corresponding translational root-mean-square displacements amount to about 10 Å or even less, i.e. translational dynamics is probed solely within the cage formed by neighboring molecules. This situation corresponds to the time scale of rotational diffusion. Both rotational and *short-time* translational diffusion are considered, as mentioned above, to be related to  $\eta_{\infty}(\phi)$ ; hence, one may expect a similar concentration dependence for these two diffusion processes. To address this point, Figure 5 also presents the short-time translational diffusion data for BSA measured by quasielastic neutron backscattering.<sup>33</sup> Indeed, the concentration dependence of short-time translational diffusion coincides with our data on rotational diffusion within the experimental uncertainty. This coincidence also reinforces that NMR relaxometry and FCS provide reliable results regardless of the polydispersity of the protein solution.

**Role of Protein–Protein Interactions.** HEWL, being a strongly charged protein under our conditions ( $\text{pD} = 3.8$ , no buffer), behaves qualitatively differently as compared to  $\alpha\text{Bc}$  and BSA; its rotation is fully coupled to long-time translational diffusion and macroscopic viscosity. Here, when using the *same* effective HS volume fraction as for translational diffusion, the hard-sphere model, eq 6, clearly fails in accounting for the concentration dependence of rotational diffusion (Figure 5). Instead, approaching the experimental data requires a rescaling as large as  $k = 3.7$  (compared to  $k = 1.5$  for translational diffusion). Moreover, HS modeling intrinsically predicts a



**Figure 6.** Protein interactions as assessed by SAXS experiments. The plotted inverse maximum positions of the structure factor ( $2\pi/q_{\max}$ ) in dependence of the volume fraction  $\phi$  are expected to decrease according to a power law for strong repulsive systems ( $q_{\max}^{-1} \propto \phi^{-1/3}$ , solid lines; BSA at  $\phi = 1\%$  was excluded in the fit). The value of  $2\pi/q_{\max}$  then relates to the average center to center distance between first neighbors. However, in case of screened, hard-sphere-like particles the concentration dependence of  $2\pi/q_{\max}$  is strongly reduced (the dashed line gives the prediction for spheres with  $R = 58 \text{ \AA}$ ).

decoupling of rotational from long-time translational diffusion under crowding conditions, see above. Thus, with regard to the same concentration dependence of long-time translational and rotational diffusion, effective-sphere behavior cannot even qualitatively describe the HEWL experimental data. Our finding also stands in stark contrast to results of Brownian dynamics simulations of crowded HEWL solutions when shape effects and only excluded-volume interactions are considered.<sup>38</sup> Consequently, specific interactions must be of major importance and evidently lead to a correlation of the tumbling of a single HEWL molecule with the dynamics of its surroundings.

Our SAXS data (Figure 4) provide direct evidence of the different nature of the interactions present in the three protein solutions studied. A qualitative but easily assessed piece of information is the concentration dependence of the low- $q$  maximum of the structure factor,  $q_{\max}$  (Figure 6). For a simple HS liquid such as  $\alpha$ Bc,<sup>29</sup> the maximum of the structure factor does not scale with the mean center-to-center distance of the particles, leading to a constant  $q_{\max}$ . In contrast, for long-range Coulomb repulsion particles tend to maximize their interparticle distances. Then,  $q_{\max}$  can be related to the inverse interparticle distance by  $q_{\max} = 2\pi/R_{cc}$  and scales as  $\sim \phi^{1/3}$ , as is observed for both HEWL and BSA.

Qualitative differences in the charge–charge interactions for our three samples are further documented by the SAXS structure factors (Figure 4). For  $\alpha$ Bc, again, the HS behavior is supported by the perfect agreement of the experimentally derived structure factor with HS model predictions. In contrast, the structure factor of BSA agrees with the prediction of a screened Coulomb (SC) potential. In case of HEWL, the latter fit still exhibits systematic deviations at small  $q$ , indicating the presence of additional short-range attractive forces. This is demonstrated by comparison of the experimentally derived structure factor with the structure factor prediction of a SC potential and a SC potential with an additional short-range attractive Yukawa potential. The obtained ranges of about 39 and 2 Å for the long-range repulsive and short-range attractive interactions, respectively, are consistent with literature<sup>51,52</sup> and match the behavior of patchy charged colloids.<sup>53,54</sup>

The different nature of intermolecular interactions in  $\alpha$ Bc, BSA, and HEWL solutions is also corroborated by the small but

significant differences in the temperature dependence of viscosity and long-time translational diffusion shown in Figure 2. The remarkable non-Arrhenius behavior and the somewhat increased (apparent) activation energy of HEWL at low temperatures as compared to that of pure water reflects significant inter-protein interactions and correlated motions, which ultimately lead to a calorimetric glass transition at high concentrations.

As a result of the short-range attraction, HEWL has long been discussed to form transient clusters upon increasing the concentration.<sup>55,56</sup> The fact that attractive interactions are known to retard rotational diffusion beyond the limit of HS behavior<sup>57</sup> suggests that the phenomena may have a common origin. Note, however, that at all HEWL concentrations, our data on long-time translational diffusion (ms time scale) agree with the diffusion of monomers, which is not found in systems characterized by transient clustering.<sup>37</sup> Further work is certainly necessary to explore this issue.

Since the high overall charge of the HEWL monomers leads to repulsion, this in turn leading to decoupling rather than coupling of rotation and macro-viscosity in the case of isotropic spheres,<sup>57</sup> we suggest that the behavior of HEWL may be related to either a non-uniform charge distribution or charge-enhanced hydrodynamic effects combined with shape anisotropy, to be discussed below. Note that electrostatic multipole interactions and alignment effects give rise to net attractive interactions,<sup>53</sup> providing a rationale for the short-range attraction discussed above.

The quantitative understanding of anisotropic interactions between “patchy” charged colloids as a suitable model for proteins is of substantial current interest,<sup>58</sup> yet recent reports only focus on static structural and thermodynamic properties,<sup>53,54,59</sup> in particular protein solution-phase behavior.<sup>60–63</sup> We are so far not aware of any theoretical assessment of tumbling motion in such cases. The only experimental observation of rotational–translational coupling in a colloidal system was recently made in a suspension of homogeneously charged platelets.<sup>64</sup> Also in this case, effective-sphere models failed to explain the observation, and it was attributed to electro-hydrodynamic coupling effects.

Additional support of our hypothesis is provided by the increasingly non-exponential, tailed nature of the TCF as detected by NMR relaxometry.<sup>19,32</sup> A similar phenomenon has been described even for spherical colloids at high concentrations,<sup>18,65</sup> where it is most likely due to local concentration fluctuations. As mentioned under Results, a “slow tail” was modeled empirically by a second exponential component with a slower isotropization time  $\tau_s$  and (small) amplitude  $a_s$ . We stress that our previous interpretation of  $a_s$  in terms of a “model-free” order parameter  $S_{\text{rot}}^2$  should be considered critical, in view of the unknown shape of the TCF.<sup>39</sup> More detailed analyses are certainly required to extract physically more meaningful parameters. We just note that the parameter  $a_s$  increases with concentration, as expected,<sup>32</sup> and is significantly larger for HEWL than for BSA. The relevance of charge for the apparent “slow tail” was proven by NMR experiments on HEWL solutions at different pH.<sup>19</sup> Along this line, experiments<sup>66</sup> and simulations<sup>67</sup> have demonstrated that HEWL orients along the electric field exerted by another protein ( $\alpha$ -lactalbumin, 14 kDa).

In summary, we have discussed evidence that the observed coupling between protein rotational and translational diffusion (and also macroscopic viscosity) may be explained by shape-



and charge-related anisotropic protein–protein interactions, possibly combined with specific hydrodynamic coupling effects. With this, our results emphasize the importance of a prudent choice of the crowding agent in order to mimic *in vivo* conditions. Macromolecular crowding by flexible, possibly branched random-coil polymers such as Ficoll or polyethylene glycol results in a physically different situation compared to crowding by unevenly charged globular proteins.<sup>11,68,69</sup> At high concentrations, the random coils of flexible-polymer crowders interpenetrate, forming an entangled medium that is better described by established polymer physics concepts rather than colloid concepts based on excluded volume only.<sup>68</sup>

## CONCLUSIONS

Upon increasing the concentration of globular proteins of widely different size and interactions, PFG NMR and FCS results on translational diffusion measured on a millisecond time scale are consistent and exhibit a scaling with the macroviscosity. Such behavior confirms the wide applicability of the generalized Stokes–Einstein relation for both mono- and polydisperse protein solutions.

In contrast, rotational diffusion, as assessed complementarily by NMR relaxometry and polarized FCS, is a short-time quantity, which is sensitive to the viscosity of the micro-environment. In case of weakly interacting, near-isotropic particles it is close to the viscosity of pure solvent, with only small corrections due to local hydrodynamics. However, a non-spherical shape and/or specific anisotropic interactions lead to a correlation between the rotations of neighboring proteins, coupling the tumbling motion to the macroscopic zero-shear viscosity.

We have found that the applicability of either scenario is protein-specific and that the whole range between these limiting cases known in colloid science is covered: our results reflect both full coupling and strong decoupling between rotational and translational diffusion (HEWL and  $\alpha$ Bc, respectively), as well as an intermediate case (BSA). SAXS measurements reflecting inter-particle interactions and previous NMR data<sup>19,32</sup> emphasize the relevance of charge effects, combined with hydrodynamic coupling and transient anisotropy arising from a complex surface charge distribution and/or a non-spherical shape. This view is supported by patchy charge models that emphasize the relevance of mutual alignment effects.<sup>53,54,70</sup>

Theoretical assessments of the tumbling motion of concentrated patchy charged colloids and especially proteins are so far not available, but we hope that our work provides a stimulus to develop a more complete physical understanding. This is also important for future NMR studies, in particular of crowded proteins, where the non-exponential character of the tumbling correlation function (TCF) with its apparent “slow tail”, which likely arises from anisotropic protein–protein interactions, challenges established data analysis models.<sup>32</sup> Precise knowledge of the TCF may enable the development of physically well motivated and thus precise approaches.<sup>39,40,71</sup> Such endeavors will likely benefit from the complementarity of NMR and FCS results presented herein.

## ASSOCIATED CONTENT

### Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/jacs.6b06615.

PFG NMR and sub-ensemble effects (SI1); polarized FCS (SI2); X-ray scattering experiments (SI3); polydispersity and anisotropy effects in NMR relaxometry (SI4); size polydispersity effect on rotational diffusion of  $\alpha$ Bc (SI5); absolute values of NMR-based translational and rotational diffusion coefficients, and extrapolation to zero concentration (SI6); volume fractions and inter-particle distances (SI7) (PDF)

## AUTHOR INFORMATION

### Corresponding Authors

\*krushelnitsky@physik.uni-halle.de

\*kay.saalwaechter@physik.uni-halle.de

### Author Contributions

<sup>§</sup>M.R. and M.O. contributed equally.

### Notes

The authors declare no competing financial interest.

## ACKNOWLEDGMENTS

We thank T. Thurn-Albrecht (Martin Luther Universität Halle-Wittenberg) for providing access to the SAXS equipment and for helpful discussions. The authors are indebted to the Deutsche Forschungsgemeinschaft (DFG, SFB-TRR 102 project A08) for funding this work.

## REFERENCES

- (1) Zimmerman, S. B.; Trach, S. O. *J. Mol. Biol.* **1991**, *222*, 599.
- (2) Medalia, O.; Weber, I.; Frangakis, A. S.; Nicastro, D.; Gerisch, G.; Baumeister, W. *Science* **2002**, *298*, 1209.
- (3) Ellis, R. J. *Trends Biochem. Sci.* **2001**, *26*, 597.
- (4) Ellis, R. J. *Curr. Opin. Struct. Biol.* **2001**, *11*, 114.
- (5) Ellis, R. J.; Minton, A. P. *Nature* **2003**, *425*, 27.
- (6) Zhou, H. X.; Rivas, G. N.; Minton, A. P. *Annu. Rev. Biophys.* **2008**, *37*, 375.
- (7) Klumpp, S.; Scott, M.; Pedersen, S.; Hwa, T. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 16754.
- (8) Cheung, M. S.; Klimov, D.; Thirumalai, D. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 4753.
- (9) Miklos, A. C.; Sarkar, M.; Wang, Y.; Pielak, G. J. *J. Am. Chem. Soc.* **2011**, *133*, 7116.
- (10) Senske, M.; Törk, L.; Born, B.; Havenith, M.; Herrmann, C.; Ebbinghaus, S. *J. Am. Chem. Soc.* **2014**, *136*, 9036.
- (11) Danielsson, J.; Mu, X.; Lang, L.; Wang, H.; Binolfi, A.; Theillet, F.-X.; Bekei, B.; Logan, D. T.; Selenko, P.; Wennerström, H.; Oliveberg, M. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112*, 12402.
- (12) Ando, T.; Skolnick, J. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 18457.
- (13) Koenderink, G. H.; Philipse, A. P. *Langmuir* **2000**, *16*, 5631.
- (14) Segre, P. N.; Meeker, S. P.; Pusey, P. N.; Poon, W. C. K. *Phys. Rev. Lett.* **1995**, *75*, 958.
- (15) Banchio, A. J.; Nägele, G.; Bergenholtz, J. *J. Chem. Phys.* **1999**, *111*, 8721.
- (16) Banchio, A. J.; Bergenholtz, J.; Nägele, G. *Phys. Rev. Lett.* **1999**, *82*, 1792.
- (17) Gupta, S.; Stellbrink, J.; Zaccarelli, E.; Likos, C. N.; Camargo, M.; Holmqvist, P.; Allgaier, J.; Willner, L.; Richter, D. *Phys. Rev. Lett.* **2015**, *115*, 128302.
- (18) Koenderink, G. H.; Zhang, H.; Aarts, D. G. A. L.; Lettinga, M. P.; Philipse, A. P.; Nägele, G. *Faraday Discuss.* **2003**, *123*, 335.
- (19) Krushelnitsky, A. *Phys. Chem. Phys.* **2006**, *8*, 2117.
- (20) Zorrilla, S.; Hink, M. A.; Visser, A.; Lillo, M. P. *Biophys. Chem.* **2007**, *125*, 298.
- (21) Wang, Y.; Li, C.; Pielak, G. J. *J. Am. Chem. Soc.* **2010**, *132*, 9392.
- (22) Roos, M.; Link, S.; Balbach, J.; Krushelnitsky, A.; Saalwächter, K. *Biophys. J.* **2015**, *108*, 98.

- (23) Doliwa, B.; Heuer, A. *Phys. Rev. Lett.* **1998**, *80*, 4915.
- (24) Kim, M.; Anthony, S. M.; Bae, S. C.; Granick, S. *J. Chem. Phys.* **2011**, *135*, 054905.
- (25) Edmond, K. V.; Elsesser, M. T.; Hunter, G. L.; Pine, D. J.; Weeks, E. R. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 17891.
- (26) Pusey, P. N.; van Megen, W. *Nature* **1986**, *320*, 340–342.
- (27) Pusey, P. N.; van Megen, W. *Phys. Rev. Lett.* **1987**, *59*, 2083.
- (28) Pusey, P. N. *J. Phys.: Condens. Matter* **2008**, *20*, 494202.
- (29) Foffi, G.; Savin, G.; Bucciarelli, S.; Dorsaz, N.; Thurston, G. M.; Stradner, A.; Schurtenberger, P. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111*, 16748.
- (30) Peschek, J.; Braun, N.; Franzmann, T. M.; Georgalis, Y.; Haslbeck, M.; Weinkauff, S.; Buchner, J. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 13272.
- (31) Licinio, P.; Delaye, M. *J. Colloid Interface Sci.* **1988**, *123*, 105.
- (32) Roos, M.; Hofmann, M.; Link, S.; Ott, M.; Balbach, J.; Rössler, E.; Saalwächter, K.; Krushelnitsky, A. *J. Biomol. NMR* **2015**, *63*, 403.
- (33) Roosen-Runge, F.; Hennig, M.; Zhang, F.; Jacobs, R. M. J.; Sztucki, M.; Schober, H.; Seydel, T.; Schreiber, F. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 11815.
- (34) Bova, M. P.; Ding, L.-L.; Fung, B. K. K.; Horwitz, J. *J. Biol. Chem.* **1997**, *272*, 29511.
- (35) Squire, P. G.; Moser, P.; O’Konski, C. T. *Biochemistry* **1968**, *7*, 4261.
- (36) Loman, A.; Gregor, I.; Stutz, C.; Mund, M.; Enderlein, J. *Photochem. Photobiol. Sci.* **2010**, *9*, 627.
- (37) Rothe, M.; Gruber, T.; Gröger, S.; Balbach, J.; Saalwächter, K.; Roos, M. *Phys. Chem. Chem. Phys.* **2016**, *18*, 18006.
- (38) Dlugosz, M.; Antosiewicz, J. M. *J. Chem. Theory Comput.* **2014**, *10*, 481.
- (39) Tugarinov, V.; Liang, Z.; Shapiro, Y. E.; Freed, J. H.; Meirovitch, E. *J. Am. Chem. Soc.* **2001**, *123*, 3055.
- (40) Meirovitch, E.; Shapiro, Y. E.; Polimeno, A.; Freed, J. H. *J. Phys. Chem. A* **2006**, *110*, 8366.
- (41) Finet, S.; Tardieu, A. *J. Cryst. Growth* **2001**, *232*, 40.
- (42) Chalikian, T. V.; Totrov, M.; Abagyan, R.; Breslauer, K. J. *J. Mol. Biol.* **1996**, *260*, 588.
- (43) Liang, J.; Dill, K. A. *Biophys. J.* **2001**, *81*, 751.
- (44) Licinio, P.; Delaye, M. *J. Phys. (Paris)* **1988**, *49*, 975.
- (45) van Blaaderen, A.; Peetermans, J.; Maret, G.; Dhont, J. K. G. *J. Chem. Phys.* **1992**, *96*, 4591.
- (46) Balbo, J.; Mereghetti, P.; Hertzen, D.-P.; Wade, R. C. *Biophys. J.* **2013**, *104*, 1576.
- (47) Mereghetti, P.; Wade, R. C. *J. Phys. Chem. B* **2012**, *116*, 8523.
- (48) Kay, L. E.; Torchia, D. A.; Bax, A. *Biochemistry* **1989**, *28*, 8972.
- (49) Cichocki, B.; Ekiel-Jezewska, M. L.; Wajnryb, E. *J. Chem. Phys.* **1999**, *111*, 3265.
- (50) Nägele, G. *J. Phys.: Condens. Matter* **2003**, *15*, S407.
- (51) Tardieu, A.; Le Verge, A.; Malfois, M.; et al. *J. Cryst. Growth* **1999**, *196*, 193.
- (52) Shukla, A.; Mylonas, E.; Di Cola, E.; Finet, S.; Timmins, P.; Narayanan, T.; Svergun, D. I. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 5075.
- (53) McClurg, R. B.; Zukoski, C. F. *J. Colloid Interface Sci.* **1998**, *208*, 529.
- (54) Yigit, C.; Heyda, J.; Dzubiella, J. *J. Chem. Phys.* **2015**, *143*, 064904.
- (55) Porcar, L.; Falus, P.; Chen, W.-R.; Faraone, A.; Fratini, E.; Hong, K.; Baglioni, P.; Liu, Y. *J. Phys. Chem. Lett.* **2010**, *1*, 126.
- (56) Cardinaux, F.; Zaccarelli, E.; Stradner, A.; Bucciarelli, S.; Farago, B.; Egelhaaf, S. U.; Sciortino, F.; Schurtenberger, P. *J. Phys. Chem. B* **2011**, *115*, 7227.
- (57) Koenderink, G. H.; Lettinga, M. P.; Philipse, A. P. *J. Chem. Phys.* **2002**, *117*, 7751.
- (58) Lund, M. *Colloids Surf., B* **2016**, *137*, 17.
- (59) Blanco, M. A.; Sahin, E.; Robinson, A. S.; Roberts, C. J. *J. Phys. Chem. B* **2013**, *117*, 16013.
- (60) Kern, N.; Frenkel, D. *J. Chem. Phys.* **2003**, *118*, 9882.
- (61) Liu, H.; Kumar, S. K.; Sciortino, F. *J. Chem. Phys.* **2007**, *127*, 084902.
- (62) Gögelein, C.; Nägele, G.; Tuinier, R.; Gibaud, T.; Stradner, A.; Schurtenberger, P. *J. Chem. Phys.* **2008**, *129*, 085102.
- (63) Platten, F.; Valadez-Pérez, N. E.; Castaneda-Priego, R.; Egelhaaf, S. U. *J. Chem. Phys.* **2015**, *142*, 174905.
- (64) Kleshchanok, D.; Heinen, M.; Nägele, G.; Holmqvist, P. *Soft Matter* **2012**, *8*, 1584.
- (65) Degiorgio, B.; Piazza, R.; Jones, R. B. *Phys. Rev. E: Stat. Phys., Plasmas, Fluids, Relat. Interdiscip. Top.* **1995**, *52*, 2707.
- (66) Salvatore, D. B.; Duraffourg, N.; Favier, A.; Persson, B. A.; Lund, M.; Delage, M.-M.; Silvers, R.; Schwalbe, H.; Croguennec, T.; Bouhallab, S.; Forge, V. *Biomacromolecules* **2011**, *12*, 2200.
- (67) Kurut, A.; Persson, B. A.; Åkesson, T.; Forsman, J.; Lund, M. *J. Phys. Chem. Lett.* **2012**, *3*, 731.
- (68) Soranno, A.; Koenig, I.; Borgia, M. B.; Hofmann, H.; Zosel, F.; Nettels, D.; Schuler, B. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111*, 4874.
- (69) Holyst, R.; Bielejewska, A.; Szymanski, J.; Wilk, A.; Patkowski, A.; Gapinski, J.; Zywockinski, A.; Kalwarczyk, T.; Kalwarczyk, E.; Tabaka, M.; Ziebac, N.; Wiczorek, S. A. *Phys. Chem. Chem. Phys.* **2009**, *11*, 9025.
- (70) Grant, M. L. *J. Phys. Chem. B* **2001**, *105*, 2858.
- (71) Meirovitch, E.; Shapiro, Y. E.; Polimeno, A.; Freed, J. H. *Prog. Nucl. Magn. Reson. Spectrosc.* **2010**, *56*, 360.

#### 4.4. Transient binding accounts for apparent violation of the generalized Stokes–Einstein relation in crowded protein solutions

In the previous article anisotropic interactions arising from shape and/or charge effects were identified as a critical factor for the behavior of protein rotational diffusion under crowding conditions. Moreover, *long-time* translational diffusion has been shown to follow the same concentration dependence as the inverse macroscopic (zero-shear) viscosity. The latter relationship holds in spite of different nature of the proteins in terms of size, polydispersity and sub-ensemble effects,<sup>2</sup> and the presence or absence of pronounced charge effects. Effective-sphere modeling sufficiently quantifies the retardation of long-time translational diffusion with increasing protein concentration. This aspect is further corroborated by a study on concentrated  $\gamma$ -globulin for which an effective-sphere behavior of *short-time* translational diffusion has been found [Grimaldo et al., 2014]. From these insights it can be concluded that crowded protein translational diffusion resembles the behavior of concentrated hard-sphere suspensions except for the potential necessity of a re-scaled apparent volume fraction of the dispersed particles.

The clarity of these results stands in stark contrast to the diversity in current literature on the same topic. As a function of the protein solution studied (synthetic crowders giving rise to entanglement effects are not considered here), not only congruency [Wang et al., 2010] but also a mismatch [Zorilla et al., 2007; Wang et al., 2010] of long-time translational diffusion with the solution viscosity has been reported. Thus, one may raise the question: What makes the difference between the protein solutions studied here and those for which an invalidity of the GSE relationship has been obtained? Since polydispersity and sub-ensemble effects were shown not to account for a distinct concentration dependence of long-time translational diffusion and the macroscopic viscosity, there *must be* another critical factor besides time scale effects and anisotropic, charge-related interactions that determines the behavior of crowded protein Brownian motion. Seemingly, features accounting for the invalidity of the GSE relationship are absent in the single-protein solutions studied so far. Unlike the concentrated solutions of  $\alpha$ B-crystallin, BSA and LYZ, in studies reporting on the *invalidity* of the long-time GSE equation protein *mixtures* were investigated. Thus it appears worthwhile to now study translational diffusion of proteins in concentrated mixtures, and to compare these results to measurements on the same proteins in single-protein solutions. In addition, the diffusivity of solvent and small co-solute molecules will be addressed.

*Author contributions*<sup>\*\*</sup>: M.Rs. designed research; M.Re. and T.G. prepared samples; M.Re. and, in minor parts, M.Rs., S.G. and T.G. performed measurements, with S.G. setting up the HSQC measurements; M.Re. analyzed the data; M.Rs., M.Re., J.B. and K.S. interpreted the findings; M.Rs. wrote the paper with refinements by K.S. and all coauthors.

---

<sup>2</sup>See also the SI of the preceding paper

<sup>\*\*</sup>M.Re.: Maik Rothe; M.Rs.: Matthias Roos

The following article [M. Rothe et al. Transient binding accounts for apparent violation of the generalized Stokes-Einstein relation in crowded protein solutions. *Phys. Chem. Chem. Phys.* 18: 18006-18014, 2016] is reproduced by permission of the PCCP Owner Societies. The link back to the article on the publisher's website is <http://dx.doi.org/10.1039/c6cp01056c> .  
No changes were made.

CrossMark  
click for updatesCite this: *Phys. Chem. Chem. Phys.*,  
2016, **18**, 18006

# Transient binding accounts for apparent violation of the generalized Stokes–Einstein relation in crowded protein solutions

M. Rothe,<sup>†</sup> T. Gruber, S. Gröger, J. Balbach, K. Saalwächter\* and M. Roos\*

The effect of high concentration, also referred to as crowding conditions, on Brownian motion is of central relevance for the understanding of the physical, chemical and biological properties of proteins in their native environment. Specifically, the simple inverse relationship between the translational diffusion coefficient and the macroscopic solution viscosity as predicted by the generalized Stokes–Einstein (GSE) relation has been the subject of many studies, yet a consensus on its applicability has not been reached. Here, we use isotope-filtered pulsed-field gradient NMR to separately assess the  $\mu\text{m}$ -scale diffusivity of two proteins, BSA and an SH3 domain, in mixtures as well as single-protein solutions, and demonstrate that transient binding can account for an apparent violation of the GSE relation. Whereas GSE behavior applies for the single-protein solutions, it does not hold for the protein mixtures. Transient binding behavior in the concentrated mixtures is evidenced by calorimetric experiments and by a significantly increased apparent activation energy of diffusion. In contrast, the temperature dependence of the viscosity, as well as of the diffusivity in single-component solutions, is always dominated by the flow activation energy of pure water. As a practically relevant second result, we further show that, for high protein concentrations, the diffusion of small molecules such as dioxane or water is not generally a suitable probe for the viscosity experienced by the diffusing proteins.

Received 16th February 2016,  
Accepted 6th June 2016

DOI: 10.1039/c6cp01056c

www.rsc.org/pccp

## Introduction

The evaluation of experimental and simulation results on translational diffusion usually relies on the Stokes–Einstein (SE) relationship. Its applicability is in principle limited to a relatively large solute, effectively a colloidal particle, experiencing Stokes friction within a continuous effective medium – a condition that is safely fulfilled in dilute protein solutions. *In vivo*, however, organic matter is as highly concentrated as several 100 grams of biomolecules per liter,<sup>1–3</sup> rendering macromolecular crowding an important aspect of molecular processes in the living cell.<sup>4–8</sup> Given such a high concentration of dispersed particles, each protein is surrounded by several others of similar size, leading to phenomena well known from colloid science, such as cage effects<sup>9,10</sup> and restricted motion.<sup>11–13</sup> Such crowding effects dominate *in vivo* protein diffusivity<sup>12</sup> and cause considerably reduced diffusion rates.<sup>5,14</sup> Given the altered protein diffusivity in biological systems, the motion of globular proteins has been studied in the cyto-<sup>15–19</sup> and nucleoplasm<sup>20</sup> as well as in the mitochondrial matrix.<sup>21</sup> Along these lines, in

shrunken cells with reduced cellular volume, going along with an even higher concentration of obstacles, protein diffusivity was shown to be more reduced as for swollen cells.<sup>15</sup> Such findings further stress the relevance of high biomass concentration and excluded-volume effects for protein diffusion inside the cellular environment. Yet, despite the increased interest in the consequences of macromolecular crowding, knowledge on the effect of crowding on protein diffusion beyond the general slow-down of diffusion rates remains still sparse. In addition to the effect of obstacles on molecular mobility, binding of proteins to other proteins or cell compartments<sup>17,20</sup> may additionally retard diffusion.

In the absence of binding events, short-time (in-cage) translational displacements are governed by local hydrodynamic effects that depend on the pure-solvent viscosity,<sup>22</sup> whereas long-time translational diffusion provides sufficient averaging over different local surroundings, combined with a multitude of statistically independent encounters of the caged tracer particle with other host particles. Then, the friction experienced by tracer particles can be expected to scale with the macroscopic (zero-shear) solution viscosity  $\eta(c)$ , such that the long-time diffusion coefficient  $D^L$  may be considered to behave as predicted by the generalized Stokes–Einstein (GSE) relation<sup>23–26</sup>

$$D^L(c) = \frac{k_B T}{6\pi\eta(c)R_H}, \quad (1)$$

Martin-Luther-Universität Halle-Wittenberg, Institut für Physik, 06120 Halle (Saale), Germany. E-mail: matthias.roos@physik.uni-halle.de,

kay.saalwaechter@physik.uni-halle.de; Tel: +49-345-55-28564, +49-345-55-28560

<sup>†</sup> Present address: Deutsches Diabetes Zentrum, 40225 Düsseldorf, Germany.

where  $k_B T$  denotes the thermal energy, and  $R_H$  is the hydrodynamic (Stokes) radius of the tracer particle. Provided that only the viscosity depends on the concentration  $c$  of the dispersed particles but not the particle's (apparent) hydrodynamic size, long-time translational diffusion coefficients may thus be expected to follow the same concentration dependence as the inverse viscosity of the same solution.

The validity of eqn (1) has been demonstrated for concentrated monodisperse colloids<sup>23,25</sup> as well as for colloidal mixtures<sup>26,27</sup> and even soft colloid systems.<sup>28</sup> However, for charge-stabilized colloids deviations have been noticed.<sup>24,29</sup> For proteins under macromolecular crowding the applicability of the GSE relationship is either confirmed<sup>10,30,31</sup> or shown to be invalid,<sup>31,32</sup> depending on the specific proteins and crowding agents used. In a recent study,<sup>13</sup> we have demonstrated that concentrated single-component protein solutions follow the GSE relationship independently of the presence or absence of time-stable oligomers, *i.e.* size polydispersity. Studies of protein behavior in which crowding is realized by open random-coil structures such as Ficoll,<sup>33,34</sup> dextran<sup>35,36</sup> or other (semi)flexible polymers,<sup>37,38</sup> however, have to be distinguished from those studying crowding by globular proteins: at intermediate to high concentrations relevant for crowding effects, random-coil polymers interpenetrate each other, leading to long-range entanglement and mesh-size effects that are of different physical nature as compared to crowding by globular particles.<sup>39,40</sup> Thus, in the presence of random-coil polymers, protein diffusion is often<sup>33,36,37,41,42</sup> but not always<sup>34</sup> found to be faster than predicted by the solution viscosity, because the latter is much more affected by entanglement effects than the local diffusivity.

Here, we investigate a protein model mixture of bovine serum albumin (BSA) and the Src-homology 3 (SH3) domain, and show that concentration-dependent, transient binding can account for an apparent violation of the GSE relationship in crowded protein solutions. The SH3 domain and BSA were initially chosen due to their different sizes, combined with the potential affinity of SH3 to bind to proline and arginine rich regions like PxxP and RxxR motives<sup>43,44</sup> (where x can be any amino acid). Both motives are found on the surface of BSA. These features suggest that our model mixture can be taken to represent a typically polydisperse mixture of distinct proteins subject to protein-specific interactions. Additionally, we provide evidence that the diffusivity of small molecules such as dioxane or water does not represent a reliable means to estimate the effective viscosity of highly concentrated protein solutions, challenging theory-based interpretations of protein (or similar macro-solute) diffusivity.

## Materials and methods

### Materials

The SH3 domain of human amphiphysin II/Bin1<sup>45</sup> and BSA were studied both as single-protein solutions and in mixtures. The SH3 domain has a molecular weight of about 10 kDa only, and was <sup>13</sup>C and <sup>15</sup>N enriched. BSA (molecular weight of 66.4 kDa)

was used in unlabeled form. In mixtures, the mass ratio of SH3 to BSA was 1 : 2, resulting in a molar ratio of SH3 to BSA of about 3 : 1. To assess potential specific binding between the two proteins, an <sup>1</sup>H-<sup>15</sup>N HSQC spectrum of 3 mg ml<sup>-1</sup> SH3 was recorded in presence and absence of BSA, the latter at a concentration of 6 mg ml<sup>-1</sup> and 19.8 mg ml<sup>-1</sup>, corresponding to a molar ratio of SH3 : BSA of about 3 : 1 and 1 : 1, respectively. For this particular sample, 225 mM sodium phosphate buffer at pH 7.4 was used to avoid differences in the pH value between the samples. However, comparing the <sup>1</sup>H-<sup>15</sup>N HSQC spectra of the SH3 domain with and without adding BSA revealed no significant changes in the chemical shifts, indicating that no detectable amount of time-stable SH3-BSA complexes are formed under dilute conditions.

SH3 (Src-Homology 3) Bin1 was adopted from amphiphysin II. For recombinant expression of SH3, the gen of the AmpII/SH3 domain was cloned into a pET14b vector; the plasmid was then transformed into BL21(DE3) cells. Overexpression was induced with 1 mM IPTG at OD<sub>600</sub> 0.7 before growing for three hours at 37 °C. After harvesting, the cell pellet was resuspended in 50 mM sodium phosphate, 300 mM NaCl, 20 mM imidazol and 1 mM DTT at pH 8.0, containing protease inhibitor (Sigma Aldrich, product number P2714). Cells were lysed using lysozyme for 30 min, further treated in a microfluidic fluidizer and clarified by centrifugation. The supernatant was loaded on a nickel Sepharose column, washed with resuspension buffer and eluted with a gradient up to 350 mM imidazol. Fractions containing SH3 were pooled and treated with thrombin to cleave the His<sub>6</sub>-tag, where a second nickel column was used to separate cleaved from non-cleaved SH3. After that the flow through was purified using size exclusion (S75) chromatography, in which SH3 containing fractions were again pooled, dialyzed for removing salt, and then lyophilized. Blue-native PAGE depicted the presence of SH3 monomers only.

Fatty acid-free bovine serum albumin (BSA) was received from Sigma-Aldrich (product number A7030) as lyophilized powder, and was dissolved in D<sub>2</sub>O without further purification. Unpurified, commercial BSA contains about 50% permanent oligomers that likely arise from cystein-related disulfide bonds, see ref. 46 and also ref. 47 for evidence regarding the sample used.

In our previous studies of the Brownian dynamics of concentrated BSA solutions,<sup>13,47</sup> no buffer was added to avoid screening of electrostatic interactions among the proteins. To compare with our previous results, the diffusion measurements presented herein were performed under the same conditions, *i.e.* the proteins were dissolved in pure D<sub>2</sub>O. With increasing concentration of the BSA-SH3 mixture, only a slight shift of the pD was observed (6.0 at lowest concentration to 6.4 at highest concentration). After the measurements of protein diffusion were completed, 1% of dioxane was added to the samples containing the protein mixtures for also measuring the diffusivity of small probe molecules in presence of highly concentrated proteins.

### Isotope-filtered pulsed-field gradient (PFG) NMR

Translational diffusion measurements using the pulsed-gradient stimulated-echo (PGSTE) pulse sequence were performed on an

Avance II spectrometer from Bruker (Karlsruhe, Germany) at 400 MHz  $^1\text{H}$  resonance frequency, equipped with a Diff60 probehead. The PFG NMR diffusion decays were in all cases singly exponential within experimental precision, as confirmed by fits to

$$A(g) = A(0) \cdot \exp(-\gamma^2 g^2 D^L \delta^2 (\Delta - \delta/3)), \quad (2)$$

where  $A(g)$  is the signal intensity in dependence of the applied gradient strength  $g$ ;  $\delta = 1$  ms is the duration of the pulsed gradient, and  $\gamma$  is the  $^1\text{H}$  gyromagnetic ratio.<sup>48</sup> The diffusion time  $\Delta$  was set to 22 ms, providing mean-square displacements (MSDs) in the range of 1  $\mu\text{m}$ . Such MSDs are much larger than the length scale of cage effects by surrounding particles, thus providing diffusion coefficients in the long-time limit. In mixtures of BSA and SH3, the STE for quantification of self-diffusion was combined with an isotope filter, version (C) in Fig. 1 of ref. 49, to selectively detect the signal of either SH3 or BSA. We used  $^{13}\text{C}$ -filtered  $^1\text{H}$  detection rather than  $^{15}\text{N}$  isotope filtering to avoid the influence of chemical-exchange effects of amide protons during the diffusion delay. Depending on the chosen receiver phase in the pulse sequence, either only protons coupled to  $^{13}\text{C}$  or all protons that are not coupled to  $^{13}\text{C}$  nuclei are detected. As SH3 was uniformly  $^{13}\text{C}$  enriched, but BSA was not, detection of  $^{13}\text{C}$ -bound protons provides a signal dominated by the aliphatic protons of SH3. In contrast, if the receiver phase is changed and only the aliphatic region of the NMR signal is evaluated (as was also done when measuring SH3 diffusion), only  $^{12}\text{C}$ -bound protons will be detected, providing an NMR signal in which the SH3 signal is absent, and solely BSA diffusion is observed.

PFG-NMR diffusion experiments are potentially biased by  $T_2$  filtering effects. The STE pulse sequence consists of three parts: encoding of the signal, the diffusion period, and signal decoding. During the encoding and decoding periods of the experiment, transverse ( $T_2$ ) relaxation reduces the NMR-signal according to  $A(0) = \sum_k \hat{A}_k(0) \exp[-\tau/T_{2,k}]$ , where  $T_{2,k}$  and  $\hat{A}_k$  are the transverse relaxation time and the intrinsic intensity (before  $T_2$  relaxation) of the individual protein proton sites, respectively, and  $\tau = 2 \times 3.6$  ms is the overall time span of the encoding and decoding periods of the NMR pulse sequence during which  $T_2$  relaxation occurs.  $T_2$  relaxation times are inversely proportional to the rotational correlation time of the particle. Larger particles with consequently longer rotational correlation times have – averaged over the different proton sites of the protein – shorter transverse relaxation times  $T_2$  than smaller particles. If  $T_2$  is comparable to or even shorter than  $\tau$  (as is the case in concentrated protein solutions) and the solution is polydisperse, the signal obtained can be biased by monomers instead of representing the full ensemble average. This is the case for the BSA fraction, see below and ref. 13. Assuming a log-normal distribution of relaxation times, the median relaxation time represents the most populated one. For BSA and SH3 in mixture, the median relaxation time ranges between only 0.5 ms at high concentration and low temperature, and 5 ms at low concentration and high temperature. For comparison, SH3 at a concentration of 15  $\text{mg ml}^{-1}$  has a

temperature-dependent relaxation time ranging between 7 and 13 ms. For transverse relaxation data of concentrated BSA, see ref. 47.

### Viscosity measurements

Steady-shear viscosities were measured using the capillary microfluid viscometer m-VROC (Rheosense, Sam Ramon, CA) *via* detection of the fluid pressure gradient along the capillary, applying a shear rate of 2000  $\text{s}^{-1}$ . When decreasing the shear rate to 100  $\text{s}^{-1}$ , no change in viscosity was detected, yet the signal-to-noise ratio decreased due to a lower pressure gradient inside the capillary. To avoid a bias by isotope effects, samples for viscosity measurements were also prepared with  $\text{D}_2\text{O}$ .

### Isothermal titration calorimetry

To measure the interaction between SH3 and BSA, isothermal titration calorimetry (ITC) measurements were performed, titrating BSA into an SH3 solution. Measurements were performed on a VP-ITC calorimeter (Microcal LLC, Northampton, MA) at 20  $^\circ\text{C}$ . To separate dilution effects, BSA was also titrated into pure buffer as a reference.

## Results and discussion

### Single-protein solutions

In single-protein solutions, both proteins behave as predicted from the GSE equation, eqn (1); see Fig. 1(a). Using eqn (1), the hydrodynamic radii for SH3 and BSA are calculated to be 2.0 nm

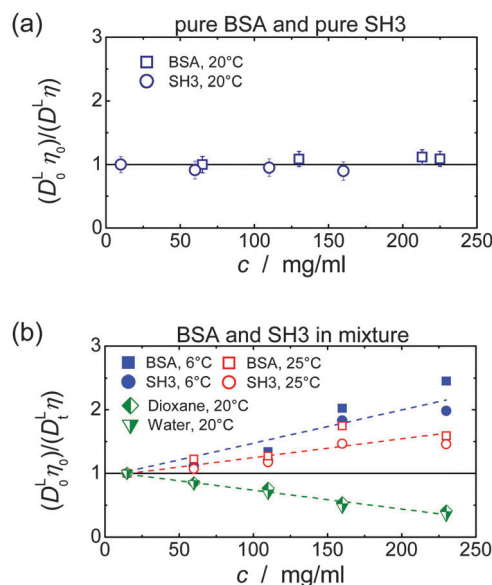


Fig. 1 Test of the applicability of the GSE relation in concentrated solutions of (a) only SH3 or only BSA, and (b) in mixtures of SH3 and BSA, including the behavior of water and dioxane. At a fixed temperature, applicability of the GSE relationship requires  $D^L \eta = \text{const.}$  (solid horizontal line). Note that increased  $(D_0^L \eta_0)/(D^L \eta) = R_{H,\text{app}}/R_{H,0}$  suggest an increase in apparent hydrodynamic size. Data of pure BSA are taken from ref. 13. In part (b) error bars were skipped for clarity; dashed lines are linear fits to indicate the trend. In each case, the lowest concentration measured was used for normalization ( $D_0^L \eta_0$ ).

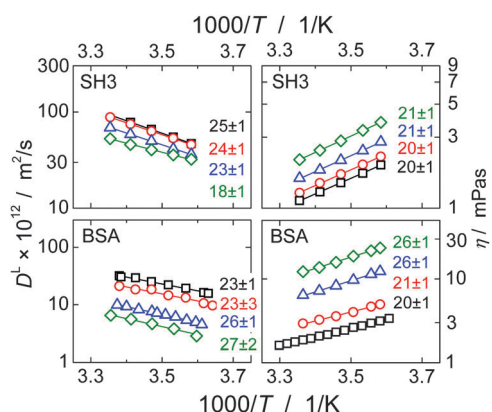


Fig. 2 Temperature dependence of translational long-time diffusion coefficients  $D^L$  (left) and viscosity (right) of the single-protein solutions in an Arrhenius plot. Numbers next to the data points give the activation energy and fitting error in  $\text{kJ mol}^{-1}$  as obtained from fitting the experimental data (solid lines). Protein concentrations are  $10 \text{ mg ml}^{-1}$  (squares),  $60 \text{ mg ml}^{-1}$  (circles),  $110 \text{ mg ml}^{-1}$  (triangles) and  $160 \text{ mg ml}^{-1}$  (rhombs) for SH3, and, for BSA,  $65 \text{ mg ml}^{-1}$  (squares),  $130 \text{ mg ml}^{-1}$  (circles),  $213 \text{ mg ml}^{-1}$  (triangles) and  $255 \text{ mg ml}^{-1}$  (rhombs). Data on translational diffusion of BSA as well as its viscosity at  $213 \text{ mg ml}^{-1}$  were taken from our study on single-protein solutions.<sup>13</sup>

and  $3.4 \text{ nm}$ ,<sup>13</sup> respectively. These values are in accordance with the predictions of  $2.2 \text{ nm}$  for SH3 using the HydroPro software<sup>50</sup> (applied to protein data bank file 1BB9) and the known value of  $3.44 \text{ nm}$  for the BSA monomer.<sup>51</sup> As indicated by the hydrodynamic size obtained, the PFG-NMR diffusion signal of BSA provides the translational displacement of mostly the monomers within the polydisperse (monomers and oligomers) surrounding.<sup>13</sup> A very important observable is the temperature dependence of translational diffusion and viscosity (Fig. 2): for both quantities and for both proteins, the activation energy  $E_A$  was found to be in the range of  $18\text{--}27 \text{ kJ mol}^{-1}$ . These values of around  $20 \text{ kJ mol}^{-1}$  agree well with the activation energy of the viscous flow of pure water,<sup>52</sup> indicating that local hydrodynamic interactions with the solvent dominate viscosity and translational diffusion. Somewhat enhanced activation energies at higher protein concentrations (e.g. BSA) suggest somewhat increased energetic barriers for particles diffusing in a crowded environment: under such conditions, surrounding particles have to get out of the way of the tracer particle to allow for translational particle rearrangements,<sup>53</sup> and further, confined, surface-related water fractions arise. For SH3 diffusion, the decrease of  $E_A$  may result from partial unfolding<sup>54,55</sup> at high concentration and highest temperatures. This effect is expected to increase the ensemble-averaged hydrodynamic radius, and thus reduces diffusivity. The validity of eqn (1) is in accordance with colloid science concepts<sup>23</sup> and agrees with our study<sup>13</sup> on single-protein solutions (lysozyme, BSA,  $\alpha$ B-crystallin).

### SH3 and BSA in mixtures

In mixtures, the diffusion coefficients of SH3 and BSA are different but have similar concentration dependencies (Fig. 3). Now, however, neither BSA nor SH3 follows the prediction of the GSE relation, see Fig. 1(b). Instead, with increasing protein

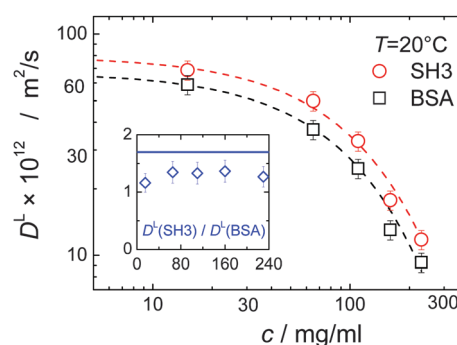


Fig. 3 Concentration dependence of long-time translational diffusion coefficients  $D^L$  of SH3 and BSA in mixtures. Dashed lines are exponential fits to guide the eye. Within the experimental uncertainty, the ratio of the diffusion coefficients (inset) stays constant but below the expected ratio of 1.7 (solid blue line in inset) as estimated from the hydrodynamic size of  $2.0 \text{ nm}$  and  $3.4 \text{ nm}$  for SH3 and BSA, respectively.

concentration an increasing value of  $(D^L\eta)^{-1} \propto R_{H,\text{app}}$  is observed at both  $6 \text{ }^\circ\text{C}$  and  $25 \text{ }^\circ\text{C}$ , meaning that translational diffusion is more retarded than expected from the solution viscosity  $\eta(c)$ . Thus, the apparent, concentration-dependent hydrodynamic radius  $R_{H,\text{app}}$  is about twice as large at high protein concentrations as in dilute solution. As follows from the measurements on single-protein solutions, neither the difference in size or shape between the two proteins, nor the presence of permanent oligomers (BSA) can account for this violation of the GSE relationship; see also ref. 13.

Remember that in the long-time limit of diffusion, the tracer particle is subject to a multitude of statistically independent encounters with surrounding host particles, meaning that translational diffusion is averaged over different micro-environments. This renders long-time translational diffusion of colloidal particles sensitive to the macroscopic viscosity.<sup>26</sup> The deviating behavior of the protein mixture from the behavior of both ordinary colloidal particles – that typically show GSE behavior<sup>23,25–28</sup> – and single-protein solutions suggests that inter-protein interactions and/or concentration-dependent binding are of critical importance for the mixture of BSA and SH3, accounting for apparent hydrodynamic radii that seemingly increase with increasing protein concentration.

Indeed, for the mixtures of SH3 and BSA, ITC reveals an increasing protein binding tendency with increasing protein concentration, see Fig. 4: adding BSA to SH3 results in an exothermal process that becomes more pronounced with increasing concentration of BSA until saturation is reached. This increasing reactivity with increasing concentration indicates concentration-induced binding. Dilution of BSA in absence of SH3 does not reveal such an effect; besides the free energy of dilution, no indication of an exothermal event is observed.

Considering that, at least in dilute solutions amenable to high-resolution NMR, we were not able to detect direct indications for site-specific binding, we should also consider more general, crowding-related reasons for protein self- and hetero-association.<sup>56,57</sup> Bound states generally require less space than individual, monomeric proteins. Thus, entropic excluded-volume



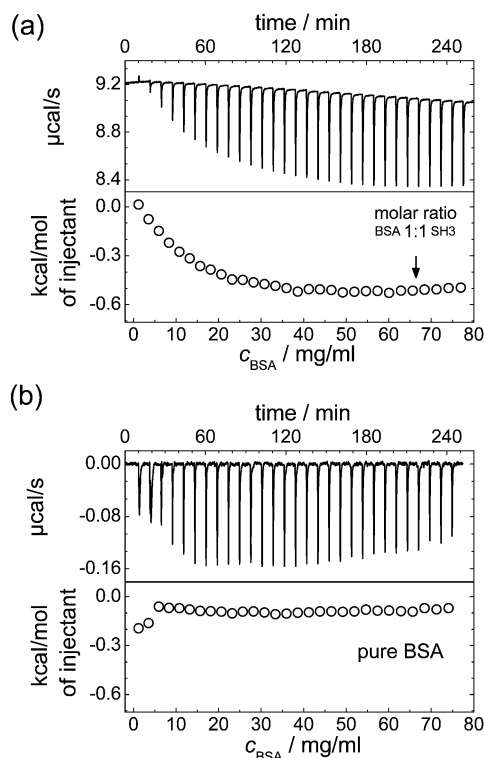


Fig. 4 ITC results for (a) addition of BSA to a 10 mg ml<sup>-1</sup> SH3 solution, and (b) dilution of BSA in PBS buffer. The dilution of BSA in absence of SH3 also acts as a reference measurement reflecting the low contribution of the free energy of dilution to the measurement result in (a).

effects can also account for increased inter-protein binding affinity. The difference in size between SH3 and BSA combined with the high protein concentration renders this depletion interaction particularly relevant. The depletion effect promotes a clustering of the larger particles, such that the presence of SH3 may cause BSA molecules to form transient clusters. However, the diffusivity of the smaller SH3 is more affected by the transient binding, as seen from the ratio of the diffusion coefficients  $D^L(\text{SH3}) : D^L(\text{BSA})$ , which is smaller than expected from the hydrodynamic radii (inset in Fig. 3). This is contrary to the expectation based upon a depletion attraction that should only affect the larger BSA.

Theory<sup>58</sup> predicts that dimerization provides an energy gain in the order of  $k_B T$  in case that 10–30% of the totally available volume is occupied by other particles. Indeed, an experimental study<sup>59</sup> showed that crowding stabilizes protein–protein binding by about 1 kcal mol<sup>-1</sup>, *i.e.* by  $1-2k_B T$ . Given the non-specific nature of crowding-induced complex formation combined with its low binding energy, such complex formation can only be considered to be transient rather than permanent.

Irrespective of the potential (non-)specificity of the BSA–SH3 interaction, it is clear that, if the lifetime of transient binding is faster than or comparable to the experimental time scale of diffusion (of the order of 20 ms), the trajectory of the tracer particle will average over bound and unbound states.<sup>60</sup> Thus, it will provide an effective, but enlarged hydrodynamic size that increases monotonically with the overall protein concentration.

Our observations are consistent with this picture in that mono-exponential diffusion decays according to eqn (2) were observed in all cases, *i.e.*, only a single diffusion coefficient is detected. We stress that  $T_2$ -filtering effects cannot explain our observations: as long as bound and unbound populations are exchanging on the 20 ms timescale of the diffusion delay  $\Delta$ , a potential  $T_2$  bias would not affect the measured result. Note that  $T_2$  of an <sup>1</sup>H protein signal integrated over a broad chemical shift range is generally non-exponential due to differences in spin–spin dipolar couplings and differences in the local internal mobility of side chains, such that a size polydispersity (that may be stable on the shorter 1 ms timescale of the transverse encoding delay  $\delta$ ) cannot simply be judged on the basis of simple  $T_2$  decays.

The concept of crowding-induced transient binding is most directly corroborated by the temperature dependence of translational diffusion, see Fig. 5: for both BSA and SH3, the apparent activation energy  $E_A$  of translational diffusion is significantly increased by a factor of about 1.5 to 2 as compared to the single-protein solutions. The apparent  $E_A$  thus reflects the thermodynamics of the binding equilibrium. As we lack information on the specificity and stoichiometry of the binding and the structures of the transient protein assemblies, it is not possible to provide a more detailed analysis of the measured values. However, the effect of the binding equilibrium, with bound states being favored at lower temperature, is directly supported by the fact that the deviations from the GSE prediction are more pronounced at lower temperatures; see again Fig. 1(b).

For viscosity, in turn, the apparent activation energy matches within the experimental precision the value found for only BSA and only SH3. In contrast to translational diffusion, viscosity does

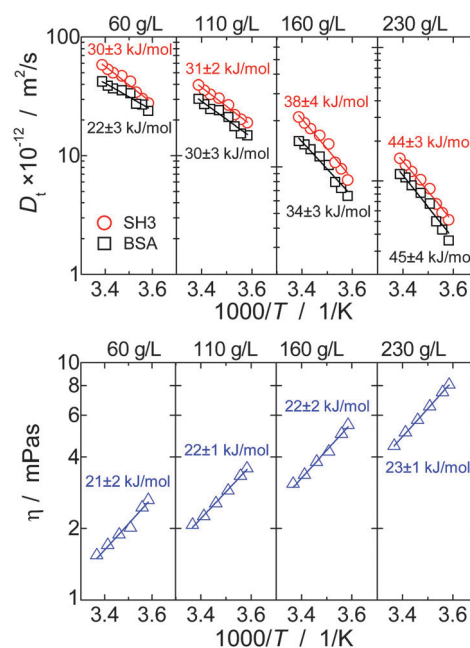


Fig. 5 Temperature dependence of long-time translational diffusion coefficients  $D^L$  and viscosity  $\eta$  of the mixture of BSA and SH3 at different concentrations, plotted in Arrhenius coordinates. Solid lines are fits using the Arrhenius law.

not depend on hydrodynamic sizes of the individual diffusing particles, but is determined by the overall volume fraction of all dispersed particles. The extent to which (unspecific) binding reduces the volume occupied by the proteins thus appears to be reasonably small, and, as a result, the effect of oligomerization on the apparent  $E_A$  is much weaker for viscosity than for translational diffusion.

We finally stress that our general observation is in qualitative agreement with previous work. In studies in which tracer protein diffusion was studied in the presence of another protein species, long-time translational diffusion was either found to be slower than expected from viscosity,<sup>31,32</sup> or the validity of the GSE equation was retained.<sup>30</sup> In contrast, for the case of crowding by globular proteins, we are not aware of any studies reporting on protein translational diffusion being faster than expected from the macroviscosity.

### Diffusion of small probe molecules

Fig. 1(b) clearly demonstrates that the concentration dependence of dioxane and water diffusion is qualitatively different from that of the proteins. Thus, it is immediately clear that small molecules are not reliable reporters on the viscosity experienced by highly concentrated macromolecules. Instead, the small solvent molecules diffuse faster than predicted from the SE relation, at least with regards to the (macroscopic) solution viscosity. Under crowding conditions, these small probe molecules diffuse in an environment of much larger surrounding particles, which renders the validity of the effective-medium approach questionable with regards to estimating the macroscopic dispersion viscosity. On the time-scale of translational motion of the small probe molecules, proteins appear as quasi-static objects that obstruct the diffusion of the much smaller solvent molecules, which is otherwise governed by the pure-solvent viscosity. Such a scenario is known to go along with a linear decrease of the translational diffusion coefficient with the volume fraction  $\phi$  of obstructing particles,<sup>61</sup>

$$D^L = D_0^L (1 - \beta\phi), \quad (3)$$

where  $\beta$  is a shape factor, and  $\phi = c\theta$  is connected to the protein mass concentration  $c$  via the protein's specific volume  $\theta$ . In our case, the diffusion of solvent molecules is hindered by both SH3 and BSA, such that  $D^L = D_0^L(1 - \beta_{\text{SH3}}\phi_{\text{SH3}} - \beta_{\text{BSA}}\phi_{\text{BSA}}) = D_0^L(1 - \bar{\beta}\bar{\theta}c)$ . Here,  $\bar{\beta} = \beta_{\text{SH3}}\phi_{\text{SH3}}/(\phi_{\text{SH3}} + \phi_{\text{BSA}}) + \beta_{\text{BSA}}\phi_{\text{BSA}}/(\phi_{\text{SH3}} + \phi_{\text{BSA}})$  and  $\bar{\theta} = \theta_{\text{SH3}}c_{\text{SH3}}/c + \theta_{\text{BSA}}c_{\text{BSA}}/c$  are independent of the overall protein concentration  $c = c_{\text{SH3}} + c_{\text{BSA}}$  since the ratio of SH3 to BSA molecules is the same for all concentrations. Fig. 6 confirms the expected linear obstruction effect with protein concentration for water and dioxane diffusion.

The linear decrease of the diffusivity of small molecule stands in stark contrast to the (approximately) exponential slow-down of protein diffusion with increasing protein concentration; cf. Fig. 3 and ref. 53, 64 and 65. Likewise, the (macroscopic) solution viscosity commonly increases more than linearly with protein concentration.<sup>66,67</sup>

For simple liquids, in which solute and solvent are of equal or similar size, the validity of the SE relationship can often be

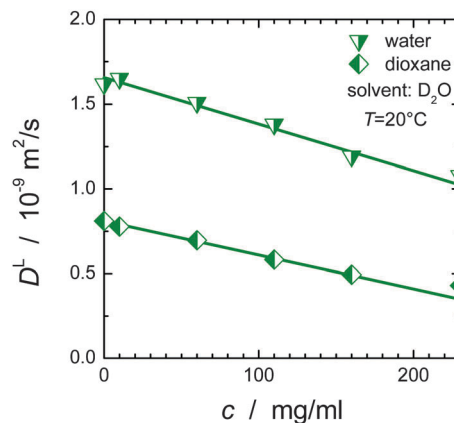


Fig. 6 Obstruction effect for water and dioxane as a function of the overall protein concentration  $c$ . The diffusion coefficient of dioxane at zero protein concentration was calculated via the Stokes–Einstein relation using  $R_H = 2.12 \text{ \AA}^{62}$  and  $\eta_{\text{D}_2\text{O}} = 1.25 \text{ mPa s}$ ; for water the diffusion coefficient of  $2.02 \times 10^{-9} \text{ m}^2 \text{ s}^{-1}$  at  $20 \text{ }^\circ\text{C}^{63}$  was corrected by a factor of 1.25 to account for the different viscosity of  $\text{D}_2\text{O}$  and  $\text{H}_2\text{O}$ .

retained via introducing slip boundary conditions,<sup>68,69</sup> i.e., the viscous friction  $\zeta$  in  $D^L = k_B T / \zeta$  scales as  $\zeta = 4\pi\eta R_H$  rather than the usual case of  $\zeta = 6\pi\eta R_H$  (stick boundary conditions). However, the maximal factor of 2/3, even when taken to vary between this value and 1 as a function of concentration, cannot account for such a difference in the concentration dependence of the diffusivity of solvent molecules and the macroscopic viscosity (and protein diffusion).

For small molecules, the deviation from SE behavior is reminiscent of neutral solutes in ionic liquids, where the solute-to-solvent size ratio plays a key role.<sup>70–73</sup> Despite noting that charge–charge interactions are much more relevant for ionic liquids as for the diffusivity of water in presence of proteins, the deviations from SE behavior are similar in origin. In ionic liquids, regions of locally increased charge concentration exist next to regions of lower charge density, accounting for dynamical heterogeneity.<sup>71,72</sup> Depending on the size of the tracer particle relative to the length scale of such heterogeneity, the effect is either averaged out, recovering SE behavior,<sup>72</sup> or the neutral solute is small enough to predominately diffuse within regions of high mobility, accounting for diffusion rates higher than expected from the macroscopic viscosity. Although dynamical heterogeneity is not of relevance for long-range diffusion of water or dioxane in protein solutions, the structurally inhomogeneous surroundings (large, almost immobile proteins compared to the small solvent molecules) provide a similar scenario as for ionic liquids: despite of the obstruction effect mediated by the proteins, the environment probed by small solvent particles is dominated by the solvent itself, and contrasts with the (sterical) hindrance relevant for the proteins. Whereas a small molecule escapes a “macromolecular cage” almost effortlessly, proteins are much more sensitive to macromolecular caging, introducing dynamic correlations on larger scales, and thus experience a stronger retardation in diffusion than does water or dioxane.

This latter effect even holds with regards to the hydration shell of the proteins: the life-time of water molecules within the hydration shell is usually on the timescale of picoseconds,<sup>74</sup> such that trapping of a water molecule within the hydration shell does not significantly hinder water diffusion. Thus, small probe molecules experience a local, solvent-related viscosity with some obstruction by the presence of the proteins, rather than being sensitive to the (macroscopic) dispersion viscosity. As a result, the GSE relationship cannot be applied (see Fig. 1(b)). Notably, this effect can already be seen at low concentrations, although the deviation is reasonably weak. On the other extreme, in case of a hydrated lyophilized protein powder, confinement effects become relevant even for small molecules, and surface effects must be taken into account.<sup>75</sup>

Protein diffusion within an entangled polymer matrix is phenomenologically similar to diffusion of water or dioxane under macromolecular crowding. Indeed, protein diffusion is usually faster than predicted by the GSE relationship in the presence of Ficoll or other unstructured polymers.<sup>33,36,37</sup> For globular proteins crowded by other globular proteins, size effects are usually not that dramatic, such that protein diffusion is well described by the macroscopic (zero-shear) viscosity,<sup>13</sup> given that no transient binding occurs.

## Conclusions

In the absence of transient binding, long-time self-diffusion scales with the solution viscosity. Such a behavior is expected on the basis of established concepts in colloid science, and was demonstrated here for single-protein solutions of the SH3 domain and BSA, the latter even having a distribution in size and shape of the Brownian particles. In both cases, the activation energies of translational diffusion and macro-viscosity coincide within experimental uncertainty, being close to the activation energy for viscous flow of pure water. In contrast, upon mixing SH3 and BSA, translational diffusion of either species was found to be single-component but more retarded than expected from the solution viscosity: the hydrodynamic radius thus appeared to increase with concentration, indicating the relevance of attractive inter-protein interactions and concentration-dependent, transient binding.

While specific binding sites could not be resolved in <sup>1</sup>H-<sup>15</sup>N HSQC NMR spectra of a necessarily rather dilute solution of isotope-labeled SH3 in the presence of BSA, calorimetric measurements did prove an exothermal reaction between SH3 and BSA. The relevance of transient binding is most clearly corroborated by the observed temperature dependence of translational diffusion. In a mixture of both proteins, the apparent activation energy increased significantly with protein concentration, with values being notably larger than the again nearly constant activation energy of the macroscopic viscosity of the same mixture. This is in line with the fact that solution viscosity mainly depends on the overall volume fraction of the solutes rather than on individual particle sizes (Einstein model). The relevance of transient binding is further supported by the fact

that the apparent increase in hydrodynamic radius with concentration was found to be larger at low temperatures, at which the binding equilibrium is shifted towards the bound state.

Moreover, we showed that translational diffusion of small probe molecules (water, dioxane) does not allow for conclusions on the viscosity relevant for long-time translational diffusion of highly concentrated proteins: the local viscosity probed by dioxane or water is in-between the pure solvent viscosity and the macroscopic solution viscosity, and is mainly affected by geometric obstruction effects.

## Acknowledgements

We are grateful to Alexey Krushelnitsky for very helpful discussions, and thank Sebastian Kirschke for preparatory work during his Bachelor thesis, implementing and testing the pulse sequence for isotope-filtered PFG NMR. Funding of this work was provided by the Deutsche Forschungsgemeinschaft (DFG) in the framework of the collaborative research center SFB/TRR 102 (project A08). We also acknowledge significant investments in our NMR facility by the European Regional Development Fund (ERDF) of the European Union.

## References

- 1 A. B. Fulton, How crowded is the cytoplasm?, *Cell*, 1982, **30**, 345–347.
- 2 S. B. Zimmerman and S. O. Trach, Estimation of macromolecule concentrations and excluded volume effects for the cytoplasm of *Escherichia coli*, *J. Mol. Biol.*, 1991, **222**, 599–620.
- 3 R. J. Ellis and A. P. Minton, Cell biology: Join the crowd, *Nature*, 2003, **425**, 27–28.
- 4 R. J. Ellis, Macromolecular crowding: obvious but underappreciated, *Trends Biochem. Sci.*, 2001, **26**(10), 597–604.
- 5 R. J. Ellis, Macromolecular crowding: an important but neglected aspect of the intracellular environment, *Curr. Opin. Struct. Biol.*, 2001, **11**, 114–119.
- 6 O. Medalia, I. Weber, A. S. Frangakis, D. Nicastro, G. Gerisch and W. Baumeister, Macromolecular architecture in eukaryotic cells visualized by cryoelectron tomography, *Science*, 2002, **298**, 1209–1213.
- 7 H. X. Zhou, G. N. Rivas and A. P. Minton, Macromolecular crowding and confinement: biochemical, biophysical, and potential physiological consequences, *Annu. Rev. Biophys.*, 2008, **37**, 375–397.
- 8 S. Klumpp, M. Scott, S. Pedersen and T. Hwa, Molecular crowding limits translation and cell growth, *Proc. Natl. Acad. Sci. U. S. A.*, 2013, **110**(42), 16754–16759.
- 9 B. Doliwa and A. Heuer, Cage effect, local anisotropies, and dynamic heterogeneities at the glass transition: A computer study of hard spheres, *Phys. Rev. Lett.*, 1998, **80**(22), 4915–4918.
- 10 M. Roos, S. Link, J. Balbach, A. Krushelnitsky and K. Saalwächter, NMR-detected Brownian dynamics of

- $\alpha$ B-crystallin over a wide range of concentrations, *Biophys. J.*, 2015, **108**, 98–106.
- 11 T. J. O'Leary, Concentration dependence of protein diffusion, *Biophys. J.*, 1987, **52**, 137–139.
  - 12 T. Ando and J. Skolnick, Crowding and hydrodynamic interactions likely dominate *in vivo* macromolecular motion, *Proc. Natl. Acad. Sci. U. S. A.*, 2010, **107**(43), 18457–18462.
  - 13 M. Roos, M. Ott, M. Hofman, S. Link, E. Rössler, J. Balbach, A. Krushelnitsky and K. Saalwächter, under revision.
  - 14 K. Luby-Phelps, Cytoarchitecture and physical properties of cytoplasm: volume, viscosity, diffusion, intracellular surface area, *Int. Rev. Cytol.*, 2000, **192**, 189–221.
  - 15 J. W. Wojcieszyn, R. A. Schlegel, E.-S. Wu and K. A. Jacobson, Diffusion of injected macromolecules within the cytoplasm of living cells, *Proc. Natl. Acad. Sci. U. S. A.*, 1981, **78**(7), 4407–4410.
  - 16 R. Swaminathan, C. P. Hoang and A. S. Verkman, Photo-bleaching recovery and anisotropy decay of green fluorescent protein GFP-S65T in solution and cells: cytoplasmic viscosity probed by green fluorescent protein translational and rotational diffusion, *Biophys. J.*, 1997, **72**(4), 1900–1907.
  - 17 J. Lippincott-Schwartz, E. Snapp and A. Kenworthy, Studying Protein Dynamics in Living Cells, *Nat. Rev. Mol. Cell Biol.*, 2001, **2**, 444–456.
  - 18 A. S. Verkman, Solute and macromolecule diffusion in cellular aqueous compartments, *Trends Biochem. Sci.*, 2002, **27**, 27–33.
  - 19 M. Kumar, M. S. Mommer and V. Sourjik, Mobility of Cytoplasmic, Membrane, and DNA-Binding Proteins in *Escherichia coli*, *Biophys. J.*, 2010, **98**, 552–559.
  - 20 R. D. Phair and T. Misteli, High mobility of proteins in the mammalian cell nucleus, *Nature*, 2000, **404**, 604–609.
  - 21 A. Partikian, B. Ölveczky, R. Swaminathan, Y. Li and A. S. Verkman, Rapid Diffusion of Green Fluorescent Protein in the Mitochondrial Matrix, *J. Cell Biol.*, 1998, **140**, 821–829.
  - 22 F. Roosen-Runge, M. Hennig, F. Zhang, R. M. J. Jacobs, M. Sztucki, H. Schober, T. Seydel and F. Schreiber, Protein self-diffusion in crowded solutions, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, **108**(29), 11815–11820.
  - 23 P. N. Segre, S. P. Meeker, P. N. Pusey and W. C. K. Poon, Viscosity and structural relaxation in suspensions of hard-sphere colloids, *Phys. Rev. Lett.*, 1995, **75**(5), 958–961.
  - 24 A. J. Banchio, G. Nägele and J. Bergenholtz, Viscoelasticity and generalized Stokes–Einstein relations of colloidal dispersions, *J. Chem. Phys.*, 1999, **111**(18), 8721–8740.
  - 25 A. J. Banchio, J. Bergenholtz and G. Nägele, Rheology and dynamics of colloidal suspensions, *Phys. Rev. Lett.*, 1999, **82**(8), 1792–1795.
  - 26 G. H. Koenderink and A. P. Philipse, Rotational and translational self-diffusion in colloidal sphere suspensions and the applicability of generalized Stokes–Einstein relations, *Langmuir*, 2000, **16**, 5631–5638.
  - 27 W. Richtering and H. Müller, Comparison between Viscosity and Diffusion in Monodisperse and Bimodal Colloidal Suspensions, *Langmuir*, 1995, **11**, 3699–3704.
  - 28 S. Gupta, J. Stellbring, E. Zaccarelli, C. N. Likos, M. Camargo, P. Holmqvist, J. Allgaier, L. Willner and D. Richter, Validity of the Stokes–Einstein relation in soft colloids up to the glass transition, *Phys. Rev. Lett.*, 2015, **115**, 128302.
  - 29 A. Imhof, A. van Blaaderen, G. Maret, J. Mellema and J. K. G. Dhont, A comparison between the long-time self-diffusion and low shear viscosity of concentrated dispersions of charged colloidal silica spheres, *J. Chem. Phys.*, 1994, **100**, 2170–2181.
  - 30 P. Licinio and M. Delaye, Mutual and self-diffusion in concentrated  $\alpha$ -crystallin protein dispersion. A dynamic light-scattering study, *J. Phys.*, 1988, **49**, 975–981.
  - 31 Y. Wang, C. Li and G. J. Pielak, Effects of proteins on protein diffusion, *J. Am. Chem. Soc.*, 2010, **132**(27), 9392–9397.
  - 32 S. Zorilla, M. A. Hink, A. J. Visser and M. P. Lillo, Translational and rotational motions of proteins in a protein crowded environment, *Biophys. Chem.*, 2007, **125**, 298–305.
  - 33 N. A. Busch, T. Kim and V. A. Bloomfield, Tracer diffusion of proteins in DNA Solutions. 2. Green fluorescent protein in crowded DNA solutions, *Macromolecules*, 2000, **33**, 5932–5937.
  - 34 E. Dauty and A. S. Verkman, Molecular crowding reduces to a similar extent the diffusion of small solutes and macromolecules: measurement by fluorescence correlation spectroscopy, *J. Mol. Recognit.*, 2004, **17**, 441–447.
  - 35 D. Lavalette, M. A. Hink, M. Tourbez, C. Tétreau and A. J. Visser, Proteins as micro viscosimeters: Brownian motion revisited, *Eur. Biophys. J.*, 2006, **35**, 517–522.
  - 36 S. P. Zustiak, R. Nossal and D. L. Sackett, Hindered diffusion in polymeric solutions studied by fluorescence correlation spectroscopy, *Biophys. J.*, 2011, **101**, 255–264.
  - 37 Y. Y. Kuttner, N. Kozer, E. Segal, G. Schreiber and G. Haran, Separating the Contribution of Translational and Rotational Diffusion to Protein Association, *J. Am. Chem. Soc.*, 2005, **127**(27), 15138–15144.
  - 38 C. Li, Y. Wang and G. J. Pielak, Translational and rotational diffusion of a small globular protein under crowded conditions, *J. Phys. Chem. B*, 2009, **113**, 13390–13392.
  - 39 A. Soranno, I. Koenig, M. B. Borgia, H. Hofmann, F. Zosel, D. Nettels and B. Schuler, Single-molecule spectroscopy reveals polymer effects of disordered proteins in crowded environments, *Proc. Natl. Acad. Sci. U. S. A.*, 2014, **111**(13), 4874–4879.
  - 40 J. Danielsson, X. Mu, L. Lang, H. Wang, A. Binolfi, F.-X. Theillet, B. Bekei, D. T. Logan, Ph. Selenko, H. Wennerström and M. Oliveberg, Thermodynamics of protein destabilization in live cells, *Proc. Natl. Acad. Sci. U. S. A.*, 2015, **112**(40), 12402–12407.
  - 41 H. Tanaka, I. Yoshizaki, S. Takahashi, M. Yamanaka, S. Fukuyama, M. Sato, S. Sano, M. Motohara, T. Kobayashi, S. Yoshitomi and T. Tanaka, Diffusion coefficient of the protein in various crystallization solutions: the key to growing high-quality crystals in space, *Microgravity Sci. Technol.*, 2006, **18**(3–4), 91–94.
  - 42 A. B. Goins, H. Sanabria and M. N. Waxham, Macromolecular Crowding and Size Effects on Probe Microviscosity, *Biophys. J.*, 2008, **95**, 5362–5373.
  - 43 K. Elliott, D. Sakamuro, A. Basu, W. Du, W. Wunner, P. Staller, S. Gaubatz, H. Zhang, E. Prochownik, M. Eilers and G. C. Prendergast, Bin1 functionally interacts with Myc

- and inhibits cell proliferation *via* multiple mechanisms, *Oncogene*, 1999, **18**, 3564–3573.
- 44 V. Anggono and P. J. Robinson, Syndapin I and endophilin I bind overlapping proline-rich regions of dynamin I: role in synaptic vesicle endocytosis, *J. Neurochem.*, 2007, **102**, 931–943.
- 45 S. Suetsu, K. Toyooka and Y. Senju, Subcellular membrane curvature mediated by the BAR domain superfamily proteins, *Semin. Cell Dev. Biol.*, 2010, **21**, 340–349.
- 46 P. G. Squire, P. Moser and C. T. O’Konski, The hydrodynamic properties of bovine serum albumin monomer and dimer, *Biochemistry*, 1968, **7**, 4261–4272.
- 47 M. Roos, M. Hofmann, S. Link, M. Ott, J. Balbach, E. Rössler, K. Saalwächter and A. Krushelnitsky, The “long tail” of the protein tumbling correlation function: observation by <sup>1</sup>H NMR relaxometry in a wide frequency and concentration range, *J. Biomol. NMR*, 2015, **63**, 403–415.
- 48 E. O. Stejskal and J. E. Tanner, Spin diffusion measurements: spin echoes in the presence of a time-dependent field gradient, *J. Chem. Phys.*, 1965, **42**, 288–292.
- 49 M. L. Tillet, M. A. Horsfield, L.-Y. Lian and T. J. Norwood, Protein-ligand interactions measured by <sup>15</sup>N-filtered diffusion experiments, *J. Biomol. NMR*, 1999, **13**, 223–232.
- 50 J. G. de la Torre, M. L. Huertas and B. Carrasco, Calculation of hydrodynamic properties of globular proteins from their atomic-level structure, *Biophys. J.*, 2000, **78**, 719–730.
- 51 I. Axelsson, Characterization of proteins and other macromolecules by agarose gel chromatography, *J. Chromatogr. A*, 1978, **152**(1), 21–32.
- 52 R. A. Horne, R. A. Courant, D. S. Johnson and F. F. Margosian, The activation energy of viscous flow of pure water and sea water in the temperature region of maximum density, *J. Phys. Chem.*, 1965, **69**, 3988–3991.
- 53 N. Muramatsu and A. P. Minton, Tracer diffusion of globular proteins in concentrated protein solutions, *Proc. Natl. Acad. Sci. U. S. A.*, 1988, **85**, 2984–2988.
- 54 O. Zhang and J. D. Forman-Kay, Structural characterization of folded and unfolded states of an SH3 domain in equilibrium in aqueous buffer, *Biochemistry*, 1995, **34**, 6784–6794.
- 55 A. E. Smith, L. Z. Zhou, A. H. Gorenske, M. Senske and G. J. Pielak, In-cell thermodynamics and a new role for protein surfaces, *Proc. Natl. Acad. Sci. U. S. A.*, 2016, **113**(7), 1725–1730.
- 56 S. B. Zimmerman and A. P. Minton, Macromolecular crowding: biochemical, biophysical, and physiological consequences, *Annu. Rev. Biophys. Biomol. Struct.*, 1993, **22**, 27–65.
- 57 A. P. Minton, Implications of macromolecular crowding for protein assembly, *Curr. Opin. Struct. Biol.*, 2000, **10**, 34–39.
- 58 O. G. Berg, The influence of macromolecular crowding on thermodynamic activity: solubility and dimerization constants for spherical and dumbbell-shaped molecules in a hard-sphere mixture, *Biopolymers*, 1990, **30**, 1027–1037.
- 59 J. Batra, K. Xu, S. Qin and H.-X. Zhou, Effect of macromolecular crowding on protein binding stability: modest stabilization and significant biological consequences, *Biophys. J.*, 2009, **97**, 906–911.
- 60 T. Geyer, Mixing normal and anomalous diffusion, *J. Chem. Phys.*, 2012, **137**, 115101.
- 61 R. Kimmich, F. Klammler, V. D. Skirda, I. A. Serebrennikova, A. I. Maklakov and N. Fatkullin, Geometrical restrictions of water diffusion in aqueous protein systems. A study using NMR field-gradient techniques, *Appl. Magn. Reson.*, 1993, **4**, 425–440.
- 62 D. K. Wilkins, S. B. Grimshaw, V. Receveur, C. M. Dobson, J. A. Jones and L. J. Smith, Hydrodynamic Radii of Native and Denatured Proteins Measured by Pulse Field Gradient NMR Techniques, *Biochemistry*, 1999, **38**, 16424–16431.
- 63 M. Holz, S. R. Heil and A. Sacco, Temperature-dependent self-diffusion coefficients of water and six selected molecular liquids for calibration in accurate <sup>1</sup>H NMR PFG Measurements, *Phys. Chem. Chem. Phys.*, 2000, **2**, 4740–4742.
- 64 T. J. O’Leary, Concentration dependence of protein diffusion, *Biophys. J.*, 1987, **52**, 137–139.
- 65 J. Han and J. Herzfeld, Macromolecular diffusion in crowded solutions, *Biophys. J.*, 1993, **65**, 1155–1161.
- 66 P. Pradipasena and C. Rha, Effect of concentration on apparent viscosity of a globular protein solution, *Polym. Eng. Sci.*, 1977, **17**(12), 861–864.
- 67 W. J. Galush, L. N. Le and J. M. R. Moore, Viscosity Behavior of High-Concentration Protein Mixtures, *J. Pharm. Sci.*, 2012, **110**(3), 1012–1020.
- 68 B. J. Alder, D. M. Gass and T. E. Wainwright, Studies in Molecular Dynamics. VIII. The Transport Coefficients for a Hard-Sphere Fluid, *J. Chem. Phys.*, 1970, **53**, 3813–3826.
- 69 F. Ould-Kaddour and D. Levesque, Molecular-dynamics investigation of tracer diffusion in a simple liquid: Test of the Stokes-Einstein law, *Phys. Rev. E*, 2000, **63**, 011205.
- 70 A. Kaintz, G. Baker, A. Benesi and M. Maroncelli, Solute Diffusion in Ionic Liquids, NMR Measurements and Comparisons to Conventional Solvents, *J. Phys. Chem. B*, 2013, **117**, 11697–11708.
- 71 J. C. Araque, S. K. Yadav, M. Shadeck, M. Maroncelli and C. J. Margulis, How Is Diffusion of Neutral and Charged Tracers Related to the Structure and Dynamics of a Room-Temperature Ionic Liquid? Large Deviations from Stokes-Einstein Behavior Explained, *J. Phys. Chem. B*, 2015, **119**, 7015–7029.
- 72 J. C. Araque, J. J. Hettige and C. J. Margulis, Ionic liquids-Conventional solvent mixtures, structurally different but dynamically similar, *J. Chem. Phys.*, 2015, **143**, 134505.
- 73 Y. Kimura, Y. Kida, Y. Matsushita, Y. Yasaka, M. Ueno and K. Takahashi, Universality of Viscosity Dependence of Translational Diffusion Coefficients of Carbon Monoxide, Diphenylacetylene, and Diphenylcyclopropenone in Ionic Liquids under Various Conditions, *J. Phys. Chem. B*, 2015, **119**, 8096–8103.
- 74 V. P. Denisov and B. Halle, Protein hydration dynamics in aqueous solution, *Faraday Discuss.*, 1996, **103**, 227–244.
- 75 M. Rosenstihl, K. Kämpf, F. Klameth, M. Sattig and M. Vogel, Dynamics of interfacial water, *J. Non-Cryst. Solids*, 2015, **407**, 449–458.



# 5 The key factors that determine the Brownian motion of crowded proteins

Knowledge on the consequences of macromolecular crowding is of particular relevance for elucidating protein behavior in the cellular milieu – a reason for which current research on proteins increasingly pays attention to crowding effects [Ellis, 2001a,b; Ellis and Minton, 2003; Gnutt and Ebbinghaus, 2016]. Current literature results on crowded protein Brownian motion are at points contradictory to each other [Zorilla et al., 2007; Li et al., 2009; Wang et al., 2010] with the key factors accounting for such sample specificity not being identified. Based upon NMR spin relaxation in *non-dilute* protein solutions, the protein rotational auto-correlation function (RACF) has been concluded to deviate from singly exponential behavior [Krushelnitsky, 2006], which may provide a rationale for some of the observations. Nonetheless, data that *directly* resolve the non-exponential nature of non-dilute protein rotational diffusion were missing till now.

To elucidate the effects of macromolecular crowding on protein Brownian motion, NMR spectroscopic methods were combined with viscosity measurements. Long-time translational self-diffusion was measured by pulsed-field gradient (PFG) NMR; insights into protein rotational diffusion rely on NMR spin relaxation measurements. For the latter case, field-cycling protein proton  $R_1$  data recorded in cooperation with the research group of ERNST RÖSSLER (University of Bayreuth) were combined with  $R_{1\rho}$  and  $R_2$  measurements on the same samples. By these techniques, NMR  $^1\text{H}$  spin relaxation was recorded at low magnetic fields, which significantly increases the sensitivity to global rotational diffusion rates. This approach [Roos et al., 2015a,b] overcame obstacles encountered in previous NMR relaxometry studies. The accuracy of this procedure has been confirmed by independent data from polarized fluorescence correlation spectroscopy (FCS) [Roos et al., 2016], the latter experiments were performed and evaluated by MARIA OTT. In contrast, the standard protocol for evaluating NMR relaxation data – the  $R_1/R_2$  approach – provides incorrect results under crowding conditions [Roos et al., 2015a]. Even if the deviation of the RACF from single-exponential behavior is weak, biased results will be obtained. The application of the standard  $R_1/R_2$  protocol should thus be restricted to highly dilute protein solutions

## 5. Summary, conclusions and outlook

only. Then, inter-protein interactions are safely negligible and the RACF conforms to a single exponential,  $C_r(t) \cong \tilde{C}_{\text{rot}} \propto \exp(-|t|/\tau_{\text{rot}})$ .

The experiments were performed on the eye-lens protein  $\alpha$ B-crystallin ( $\alpha$ Bc) that is heavily crowded under native conditions, as well as on bovine serum albumin (BSA), lysozyme (LYZ), and the Src-homology 3 (SH3) domain. Notably,  $\alpha$ Bc behaves as a hard-sphere particle, i.e., there are effectively no inter-particle interactions besides steric restrictions [Roos et al., 2015b, 2016]. Its retardation of rotational and long-time translational diffusion quantitatively matches the predictions by hard-sphere models [van Blaaderen et al., 1992; Cichocki et al., 1999] even *without* re-scaling to an effective volume fraction.

Based upon the experimental data on translational and rotational diffusion of the four test proteins, with complementary insights from small-angle X-ray scattering (SAXS) experiments performed and analyzed by MARIA OTT, a clear, comprehensive physical picture of crowded protein Brownian dynamics emerged. The central results on crowded *protein* Brownian motion are as follows:

(i) – *The relation of translational and rotational diffusion to the macroscopic viscosity* –

(a) In the absence of transient binding, *long-time* translational diffusion couples to the macroscopic viscosity [Roos et al., 2016]. This relationship holds for monomeric proteins (LYZ) as well as in presence of permanent oligomers (BSA) or in the case of a size distribution of the protein ( $\alpha$ Bc) [Roos et al., 2015b]. Given the case of transient binding (BSA+SH3), in contrast, the slow-down of *long-time* translational diffusion with increasing protein concentration can be *stronger* than the accordant increase in the macroscopic viscosity [Rothe et al., 2016].

*Short-time* translational diffusion is usually less affected by macromolecular caging and concentration effects as long-time translational diffusion. This situation is accompanied by a local viscosity that is usually lower than the macroscopic viscosity. Such a relation has been found in the case of BSA [Roos et al., 2016] by incorporating neutron scattering literature data [Roosen-Runge et al., 2011]. For LYZ, however, a shared concentration dependence of short-time and long-time translational diffusion (and hence the viscosity) has been found [Liu et al., 2011], indicating the presence of strong protein-protein interactions.

(b) Rotational diffusion indeed can but needs not to be sensitive to the macroscopic viscosity. The applicable scenario is protein specific and depends on the strength of anisotropic protein-protein interactions, including charge effects [Roos et al., 2016]. A discussion of this issue will be given below. Whereas rotational diffusion of the hard-sphere like protein  $\alpha$ B-crystallin is almost unaffected by concentration effects, rotational diffusion of LYZ is sensitive to the macroscopic solution viscosity and its concentration dependence.



(ii) – *Translational vs. rotational diffusion: Coupling or Decoupling?* –

(a) Rotational dynamics can be fully coupled to (LYZ) or completely decoupled from ( $\alpha$ Bc) *long-time* translational diffusion, with the critical factor being the sensitivity to the macro-viscosity, and hence the extent of anisotropic inter-molecular interactions including charge effects. Intermediate behavior (BSA) may also arise [Roos et al., 2016].

(b) Rotational diffusion in the absence of transient binding is not expected to be more retarded than *long-time* translational diffusion. Consistent with this expectation, none of the proteins for which rotational diffusion was studied showed such behavior [Roos et al., 2016].

(c) The data on BSA obtained in this study indicate that rotational diffusion of globular proteins is sensitive to similar viscous effects as short-time translational diffusion detected by neutron scattering experiments [Roosen-Runge et al., 2011]. A shared concentration dependence for both short-time quantities has been observed [Roos et al., 2016].

(iii) – *The behavior of the protein RACF in non-dilute solutions* –

The non-exponential nature of the protein RACF arising from inter-protein interactions has been resolved. In a minimal but sufficient model the RACF has been approximated by two superimposed exponentials. In this way, a monotonous increase of the amplitude of the “slow component” with increasing protein concentration is observed (BSA, LYZ) [Roos et al., 2015a]. This aspect stresses the impact of protein-protein interactions on the general behavior of the RACF and reinforces the importance to account for its non-exponential shape under crowding conditions.

The sensitivity of rotational diffusion to the macroscopic solution viscosity as depicted for LYZ can be generally attributed to correlated motions of neighboring particles. In the same way, the shared concentration dependence of short and long-time translational diffusion of this protein [Liu et al., 2011] requires pronounced inter-particle interactions, keeping in mind that particle encounters are less relevant for local as for long-range displacements. Consistent with such requirements, LYZ in solution has been shown to form an intermediate range order structure, the lifetime of which is longer than but comparable to the time span required for displacements over a distance of the size of a LYZ monomer [Liu et al., 2011]. Such inter-molecular alignment (“dynamic clusters”) is in agreement with so-called electro-static steering effects [Krushelnitsky, 2006; Roos et al., 2015a] and requires anisotropic protein-protein interactions. Local ordering effects can be induced by a charge

## 5. Summary, conclusions and outlook

patchiness of the colloidal particles accompanied by short-range attractive interactions despite the same net charge of the particles [McClurg and Zukoski, 1998]. Attractive interactions, in turn, retard even those displacements of a colloid that take place inside the cage formed by surrounding particles [Riest and Nägele, 2015]. This situation holds for rotational and short-time translational diffusion. This situation is also accompanied by electro-viscous effects originating from the somewhat retarded electro-kinetic response of surrounding ions to movements of the charged colloid [McPhie and Nägele, 2007]. For a colloid particle with a non-uniform charge distribution, the latter effect will not only hinder translational displacements, but also retards the rotation of the colloid. Hydrodynamic coupling effects can be considered to be equally relevant for particles of an anisotropic shape. Particularly relevant in case of a strong shape anisotropy, rotational diffusion mediates a cross-over from anisotropic translational diffusion at short times to isotropic translational diffusion at longer times, which is accompanied by a rotation-translation coupling [Han et al., 2006].

After having studied crowded protein Brownian motion, also the translational diffusivity of small solvent molecules has been addressed. These were identified as poor reporters for the viscosity experienced during long-time diffusion of macro-molecularly crowded proteins: steric hindrance is much more relevant for the large proteins than for the small co-solute or solvent particles [Rothe et al., 2016].

From the above findings and concepts it can be concluded that the key factors determining crowded protein Brownian motion are

- (a) the time scale of the motion and the sensitivity to macromolecular caging effects,
- (b) the presence or absence of shape- and/or charge-related anisotropic protein-protein interactions that may mediate correlated motions of neighboring particles, and
- (c) potential transient protein-protein binding.

For globular proteins crowded by the same or other globular proteins, the applicability of the generalized Stokes-Einstein (GSE) and the generalized Stokes-Einstein-Debye (GSED) equations for translational and rotational diffusion, respectively, depends on these factors. For crowding caused by unstructured, entangled (bio)polymers, mesh size and long-range topological effects have to be considered as well.

## Outlook

The results and concepts presented here may encourage further investigations on the impact of charge effects on rotational diffusion. Further progress along these lines may address questions regarding the extent to which electro-viscous effects are based upon direct colloid-colloid interactions, and to which extent ions of the solvent contribute to this effect. Molecular dynamics simulations that combine hydrodynamics with (patchy) charge effects would be well suited for this purpose, but may request sophisticated modeling strategies: hydrodynamics combined with charge effects give a physically complex situation. To be noted, both Coulomb and hydrodynamic interactions [Bartlett et al., 2001] are long range with an  $1/r$  dependence between colloidal particles. Certainly, the interplay between electro-kinetic effects and (rotational) hydrodynamics should also be studied further.

The impact of charge effects on the coupling of rotational and long-time translational diffusion may be further elucidated by exploring the effect of electrostatic screening. Increasing the salt concentration can cause a screening of the colloid's charges at length scales shorter than or comparable to the average inter-molecular distance. Then, for a protein that features a noticeable charge patchiness (such as LYZ), a transition from coupling to decoupling of rotational and long-time translational diffusion may occur with increasing salt concentration. Opposed to that, increasing the number density of dissolved ions may also increase those electro-kinetic friction effects that are mediated by these solvent ions. In this regard, such experiments may also provide insights into solvent-induced electro-viscous effects.

Exploring the effect of electrostatic screening on protein rotational diffusion requires a beforehand study on the stability of the protein solution. LYZ solutions, for instance, undergo phase transitions (including gelation) and aggregation with increasing concentration of monovalent ions (e.g. via adding NaCl; see Rosenbaum et al. [1996] and Sedgwick et al. [2005]). This behavior is different for multivalent ions (e.g.  $Y^{3+}$ ). For negatively charged proteins and a high concentration of multivalent ions, condensation of ions at the protein surface leads to a charge inversion of the protein-ion complex. Under such circumstances, a reentrant phase behavior occurs: the solutions are stable for low and high salt concentrations, but undergo aggregation, crystallization and liquid-liquid phase separation for intermediate concentrations of multivalent ions [Zhang et al., 2008, 2014]. Increasing the *protein* concentration is accompanied by a higher salt concentration at which reentrant phase behavior occurs [Zhang et al., 2014] – crowding studies including multivalent ions may thus require very high salt concentrations. Notably, charge-charge interactions of the protein with multivalent ions can be further tuned by competing non-specific effects of monovalent salts [Jordan et al., 2014]. These effects have to be carefully considered in the study of Brownian dynamics of concentrated LYZ solutions under the condition of strong electrostatic screening.

\* \* \*



---

## SUPPLEMENT

---

– including –

- S.1 Notes on spin precession – PAGE S-1
- S.2 Why using the arithmetically averaged relaxation rate? – PAGE S-1
- S.3 Limitations of PFG NMR in crowded protein solutions – PAGE S-4
- S.4 Viscosity determination via a simple capillary – PAGE S-11
- S.5 Supplementary material to the research articles – PAGE S-13
- S.6 Bibliography – PAGE S-53

– as well as the –

Curriculum vitae  
List of publications  
List of scientific talks and posters  
Statutory declaration

– and the –

Acknowledgments

# S Supporting material, bibliography

## S.1. Notes on spin precession

The evolution of spin states can be described by the time-dependent Schrödinger equation, with the Hamiltonian given by the Zeeman interaction,  $\hat{\mathcal{H}} = \omega_0 \hat{I}_z$ . Here,  $\hat{I}_x$ ,  $\hat{I}_y$  and  $\hat{I}_z$  are spin angular momentum operators. For the time-independent Hamiltonian, the solution of the time-dependent Schrödinger equation is given by  $|\psi(t)\rangle = \exp(-i\omega_0 t \hat{I}_z) |\psi(0)\rangle$ . The state of the spin system  $|\psi(t)\rangle$  is essentially the same as in  $|\psi(0)\rangle$ , reflecting a stationary solution. The exponential phase factor equals a rotation around the z-axis by an angle of  $\phi = \omega_0 t$ , but is *not* a measurable quantity for a *single* Zeeman state such as  $|\mathcal{I}, m_s\rangle = |\frac{1}{2}, \pm\frac{1}{2}\rangle$ : the expectation value of the spin angular momentum contains no *x/y*-components, as is seen from  $\langle \hat{I}_x \rangle = \langle \hat{I}_y \rangle = 0$ . Instead, only a mixed state such as  $\frac{1}{\sqrt{2}}(|\frac{1}{2}, +\frac{1}{2}\rangle + |\frac{1}{2}, -\frac{1}{2}\rangle)$  has a non-vanishing *x/y*-component of the expectation value of the angular momentum. For this particular example,  $\langle \hat{I}_z \rangle = \langle \hat{I}_y \rangle = 0$  and  $\langle \hat{I}_x \rangle = 1/2 \hbar$ . Accounting for the above phase factor, the *mixed* state gives a detectable rotation of a transverse component (precession of the spin ensemble), whereas for a single spin from the same spin ensemble, no information on the *x/y*-components can be obtained. As such, precession should not be attributed to single spins.

## S.2. Why using the arithmetically averaged relaxation rate?

In the following, the use of the arithmetically averaged relaxation rate for relaxometry data analysis will be motivated, with additional insight into how the mean relaxation rate is linked to the protein average internal and Brownian dynamics.

For a *single* spin, spin relaxation as described by an exponential relationship (cf. eq. (3.4)) is not valid; this treatment is only applicable to an ensemble of spins. However, each position within a protein can be found as a chemically equivalent position within

another protein of same kind. As each protein solution contains a multitude of identical proteins, site-specific relaxometry data analysis relying on an exponential relationship remains applicable. The spin relaxation rate of the  $i^{\text{th}}$  protein proton site is sensitive to the spectral density describing the motion at this particular environment,  $J_{\text{int},r}^{(i)}(\omega)$ , which, with regard to eq. (3.12), can be written as

$$J_{\text{int},r}^{(i)}(\omega) = \left(1 - S_{\text{int}}^{2(i)}\right) J_{\text{int}}^{(i)}(\omega) + S_{\text{int}}^{2(i)} J_r(\omega) \quad . \quad (\text{S.1})$$

Here  $J_r(\omega)$  is, as before, the spectral density of Brownian tumbling, and  $J_{\text{int}}^{(i)}$  is the spectral density of internal motions of the  $i^{\text{th}}$  protein proton site, along with its individual order parameter  $S_{\text{int}}^{2(i)}$ . For a sum of different protein sites, the sum of the individual correlation functions matters. The Fourier transform is a linear mathematical operator; thus, the average spectral density simply reads

$$\begin{aligned} \langle J_{\text{int},r}(\omega) \rangle &= \frac{1}{N} \sum_{i=1}^N J_{\text{int},r}^{(i)}(\omega) \\ &\cong \left(1 - \langle S_{\text{int}}^2 \rangle\right) \overline{J_{\text{int}}}(\omega) + \langle S_{\text{int}}^2 \rangle J_r(\omega) \quad . \end{aligned} \quad (\text{S.2})$$

For simplicity, the multitude of different internal correlation times has been replaced here by a single mode,

$$\begin{aligned} \tau_{\text{int}}^{(i)} &\cong \overline{\tau_{\text{int}}} \quad \forall i = \{1, \dots, N\} \\ \Rightarrow J_{\text{int}}^{(i)}(\omega) &\cong \overline{J_{\text{int}}}(\omega) \quad \forall i = \{1, \dots, N\} \quad . \end{aligned} \quad (\text{S.3})$$

Measuring spin relaxation at low frequencies and thus low resolution, the multitude of different internal modes is experimentally indistinguishable. Given an integral protein signal, the above simplification is thus inevitable. This treatment, in turn, allows the mean spectral density to be rewritten using an *average* internal order parameter,

$$\langle S_{\text{int}}^2 \rangle := \frac{1}{N} \sum_{i=1}^N S_{\text{int}}^{2(i)} \quad . \quad (\text{S.4})$$

The average spectral density obtained in this way will be used for the later data analysis.

As for spin relaxation, a similar treatment may be applied: using

$$K_{\text{HH}}^{(i)} = k_i \langle K_{\text{HH}} \rangle \quad , \quad \langle K_{\text{HH}} \rangle = \frac{1}{N} \sum_{i=1}^N K_{\text{HH}} \quad , \quad (\text{S.5})$$

and accounting for the contribution of each protein proton site to the overall spin relaxation process by the sum of the individual relaxation rates, one has

$$\frac{1}{N} \sum_{i=1}^N R_i(\omega) \sim \frac{1}{N} \sum_{i=1}^N K_{\text{HH}}^{(i)} J_{\text{int},r}^{(i)}(\omega) \quad (\text{S.6})$$

$$\stackrel{(\text{S.1})}{=} \frac{1}{N} \sum_{i=1}^N K_{\text{HH}}^{(i)} \left(1 - S_{\text{int}}^{2(i)}\right) J_{\text{int}}^{(i)}(\omega) + \frac{1}{N} \sum_{i=1}^N K_{\text{HH}}^{(i)} S_{\text{int}}^{2(i)} J_r(\omega)$$

$$\stackrel{(\text{S.5})}{=} \langle K_{\text{HH}} \rangle \frac{1}{N} \sum_{i=1}^N k_i \left(1 - S_{\text{int}}^{2(i)}\right) J_{\text{int}}^{(i)}(\omega) + \langle K_{\text{HH}} \rangle \frac{1}{N} \sum_{i=1}^N k_i S_{\text{int}}^{2(i)} J_r(\omega) \quad (\text{S.7})$$

$$\stackrel{(\text{S.2})}{\cong} \langle K_{\text{HH}} \rangle \left\{ \left(1 - \langle S_{\text{int}}^2 \rangle_k\right) \overline{J_{\text{int}}}(\omega) + \langle S_{\text{int}}^2 \rangle_k J_r(\omega) \right\}$$

$$\stackrel{(\text{S.1})}{=} \langle K_{\text{HH}} \rangle \langle J_{\text{int},r}(\omega) \rangle_k . \quad (\text{S.8})$$

Here, the identity

$$\frac{1}{N} \sum_{i=1}^N k_i = \frac{1}{N} \sum_{i=1}^N \frac{K_{\text{HH}}^{(i)}}{\langle K_{\text{HH}} \rangle} \equiv 1$$

has been used. As seen from eq. (S.8), the sum of relaxation rates directly relates to the mean spectral density of motion,  $\langle J_{\text{int},r}(\omega) \rangle_k$ , except for a re-scaled internal order parameter,

$$\langle S_{\text{int}}^2 \rangle_k := \frac{1}{N} \sum_{i=1}^N k_i S_{\text{int}}^{2(i)} \neq \langle S_{\text{int}}^2 \rangle \quad (\text{for most cases}) . \quad (\text{S.9})$$

As the value of the internal order parameter will not be of interest in this study, the issue of  $\langle S_{\text{int}}^2 \rangle_k \neq \langle S_{\text{int}}^2 \rangle$  is not of relevance. Nonetheless,  $\langle S_{\text{int}}^2 \rangle_k$  and  $\langle S_{\text{int}}^2 \rangle$  can be considered to have similar values.

Also note that eq. (S.7) directly relates to the arithmetic average of relaxation rates, as may be expected when relying on a sum of relaxation rates:

$$\begin{aligned} & \langle K_{\text{HH}} \rangle \sum_{i=1}^N \underbrace{k_i N^{-1} \left(1 - S_{\text{int}}^{2(i)}\right) J_{\text{int}}^{(i)}(\omega)}_{p_i} + \langle K_{\text{HH}} \rangle \underbrace{\frac{1}{N} \sum_{i=1}^N k_i S_{\text{int}}^{2(i)} J_r(\omega)}_{p_0} \\ & \sim \langle K_{\text{HH}} \rangle \sum_{i=1}^N p_i \frac{R_{\text{int}}^{(i)}(\omega)}{\langle K_{\text{HH}} \rangle} + \langle K_{\text{HH}} \rangle p_0 \frac{R_r(\omega)}{\langle K_{\text{HH}} \rangle} \\ & = \sum_{i=0}^N p_i R^{(i)}(\omega) , \quad R^{(0)} := R_r , \quad R^{(i)} = R_{\text{int}}^{(i)} \quad \forall i = \{1, \dots, N\} . \quad (\text{S.10}) \end{aligned}$$



Combining eqs. (S.8) and (S.10), it immediately follows that

$$-\frac{d}{dt} \sum_{i=0}^N p_i \exp(-t R^{(i)}(\omega)) \Big|_{t=0} = \sum_{i=0}^N p_i R^{(i)}(\omega) \sim \langle K_{\text{HH}} \rangle \langle J_{\text{int},r}(\omega) \rangle_{\text{k}} . \quad (\text{S.11})$$

The initial slope of the NMR relaxation curve is thus shown to be directly related to the mean spectral density  $\langle J_{\text{int},r}(\omega) \rangle_{\text{k}}$ .

For the above consideration, the simplified relationship  $R(\omega) \sim K_{\text{HH}} J(\omega)$  has been used, where  $R(0) \sim R_2$ ,  $R(\omega_1) \sim R_{1\rho}$  (with  $\omega_1$  being in the kHz regime), and  $R(\omega_0) \sim R_1$  for  $\omega_0 = |\gamma|B_0$ . Accounting for this simplified relationship, not an equality has been inferred but rather a similar scaling behavior (“ $\sim$ ”). Using the actual equations for homo-nuclear spin relaxation (eqs. (3.5)-(3.7)) the physics underlying the above treatment stays the same, with the same conclusions.

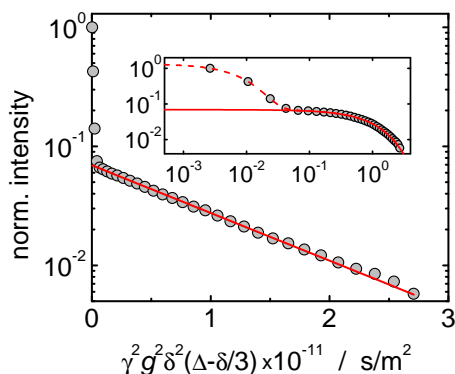
Also note that paper #2 of the RESULTS section gives an additional strong argument for using the initial slope of the relaxation curves.

### S.3. Limitations of PFG NMR in crowded protein solutions

In general, self-diffusion coefficients obtained from PFG NMR measurements are highly reliable. In the presence of slow protein Brownian tumbling, however, bias effects may occur: proton exchange between the protein and the solvent combined with spin polarization exchange inside the protein may lead to an overestimation of the protein diffusion coefficient, even if the aliphatic region of the protein spectrum is evaluated. Details of this effect are discussed below, including experimental data on this issue. The presented concept here serves as a methodical study for addressing the limiting protein concentration up to which this effect is of no relevance. The accordant experiments were performed at a Larmor frequency of 400 MHz.

At those gradient strengths which substantially resolve the protein diffusion decay, no NMR signal of the water protons remains (cf. Fig. S1). This situation enables a safe, straightforward data evaluation despite of the presence of two diffusing species. Though, for protein protons subject to chemical exchange with the solvent on a time scale comparable to or faster than the diffusion delay  $\Delta = \Delta_{\text{PFG}}$ , the protein diffusion decay will be influenced by the lack of magnetization of the solvent. Proton exchange with solvent molecules is of particular relevance for amino groups ( $-\text{NH}_2$ ,  $-\text{NH}_3^+$ ). For aliphatic protein protons, in contrast, chemical exchange with the solvent is usually negligible. Integration over the aliphatic region of the spectrum thus *appears* to be safe.

Spin polarization exchange among protein protons can cause the magnetization of aliphatic protons to not be independent from the magnetization of the amide protons; a



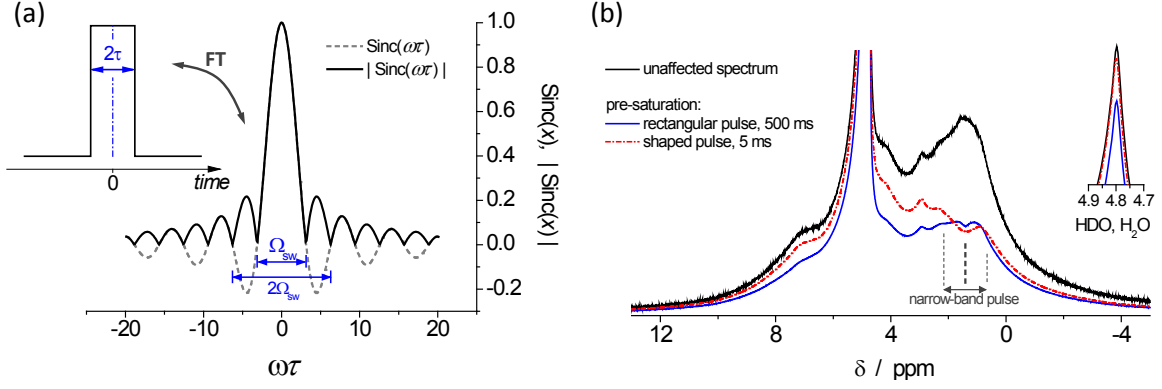
**Figure S1.:** Exemplary PFG NMR diffusion decay including a few data points of weak gradient strength at which the water signal is seen, resembling a bi-exponential decay with well separated slopes (i.e., diffusion coefficients).

magnetization exchange among different magnetization reservoirs occurs. If the period required for spin polarization exchange between amide and aliphatic protons is also comparable to the diffusion time scale, amide protons in exchange with water protons act as magnetization sinks for the aliphatic protons. Then, the diffusion decay of the aliphatic region may be distorted. Integrating over different spectral regions of the PFG NMR diffusion signal for  $\alpha$ B-crystallin the slope (and hence the diffusion coefficient obtained) in fact differs despite reflecting the diffusivity of one and the same protein. Exemplary data on this issue are shown in the Supporting Material for paper #1.

Generally, spin polarization exchange gets more pronounced with stronger spin-spin interactions. Any dipole-dipole coupling between solvent and protein protons is too weak to mediate an effective polarization transfer; J-couplings are also absent in this case. For that reason, direct spin polarization exchange can only occur among protein protons, but not between the solvent and the protein. Magnetization can only be transferred from the protein to the solvent (or *vice versa*) via mutual exchange of protons. To estimate the relevance of combined chemical and spin polarization exchange, the following experiment was performed:

Long-lasting pulses are narrow-banded regarding the frequencies they contain.<sup>1</sup> This fact can be easily understood via the Fourier transform of a rectangular function of duration  $2\tau$ , which results into a sinc function,  $\sin(\omega\tau)/(\omega\tau)$ . The spectral width  $\Omega_{\text{sw}}$  primarily excited by the rectangular pulse may be defined on basis of the first zero-crossing of the sinc function, i.e.  $\Omega_{\text{sw}} = \pi/\tau$  (cf. Fig. S2a). As a secondary effect,  $R_{1\rho}$  relaxation will become the more significant the longer the pulse, ultimately causing an (almost) complete loss of signal. As such, long-lasting pulses yield a “spectral hole burning” (cf. Fig. S2b) that is used here to affect predominately the magnetization of aliphatic protein protons. By irradiation on the resonance frequency of water, one would also affect the protein due to an intrinsic overlap of protein resonances with the resonance frequency of water. Thus, spectral hole burning was applied to the aliphatic region of the protein, and the impact on the water signal was monitored. By inserting a variable mixing time  $t_m$  in-between the

<sup>1</sup>For this reason, one usually relies on rf pulses of a few  $\mu\text{s}$  only, sufficient to excite the full spectrum.



**Figure S2.:** The principle of spectral hole burning. Each pulse of finite length contains a distribution of frequencies around the central carrier frequency of the pulse. The longer the pulse, the narrower the distribution of involved frequencies (a). This effect can be used to selectively affect only part of the spectrum (b). Vertical dashed lines give the spectral range at which the pulse is primarily active. Note that the use of a shaped pulse allows for a shorter pulse duration, with almost no effect on the water signal. The very long-lasting pulse, in contrast, causes some reduction of the water signal. For a discussion of this effect, see the text.

selective manipulation of the protein magnetization and the signal detection, the time scale of the overall spin polarization transport from protein aliphatic protons to solvent protons can be monitored.

Having previously saturated part of the magnetization by a selective pulse, not only does magnetization transfer occur, but also longitudinal spin relaxation. Thus, the experiment has to be designed in such a way that it compensates for this effect. For this purpose, the magnetization has been stored along both  $+z$  and  $-z$  during the mixing time  $t_m$ , with separate signal detection for the two cases. The detected signal belonging to the case that the magnetization was stored along the  $(-z)$ -axis will be referred to as  $M_-(t_m)$ ; for storage along the  $(+z)$ -axis, it is denoted as  $M_+(t_m)$ . Although not indicated by indices here,  $M_{\pm}(t_m)$  will always be evaluated with respect to a certain component  $i$  of the sample, e.g. the solvent peak. The pulse sequence resulting from these specifications is displayed in Fig. S3.

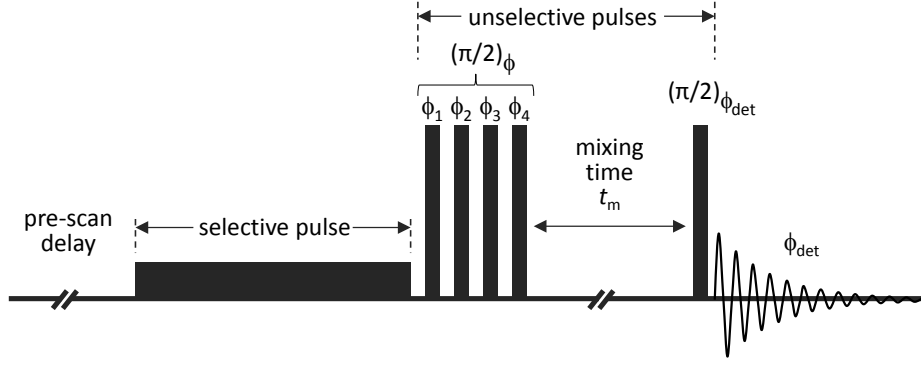
For data evaluation, consider that  $M_i^{(\text{ex})}$  is the relative magnetization of the component of interest (index  $i$ ) dependent on the mixing time, with  $0 \leq M_i^{(\text{ex})}(t_m) \leq 1 \forall t_m \in \mathbb{R}$ . Here, the case of  $M_i^{(\text{ex})} = M_w^{(\text{ex})}$  will belong to the residual water protons (HDO, H<sub>2</sub>O). The overall behavior of the magnetization can be approximated by

$$M_-(t_m)/M_{\infty} \cong 1 - \left(1 + M_i^{(\text{ex})}(t_m)\right) \cdot r(t_m) \quad (\text{S.12})$$

$$M_+(t_m)/M_{\infty} \cong 1 - \left(1 - M_i^{(\text{ex})}(t_m)\right) \cdot r(t_m) \quad , \quad (\text{S.13})$$

where  $M_{\infty}$  is the equilibrium magnetization. The factor of

$$r(t_m) \cong \exp(-t_m R_1) \quad (\text{S.14})$$



**Figure S3.:** Pulse sequence for studying spin polarization exchange between aliphatic protein and residual water protons. Duration and power level of the pulses are not to scale. The block of the four  $(\pi/2)$ -pulses acts as a flip-angle deviation compensated composite pulse serving for  $(\pm z)$ -storage, with  $\phi_1 = x$ ,  $\phi_2 = y$ ,  $\phi_3 = y$ ,  $\phi_4 = x$  for  $(-z)$ -storage, and  $\phi_1 = x$ ,  $\phi_2 = -x$ ,  $\phi_3 = -x$ ,  $\phi_4 = x$  for  $(+z)$ -storage.

accounts for longitudinal spin relaxation that is addressed here via a single relaxation time  $R_1$ . (For a distribution of relaxation times, see below.) Obviously, eqs. (S.12) - (S.14) fulfill the conditions

$$r(0) = 1 \quad , \quad \lim_{t \rightarrow \infty} r(t_m) = 0 \quad , \quad (\text{S.15})$$

$$M_-(0)/M_\infty = -M_i^{(\text{ex})}(0) \quad , \quad \lim_{t \rightarrow \infty} M_-(t_m)/M_\infty = 1 \quad , \quad (\text{S.16})$$

$$M_+(0)/M_\infty = M_i^{(\text{ex})}(0) \quad , \quad \lim_{t \rightarrow \infty} M_+(t_m)/M_\infty = 1 \quad . \quad (\text{S.17})$$

Defining

$$\Delta_M(t_m) := \frac{M_-(t_m)}{M_\infty} - \frac{M_+(t_m)}{M_\infty} = -2M_i^{(\text{ex})}(t_m)r(t_m) \quad , \quad (\text{S.18})$$

$$\Sigma_M(t_m) := \frac{M_-(t_m)}{M_\infty} + \frac{M_+(t_m)}{M_\infty} = 2 - 2r(t_m) \quad , \quad (\text{S.19})$$

it follows that

$$\frac{\Delta_M(t_m)}{\Sigma_M(t_m) - 2} = M_i^{(\text{ex})}(t_m) \quad . \quad (\text{S.20})$$

Eq. (S.20) reflects the relative magnetization of the wanted component (here: water) as a function of the mixing time  $t_m$ , with the effect of longitudinal relaxation being canceled out. Note that eqs. (S.18) and (S.19) also demonstrate that simply adding up the NMR signal observed after  $\pm z$ -storage is not sufficient for canceling out the effect of longitudinal relaxation.

The assumption of a single relaxation rate may indeed be valid for the protein, given that spin polarization exchange occurs on a time scale much faster than spin relaxation. At high magnetic fields, this is usually the case; see also paper #2 of the RESULTS section. Nevertheless, the relaxation rate of water ( $R_1^{(w)}$ ) remains different from that of the different

protein sites ( $R_1^{(p)}$ ). Usually, it holds that  $R_1^{(w)} < R_1^{(p)}$ . Thus, for a more representative treatment, eqs. (S.12) - (S.14) should be rather replaced by

$$M_{\pm}(t_m)/M_{\infty} \cong 1 - \sum_i \left(1 \mp M_i^{(\text{ex})}(t_m)\right) p_i r_i(t_m) \quad (\text{S.21})$$

$$r_i(t_m) \cong \exp(-t_m R_{1,i}) \quad , \quad \sum_i p_i = 1 \quad ,$$

using individual relaxation rates  $R_{1,i}$  and a proton fraction of  $p_i$  for each component contributing to the signal. Following the same procedure as before, it follows that

$$\frac{\Delta_M(t_m)}{\Sigma_M(t_m) - 2} \cong \frac{\sum_i M_i^{(\text{ex})}(t_m) p_i r_i(t_m)}{\sum_i p_i r_i(t_m)} \quad (\text{S.22})$$

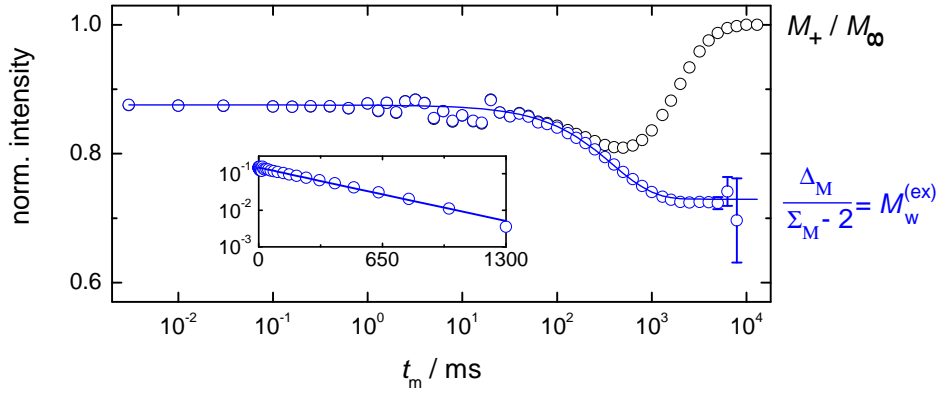
$$\cong \frac{M_w^{(\text{ex})}(t_m) p_w r_w(t_m) + M_p^{(\text{ex})}(t_m) p_p r_p(t_m)}{p_w r_w(t_m) + p_p r_p(t_m)} \quad ,$$

where the indices  $w$  (for  $i = 1$ ) and  $p$  ( $p_p = \sum_{i \geq 2} p_i$ ) reflect, as before, the water and the protein signal, respectively. Evaluating the water peak of the NMR spectrum,  $p_w \sim 0.9 > p_p \sim 0.1$  holds, and, again,

$$\frac{\Delta_M(t_m)}{\Sigma_M(t_m) - 2} \Bigg|_{\substack{\text{H}_2\text{O} \\ \text{peak}}} \approx M_w^{(\text{ex})}(t_m) \quad . \quad (\text{S.23})$$

An exemplary data set including the above treatment is shown in Fig. S4. As seen from this graph,  $M_{+}/M_{\infty}$  equals  $M_w^{(\text{ex})}$  for  $t_m \lesssim 10$  ms. Even though the selective pulse is much longer than the time scale of  $R_{1,p}$  relaxation, a small fraction of transverse relaxation remains. The precession of this minor component causes some signal distortions for mixing times in-between 1 ms and  $\sim 50$  ms. Below this time scale, no significant evolution of the transverse magnetization is observed; for longer times, the spin coherence vanishes due to  $T_2$  relaxation.

The duration of this pulse (here: 500 ms, applied to aliphatic protons) may be comparable to the time scale needed for magnetization transfer between the protein and the solvent. Under such conditions, the intensity of the water peak may be reduced even before the actual exchange time  $t_m$  starts. In fact, in the shown example (Fig. S4) the initial intensity of the water peak is only of about 88% of its equilibrium intensity  $M_{\infty}$ . Nevertheless, as soon as the pulse is switched off, the continuation of this process can be monitored. The original signal,  $M_{+}/M_{\infty}$ , further drops down by about 7%, and then rises to its equilibrium intensity; cf. again Fig. S4. Such behavior already indicates that the overall magnetization transport from the aliphatic region of the protein to the solvent is somewhat faster than  $T_1$  relaxation, but not by orders of magnitude. The  $T_1$ -corrected signal, eq. (S.20), reflects the underlying loss of water intensity due to transfer of magnetization to the protein. Compensating longitudinal relaxation by data treatment, an equilibrium is reached after



**Figure S4.:** NMR water intensity  $M_+/M_\infty$  profile (black) and the water magnetic exchange curve  $M_w^{(\text{ex})}$  (blue) of  $\alpha$ B-crystallin at a concentration of 180 g/l at  $T = 22^\circ\text{C}$ , including the fitting result (blue line; eq. (S.24)). The inset shows the magnetization exchange curve after offset subtraction on a semi-logarithmic scale.

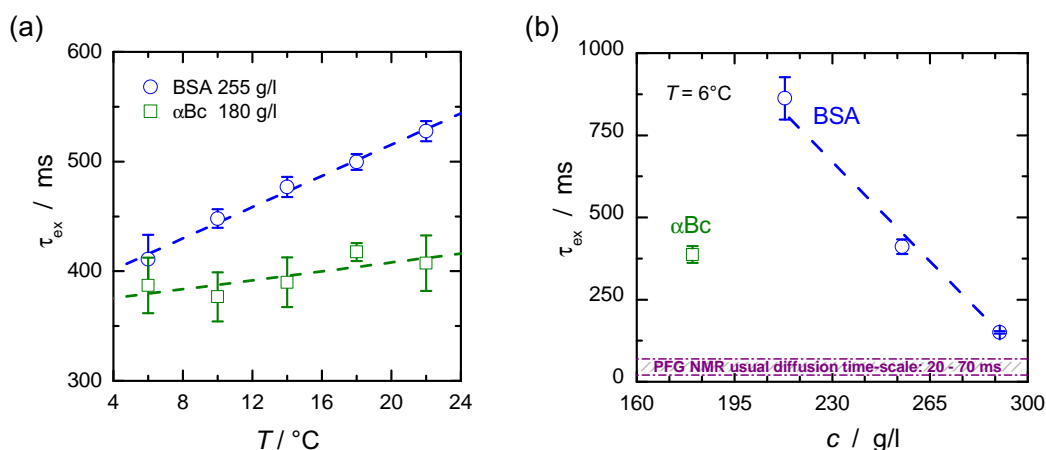
about 3 s. This reduced equilibrium magnetization reflects the overall magnetization of the sample (averaged over protein *and* water molecules) remaining after the selective pre-saturation pulse. Subtracting this reduced equilibrium value and plotting the decay on a semi-logarithmic scale, an exponential-like behavior of  $M_r$  is resolved (inset in Fig. S4). The magnetization exchange decay of the water signal may thus be modeled on the basis of a single characteristic magnetization exchange time  $\tau_{\text{ex}}$ , using

$$M_w^{(\text{ex})}(t_m) = (M_0^{(\text{ex})} - M_\infty^{(\text{ex})}) \exp(t_m/\tau_{\text{ex}}) + M_\infty^{(\text{ex})} \quad , \quad (\text{S.24})$$

where  $M_0^{(\text{ex})} := M_w^{(\text{ex})}(t_m = 0) \equiv M_+(t_m = 0)/M_\infty$  and  $M_\infty^{(\text{ex})} := \lim_{t_m \rightarrow \infty} M^{(\text{ex})}(t_m)$  are the initial and equilibrium intensity after application of the pre-saturation pulse in absence of longitudinal relaxation.

Fig. S5a displays the temperature dependence of the magnetization exchange rate of BSA and  $\alpha$ B-crystallin, each at their highest concentration addressed within the study of Brownian motion. As seen in this graph,  $\tau_{\text{ex}}$  decreases monotonically with temperature, which is contrary to the scenario that the magnetization transfer is dominated by proton exchange: chemical exchange becomes faster with increasing temperature. Such temperature dependence thus reflects a dominating impact of the spin polarization exchange rate, which becomes faster at low temperatures due to a slower tumbling of the protein. This finding is line with the magnetic exchange rate being faster for the large  $\alpha$ B-crystallin assemblies ( $\tau_r \sim 1 \mu\text{s}$ ) than compared to BSA ( $\tau_r \sim 0.1 \mu\text{s}$  at 250 g/l).

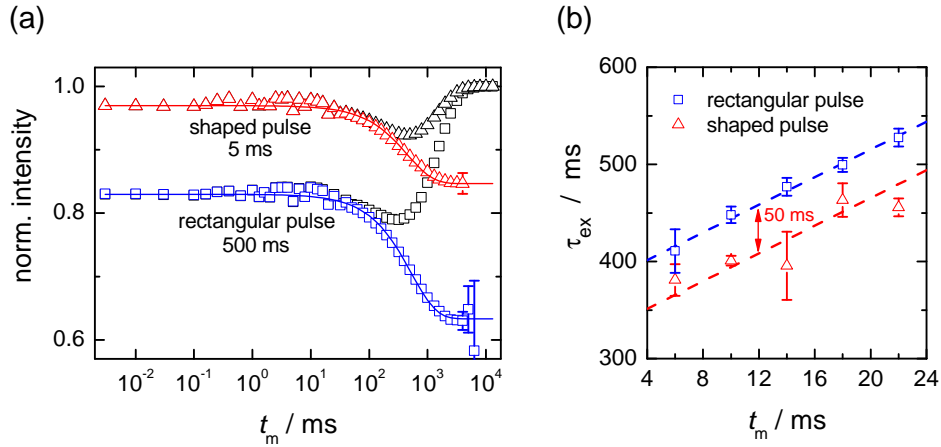
If the protein impacts the signal of water, water will also impact the protein magnetization. If the time scale of magnetization exchange is on the same order of magnitude as the PFG NMR diffusion period,  $\tau_m \sim \Delta_{\text{PFG}} \sim 20\dots70 \text{ ms}$ , then PFG NMR diffusion results



**Figure S5.:** Magnetization exchange time for BSA and  $\alpha$ B-crystallin ( $\alpha$ Bc) as a function of the temperature (a) and the protein concentration (b), with the trend indicated by linear fits (dashed lines). Error bars relate to the fitting uncertainty (eq. (S.24)).

may indeed be biased despite evaluating the signal of aliphatic protons. At  $6^\circ\text{C}$ , there is still a factor of about eight in-between  $\tau_{\text{ex}}$  and  $\Delta_{\text{PFG}}$  for both proteins. Increasing the protein concentration can decrease this factor. As BSA is much more available at high concentrations than  $\alpha$ B-crystallin, the concentration dependence of  $\tau_{\text{ex}}$  was addressed using BSA. Increasing its concentration to a value of about 300 g/l,  $\tau_{\text{ex}}$  approaches  $\Delta_{\text{PFG}}$ ; cf. Fig. S5b. Under such conditions, the PFG NMR diffusion decay can indeed be considered to be biased by the impact of water, i.e., by its faster diffusion. A value of around 250 g/l BSA thus appears to be a reasonable limit for reliable PFG NMR measurements. As is pointed out within the RESULTS section, LYZ has a steeper concentration dependence of its Brownian tumbling time than BSA, but is also smaller than BSA. At 250 g/l,  $\tau_r$  of LYZ is still smaller than  $\tau_r$  of 250 g/l BSA [Roos et al., 2016]. PFG NMR at 250 g/l LYZ can thus be considered to provide safe results as well. Note that for  $\alpha$ B-crystallin, 180 g/l is the highest concentration accessed in the study of Brownian motion.

To excite only a narrow frequency window, a long rectangular pulse was used. Applying a shaped pulse (sinc profile), the duration of the pre-saturation pulse was shortened to only 5 ms (rather than 500 ms for the rectangular pulse). Consequently, the pulse duration is now safely shorter than the characteristic time scale of magnetization exchange between the protein and the solvent. This situation leads to an almost full intensity ( $\approx 97\%$ ) of the water signal (see Figs. S2 (inset) and S6a) straight after action of the selective pre-saturation pulse. More importantly, the shorter pulse duration basically helps resolve faster magnetization exchange rates. Indeed, the magnetization exchange curves become to some extent non-exponential when applying this rather “short” pulse, which potentially results from a distribution of magnetization exchange times. Lacking knowledge on the detailed shape of the magnetization exchange curve, and with some uncertainty of the



**Figure S6.:** Exemplary experimental result of BSA 255 g/l at  $T = 22^\circ\text{C}$  (a) using a shaped pulse of 5 ms duration (triangles) as compared to the result using a long rectangular pulse of 500 ms duration (squares), together with a mono-exponential fit of the magnetization exchange curves (lines). The temperature dependence of the so obtained magnetization exchange times is shown in (b).

equilibrium value  $M_\infty^{(\text{ex})}$ , these curves were still fit using a single exponential (cf. eq. (S.24) and Fig. S6a). The agreement with the experimental data is still sufficiently good; the distribution of magnetization exchange times is not expected to cover several orders of magnitude. For BSA 255 g/l, the now obtained value of  $\tau_{\text{ex}}$  is about 50 ms shorter than that one determined from measurements using the rectangular pulse of 500 ms duration; see Fig. S6b. To estimate the order of magnitude of  $\tau_{\text{ex}}$ , and for its comparison with the PFG NMR diffusion time scale, such differences are not significant. All drawn conclusions stay thus valid for both, the rectangular 500 ms pulse and the shaped 5 ms pulse.

#### S.4. Viscosity determination via a simple capillary

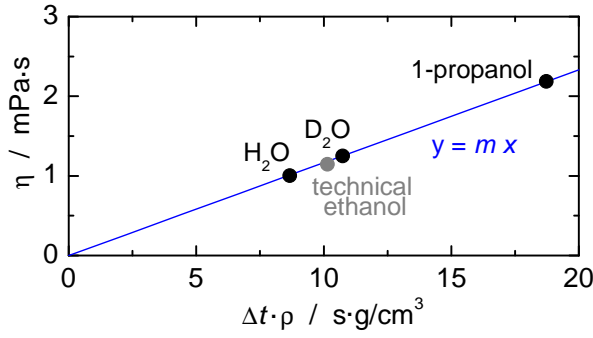
An initial study of the viscosity concentration dependence relied on a simple capillary using sample volumes of around 200  $\mu\text{l}$ . The viscosity is estimated here by the equation by HAGEN and POISEUILLE (eq. 3.22),

$$\Phi = \frac{dV}{dt} = \frac{\pi a^4 |\Delta p|}{8\eta L} .$$

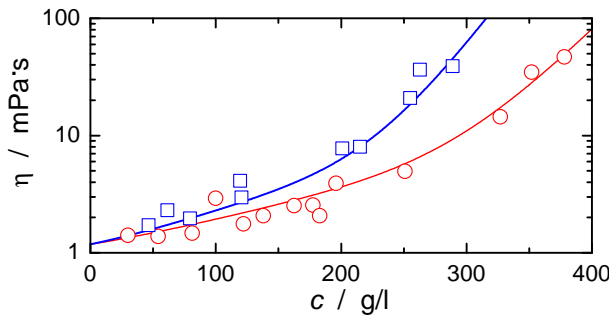
For a vertical capillary, with no force applied except of the gravitational one ( $F_G$ ), the drop of pressure  $\Delta p$  per distance  $L$  along the capillary is determined by

$$\frac{|\Delta p|}{L} = \frac{1}{L} \frac{F_G}{A} = \frac{g\rho AL}{AL} = g\rho , \quad (\text{S.25})$$





**Figure S7.:** Viscosity determination using a capillary: Calibration via substances of known viscosity and density. Technical ethanol containing impurities is slightly off the line  $y = mx$ , and was not used for calibration.



**Figure S8.:** Viscosity of bovine serum albumin (blue) and lysozyme (red) solutions as determined by a capillary at  $T = 22^\circ\text{C}$ . The solid line is a fit using the equation by Thomas [1965],  $\eta = \eta(\phi)$ , including an apparent volume fraction  $\phi_{\text{app}} = \vartheta_{\text{app}} c$ .  $\vartheta_{\text{app}}$  is the apparent specific volume of the protein.

where  $A$  is the cross-sectional area of the capillary slit, and  $\rho$  and  $g$  denote the solution density and the acceleration due to gravity, respectively. Thus, we have, averaged over a certain length of the capillary,

$$\frac{\pi g \rho \langle a^4 \rangle}{8 \eta} \cong \frac{\Delta V}{\Delta t} = A \frac{\Delta s}{\Delta t} \quad . \quad (\text{S.26})$$

Here,  $\Delta s$  is the distance along the capillary the liquid covers per time step  $\Delta t$ . In principle, this equation could be directly used to estimate the viscosity of the protein solution of interest. However, the exact mean diameter of the capillary is an uncertain parameter, therefore a calibration has to be done using a liquid of known viscosity and density (termed  $\eta_1$  and  $\rho_1$ , respectively), providing

$$\frac{\Delta s_2}{\Delta s_1} \frac{\eta_2}{\eta_1} = \frac{\Delta t_2 \rho_2}{\Delta t_1 \rho_1} \quad . \quad (\text{S.27})$$

In the experiments  $\Delta s = \Delta s_1 = \Delta s_2 = 80 \text{ cm}$  was kept constant, and  $\Delta t_{1,2}$  is the quantity to be measured. The way the calibration was performed in practice is reflected in Fig. S7; Fig. S8 shows the data obtained for BSA and LYZ.

Each time the capillary is used, a part of the protein solution remains on the inner wall of the capillary. Although the capillary was rinsed thoroughly with water immediately after each measurement, often a decrease of the effective inner diameter was observed, as

indicated by the duration water needed to pass a certain distance along the capillary. Then, urea served as a useful cleaning agent.

For protein solutions of a high viscosity,  $\Delta t$  occurred to be as long as several minutes, and drying effects were observed. This effect prevented a safe estimate of the viscosity for very highly concentrated protein solutions, with consequently some systematic deviations of the measured viscosity as compared to the measuring result using the mVROC or a rheometer. These deviations become relevant for concentrations above 250 g/l; see Fig. 3.17. Moreover, measurements could only be performed at  $T = 22^\circ\text{C}$  as the capillary was not located within a heating or cooling reservoir.

## **S.5. Supplementary material to the research articles**

Part of the research articles contain Supporting Information that are given below.

# Supporting Material

## NMR-detected Brownian dynamics of $\alpha$ B-crystallin over a wide range of concentrations

Matthias Roos, Susanne Link, Jochen Balbach, Alexey Krushelnitsky,\* Kay Saalwächter\*

Institut für Physik, Martin-Luther-Universität Halle-Wittenberg,  
Betty-Heimann-Str. 7, 06120 Halle (Saale), Germany.

### Sample preparation.

*Protein expression and purification of human  $\alpha$ B-crystallin.*

After cloning the human  $\alpha$ B cDNA into a modified, His-tag free pET16b vector and expression in *Escherichia coli* BL21(DE3), protein expression and purification were performed in a manner similar to ref. (1). Briefly,  $\alpha$ B-crystallin was expressed overnight at 22°C in the auto-induction media ZYM 5052, lysed by Microfluidizer in 20mM TrisHCL (pH 8.5) and 1mM EDTA buffer, the DNA digested by DNase1 and precipitated by protamine sulfate salt. After purification on a TMAE anion-exchange column with a stepwise NaCl gradient,  $\alpha$ B fractions were pooled, concentrated and loaded on a S200 gel filtration column with a running buffer containing 20mM TrisHCL (pH 7.5), 50mM NaCl and 1mM EDTA. The purity of the protein was confirmed by SDS-page and mass spectrometry.

*NMR sample preparation.*

After extensive dialysis against 50mM ammonium hydrogen carbonate buffer (pH 8), the protein was lyophilized, then dissolved in the smaller amount of the same buffer in D<sub>2</sub>O and lyophilized again. Note that multiple lyophilization does not affect  $\alpha$ B-crystallin properties, as confirmed by NMR spectroscopic and diffusion experiments. The lyophilized  $\alpha$ B-crystallin powder was

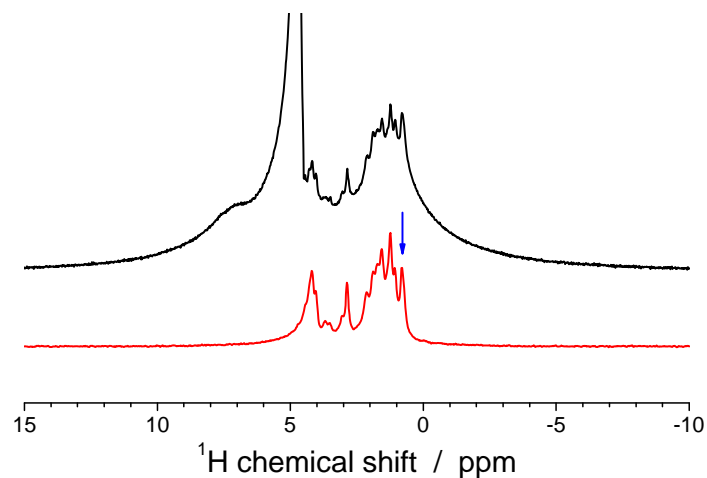
dissolved in 50mM Na-phosphate buffer, 50mM NaCl and 0.002% NaN<sub>3</sub> D<sub>2</sub>O buffer at pH 7.6 (pD=7.2) and used in the NMR and viscosity measurements. The  $\alpha$ B-crystallin concentration was determined by spectral photometry at 280 nm with MW=20027.7 Da and  $\epsilon$ =13980 M cm<sup>-1</sup>.

## **NMR experiments.**

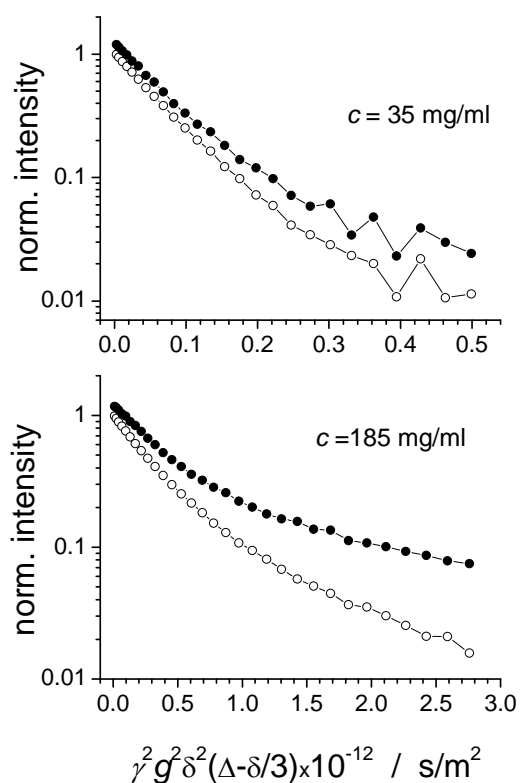
### *Translational diffusion*

Translational self-diffusion coefficients were measured using the stimulated echo technique with bipolar pulsed field gradients (2). Fig. S1 shows proton spectra of  $\alpha$ B-crystallin in D<sub>2</sub>O buffer after a 90°-pulse (top). The strong signal at ~5 ppm corresponds to residual solvent protons that cannot be fully removed during the sample preparation. The NMR spectrum seen in the PFG experiments is shown at the bottom, in which the water signal is filtered out on the basis of its fast diffusion. A large fraction of the protein signal is also filtered out due to the short  $T_2$  of the residues forming the rigid core, while the remaining signal belongs to the protons of mobile unstructured termini of  $\alpha$ B-crystallin polypeptide chains that have long  $T_2$ . It is noted that without these unstructured termini the PFG measurement of translational diffusion would be impossible:  $\alpha$ B-crystallin is a large protein, so its overall rotational tumbling is slow. If all parts of the protein were rigid, then the complete protein signal would be  $T_2$ -filtered (suppressed) during the field gradient pulse, which has a typical duration of 1 - 1.5 ms.

Analyses of the diffusion decay are based upon the right-hand side peak belonging to methyl protons, as marked by an arrow in Fig. S1. Fig. S2 compares the PFG NMR diffusion decay of the integral signal to that of the methyl protons peak. It is clearly seen that the two decays differ, with the difference becoming more pronounced at higher concentrations. Such a difference can be explained by the effect of spin diffusion between protein protons combined with hydrogen exchange of labile protein protons with the residual solvent protons. Thus, the apparent diffusion decay for the integral protein signal is distorted by the magnetic/chemical exchange processes (3,4). This effect is negligible in small and medium-sized proteins because spin diffusion is rather slow due to the much faster overall rotational tumbling, which averages out inter-proton dipole-dipole interactions. The methyl protons are less prone to such distortions since they undergo fast rotation around the C<sub>3</sub> axis, and thus have weak magnetic coupling to other protein protons.

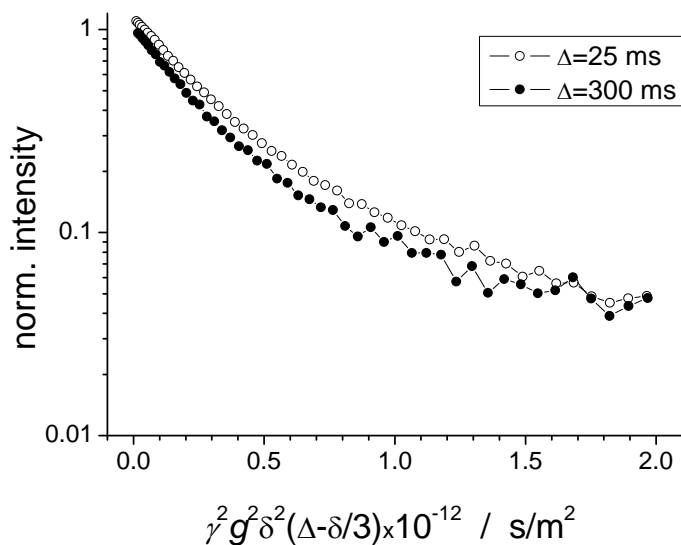


**Figure S1.** Proton NMR spectrum of  $\alpha$ B-crystallin at 400 MHz resonance frequency after a  $90^\circ$  pulse (top), and after  $T_2$ /diffusion filtering during the pulse-gradient experiment (bottom). The blue arrow indicates the peak which was used for evaluating the diffusion decays.



**Figure S2.** Diffusion decays for  $\alpha$ B-crystallin solution at two concentrations ( $T = 20^\circ\text{C}$ ) plotted for the integral signal (open circles) and the methyl peak marked by an arrow in Fig. S1 (solid circles).

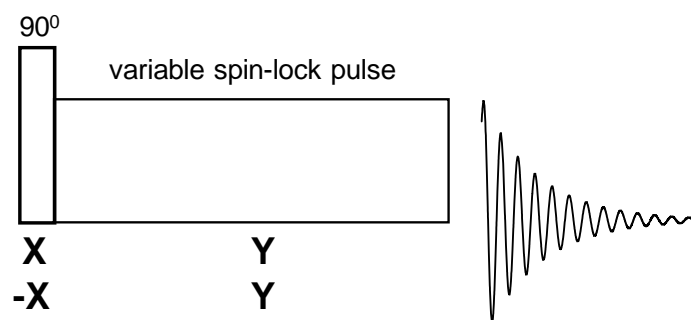
Fig. S3 compares the diffusion decays measured at two different diffusion times ( $\Delta$ ). The SDC should not depend on  $\Delta$ , however, if the characteristic time of the magnetic/chemical exchange processes is comparable to  $\Delta$ , then the apparent diffusion decay will become faster as  $\Delta$  increases. It is seen that the diffusion decay at  $\Delta = 300$  ms is indeed somewhat faster, yet the difference is negligibly small. In our PFG experiments  $\Delta$  was always between 25 and 40 ms, thus the effect of the magnetic/chemical exchange can be safely neglected.



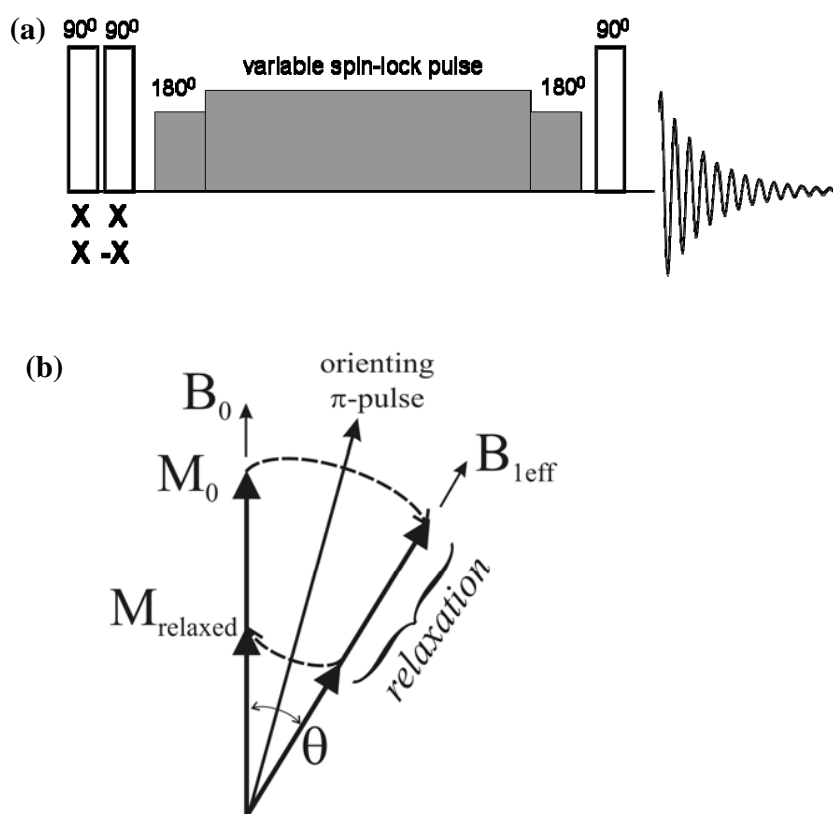
**Figure S3.** Diffusion decays measured at two different diffusion times (indicated in the plot) plotted for the methyl protons peak ( $c = 185$  mg/ml,  $T = 28$  °C).

### *Spin relaxation*

On-resonance  $T_{1\rho}$ 's at 20 and 40 kHz spin-lock frequencies were measured with the standard pulse sequence, Fig. S4. Off-resonance  $T_{1\rho}$ 's at 60 kHz were measured with the sequence shown in Fig. S5a. The latter sequence starts with two  $90^\circ$ -pulses, of which the first pulse is of fixed phase, while the second one is phase-alternated to ensure spin temperature inversion. In doing so, the relaxation signal decays exactly to zero, thus there is no need to measure the long relaxation delay plateau. In addition, the off-resonance spin-lock pulse is flanked by two orienting off-resonance  $180^\circ$ -pulses: the first orienting pulse aligns the magnetization vector along the  $B_{1e}$  field and the second one brings the magnetization back to  $B_0$  direction. By this, the orienting pulse vector forms the angle  $\theta/2$  with the  $B_0$  (and  $B_{1e}$  as well) vector, see Fig. S5b.



**Figure S4.** Pulse sequence for measuring  $T_{1\rho}$  decays with on-resonance spin-lock field.



**Figure S5.** (a) Pulse sequence for measuring off-resonance  $T_{1\rho}$ . Shaded pulses are applied with the resonance offset. (b) Vector scheme demonstrating magnetization perturbations during the off-resonance pulses.

Before each  $T_{1\rho}$  measurement the spin-lock frequencies and the angle  $\theta$  (for the off-resonance experiments) were carefully calibrated using a nutation experiment. For all the off-resonance measurements the angle  $\theta$  was equal to  $42^\circ$ . The spin-lock field duration in the  $T_{1\rho}$  experiments

was varied from few  $\mu\text{s}$  to 200 ms. For plotting the relaxation decays, the whole spectral window (see Fig. S1, top) was used for integration of the proton signal. To avoid sample heating by long spin-lock pulses, the recycling delay was 20 s. The chemical shift of the residual water protons peak did not depend on the duration of the spin-lock pulse which confirms that the sample heating effect was negligible.

Spin-spin ( $T_2$ ) relaxation decays were measured by a combination of three different experiments: Free induction decay (FID; time range from 12 (dead time) to 40  $\mu\text{s}$ ), Hahn echo (from 30  $\mu\text{s}$  to  $\sim 3$  ms) and a Carr-Purcell-Meiboom-Gill sequence (from 0.5 ms to  $\sim 0.4$  s). This was done in order to cover a wide range of the relaxation times, as the protein  $T_2$  is very short, usually tens to hundreds of  $\mu\text{s}$ , whereas the solvent  $T_2$  was of the order of hundreds of ms. Since  $T_2$  relaxation decays were measured at the resonance frequency 20 MHz and imperfect  $B_0$  field homogeneity, no spectral resolution could be achieved at these conditions and thus, the decays were recorded directly in the time domain.

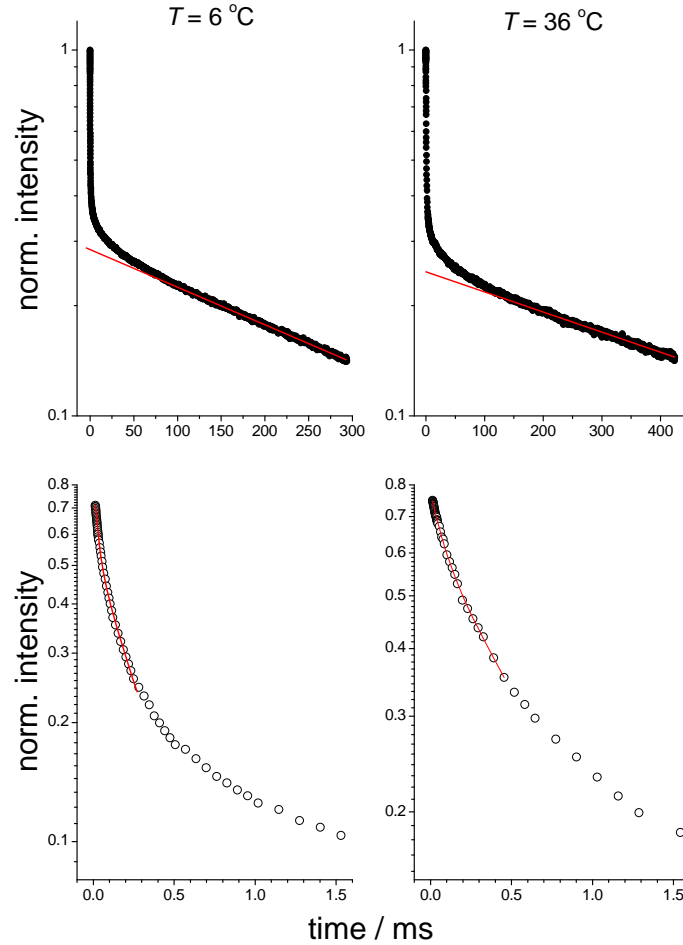
The analysis of the  $T_{1\rho}$  and  $T_2$  relaxation decays was performed according to the procedure described in ref. (5). As mentioned above, protein solutions contain a certain amount of residual solvent protons which exhibit much longer  $T_2/T_{1\rho}$ 's, thus requiring to subtract the solvent signal for analyzing the unbiased protein signal (cf. Fig. S6). For plots and analyses, the relaxation decays of the total (integral) proton signal was detected without any spectroscopic separation in both the  $T_{1\rho}$  and  $T_2$  experiments. The water signal subtraction procedures (Fig. S6) for the  $T_2$  and  $T_{1\rho}$  decays were fully identical.

After subtraction of the solvent signal we obtained the protein protons' relaxation decay, which is of multi-exponential shape (Fig. S6). This decay reflects a wide distribution of the relaxation times which is a consequence of the dynamic heterogeneity of  $\alpha$ -crystallin. For such a decay we determined the mean relaxation rate/time, which equals the slope of the initial part of the decay. In practice, we fitted the decays with a sum of two exponential components, which provides a minimal fitting model. We then determined the mean relaxation rate as

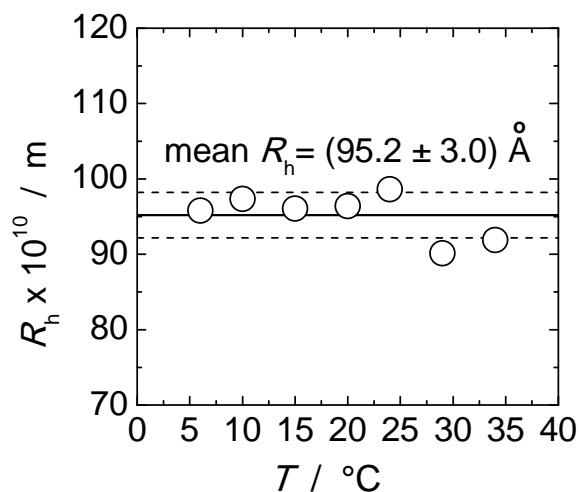


$$\left\langle \frac{1}{T_2} \right\rangle = \frac{1}{P_a + P_b} \left( \frac{P_a}{T_{2a}} + \frac{P_b}{T_{2b}} \right),$$

where  $P_{a,b}$  and  $T_{2a,b}$  are the intensities and relaxation times of the two components, respectively (taken separately, these parameters have no physical meaning).



**Figure S6.**  $T_2$ -relaxation decays measured for  $\alpha$ B-crystallin solutions at  $c = 80$  mg/ml and two temperatures (indicated in the figure).  $T_{1\rho}$ -decays have very similar shapes. *Top*: raw relaxation decays consisting of the fast (protein) and slow (solvent) relaxing components. Red lines indicate the solvent component that was defined from the exponential fit of the slow tail of the relaxation decay. *Bottom*: the protein relaxation signal after the subtraction the solvent component from the overall decay. Red curves correspond to the biexponential fit of the initial part of the decay, which was used for the determination of the mean relaxation time (initial slope of the decay) as described in the text. Possible cross-relaxation (spin diffusion) between protein protons may change the form of the relaxation decay, but the mean relaxation rate (initial slope) does not depend on this.



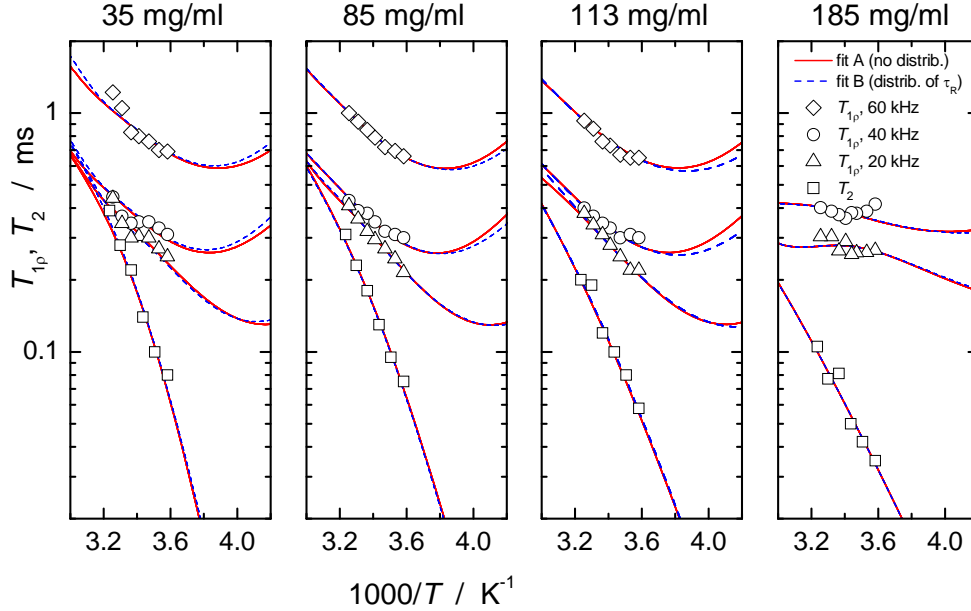
**Figure S7.** Hydrodynamic radii of  $\alpha$ B-crystallin as determined via the Stokes-Einstein relationship with independent measurements of the viscosity and translational self-diffusion at a concentration of 35 mg/ml. The solid and dashed lines indicate the mean value of all points and their standard deviation, respectively.

**Fitting the temperature dependences of the relaxation times: (i) distribution of correlation times.** Taking into account the intrinsic size distribution of  $\alpha$ -crystallin, as also reflected in the PFG NMR diffusion decays, it is worthy to estimate the impact of a distribution of  $\tau_{\text{rot}}$  on the fitting result of the  $T_2$  and  $T_{1\rho}$  data. To simulate a  $\tau_{\text{rot}}$  distribution, the fast component was represented by a spread of 3 modes on a logarithmic scale, namely  $2/3 \tau_{\text{rot}}$ ,  $\tau_{\text{rot}}$ , and  $3/2 \tau_{\text{rot}}$ , with relative amplitudes of  $1/4$ ,  $1/2$  and  $1/4$  for the faster, main and slower component, respectively. The total spectral density function addressing the fast component thus reads

$$J_{\text{fast}}(\omega, \tau_{\text{rot}}) = \frac{1}{4} J(\omega, \frac{2}{3} \tau_{\text{rot}}) + \frac{1}{2} J(\omega, \tau_{\text{rot}}) + \frac{1}{4} J(\omega, \frac{3}{2} \tau_{\text{rot}})$$

and reflects a relative standard deviation of  $\tau_{\text{rot}}$  of approx. 35%. As a consequence, this broadens the minimum of the  $T_{1\rho}$  curves, yet such a spread of  $\tau_{\text{rot}}$  values was not found to have an appreciable effect on the quality of the fitting result (cf. Fig. S8) and the fitting values obtained (cf. Table S1). In fact, the relative increase of the mean  $\tau_{\text{rot}}$  was found to be in good accordance with the previous results when modeling the experimental data using a single value of  $\tau_{\text{rot}}$  only. Thus, introducing a spread of  $\tau_{\text{rot}}$  is not reasonable since it does not change significantly the fitting

results, at the same time making fitting less certain by increasing the number of fitting parameters.



**Figure S8.**  $T_2$  and  $T_{1p}$  relaxation times (see legend) with their best fit results (lines) by use of a distribution of the fast component  $\tau_{rot}$  (dashed lines) in comparison to the fitting result using a single value of  $\tau_{rot}$  (solid lines).

**Table S1.** Summary of the fitting result by reflecting the fast component ( $\tau_{rot}$ ) by one mode only (model A) in comparison to the outcome assuming a logarithmic spread of  $\tau_{rot}$  with a standard deviation of about 35% (model B, see explanation in the text).  $\langle \tau_{rot} \rangle$  for the model (B) was calculated as  $\langle \tau_{rot} \rangle^{-1} = \frac{1}{4} \left( \frac{2}{3} \tau_{rot} \right)^{-1} + \frac{1}{2} (\tau_{rot})^{-1} + \frac{1}{4} \left( \frac{3}{2} \tau_{rot} \right)^{-1}$ .  $\langle R_{rot} \rangle$  was calculated according to Eq. 11 of the main paper.

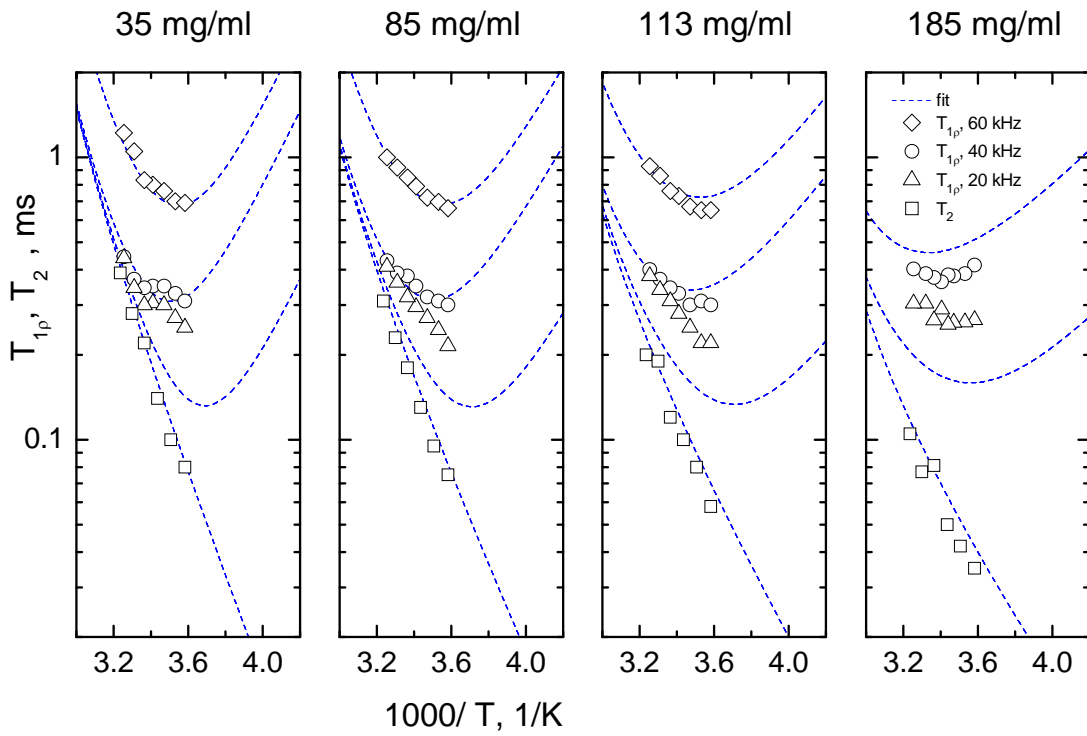
Model	$\tau$	35 mg/ml	85mg/ml	113 mg/ml	185mg/ml
(A)	$\tau_{rot} / \mu\text{s}$	$0.90 \pm 0.02$	$0.96 \pm 0.02$	$1.03 \pm 0.02$	$1.04 \pm 0.02$
	$\langle R_{rot} \rangle^{-1} / \mu\text{s}$	$0.91 \pm 0.05$	$0.98 \pm 0.05$	$1.05 \pm 0.05$	$1.31 \pm 0.05$
	$\langle R_{rot} \rangle (c_1) / \langle R_{rot} \rangle (c_i)$	<b>1.00</b>	<b>1.08 ± 0.02</b>	<b>1.15 ± 0.02</b>	<b>1.44 ± 0.04</b>
(B)	$\langle \tau_{rot} \rangle / \mu\text{s}$	$0.85 \pm 0.02$	$0.89 \pm 0.02$	$0.97 \pm 0.02$	$0.96 \pm 0.02$
	$\langle R_{rot} \rangle^{-1} / \mu\text{s}$	$0.86 \pm 0.05$	$0.91 \pm 0.05$	$0.99 \pm 0.05$	$1.17 \pm 0.05$
	$\langle R_{rot} \rangle (c_1) / \langle R_{rot} \rangle (c_i)$	<b>1.00</b>	<b>1.05 ± 0.02</b>	<b>1.16 ± 0.02</b>	<b>1.36 ± 0.04</b>

**Fitting the temperature dependences of the relaxation times: (ii)  $K_{HH}^{av}$  temperature dependence.** In the analysis we assume  $K_{HH}^{av}$  to be temperature independent although the amplitude of internal motions in proteins may depend on temperature (6,7). To check the influence of the possible  $K_{HH}^{av}$  temperature dependence on the fitting results, we performed the fitting assuming a simple linear dependence of  $K_{HH}^{av}$  on temperature. Direct measurements of  $K_{HH}^{av}$  in solid hydrated proteins at different temperatures (8,9) show that within temperature range of our experiments (from  $\sim 5$  °C to  $\sim 35$  °C)  $K_{HH}^{av}$  varies no more than 10-20%. Note that this temperature variation is caused not only by the change of motional amplitude, but mainly by the temperature dependence of the correlation times of internal motions which affects the motional averaging of the proton second moment. Therefore, we fitted the data assuming the 15% difference of  $K_{HH}^{av}$  between 5 and 35 °C. The fitting curves in this case look the same as in Fig. S8 or Fig. 3 of the main paper. The comparison of the fitting results (Table S2) with those assuming temperature independent  $K_{HH}^{av}$  (Table 1 of the main paper) demonstrates that the  $K_{HH}^{av}$  temperature dependence affects only the absolute value of  $K_{HH}^{av}$  and the activation energy  $E_{rot}$ . All other parameters remain the same.

**Table S2.** Dynamic parameter obtained from the fitting assuming  $K_{HH}^{av}$  to be temperature dependent.  $K_{HH}^{av}$  at 20 °C is  $(3.7 \pm 0.1) \cdot 10^9 \text{ s}^{-2}$ .

$c$ / mg/ml	$\tau_{rot}$ / $\mu\text{s}$ at 20 °C	$S_{rot}^2 \tau_S$ / $\mu\text{s}$ at 20 °C	$E_{rot}$ / kJ/mol	$E_s$ / kJ/mol
35	$0.92 \pm 0.02$	$0.65 \pm 0.02$	$12 \pm 1$	$63 \pm 2$
85	$0.97 \pm 0.02$	$0.84 \pm 0.02$	$13 \pm 1$	$49 \pm 2$
113	$1.04 \pm 0.03$	$1.38 \pm 0.03$	$12 \pm 1$	$38 \pm 2$
185	$1.07 \pm 0.03$	$3.73 \pm 0.07$	$2 \pm 1$	$27 \pm 1$

Here, we would like to also demonstrate that the assumption of the possibly steeper temperature dependence of  $K_{HH}^{av}$  cannot describe the data well without taking into account the slow component of the  $C_i(t)$ , i.e. assuming  $S_{rot}^2=0$ . We performed the data fitting assuming  $S_{rot}^2=0$  and the slope of the  $K_{HH}^{av}$  temperature dependence as an additional free fitting parameter. As a result, we obtained a bad fitting quality (see Fig. S9) and a rather unreasonable temperature variation of  $K_{HH}^{av}$  of about 40% between 5 and 35 °C. Thus, the assumption of the  $K_{HH}^{av}$  temperature dependence has no significant influence on our results.

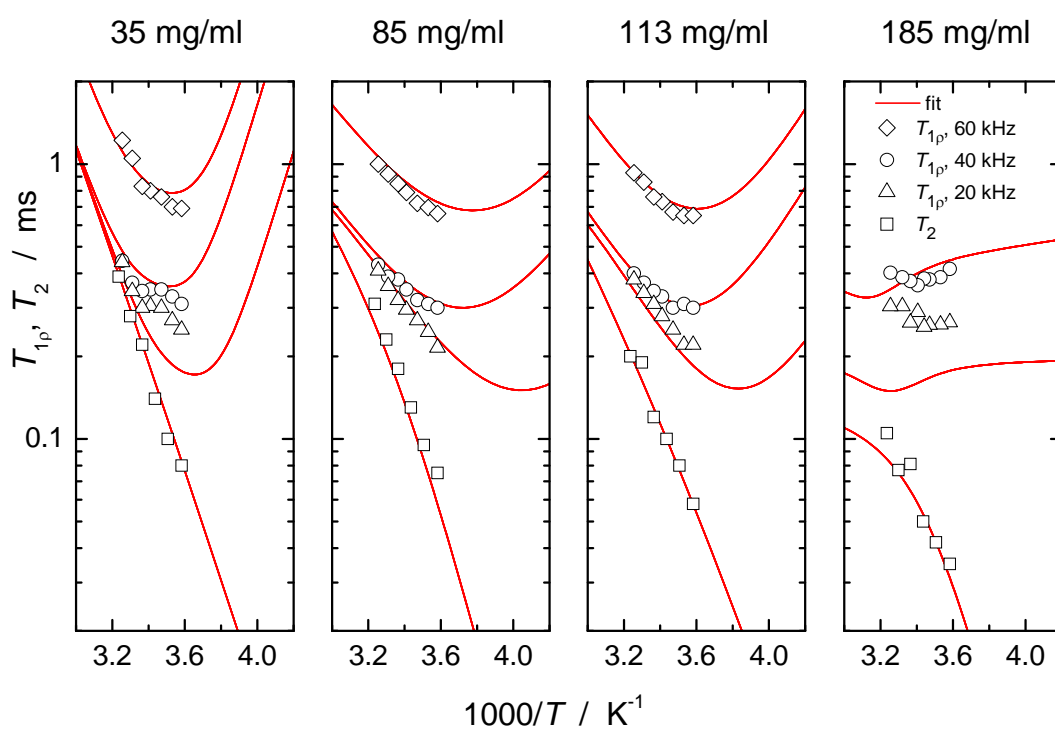


**Figure S9.** Fitting results assuming  $K_{HH}^{av}$  to be temperature-dependent, and  $S_{rot}^2=0$ .

**Fitting the temperature dependences of the relaxation times: (iii) fixed ratio between the correlation times.** Fig. S10 shows fits for which we tried a ratio between the correlation times fixed to

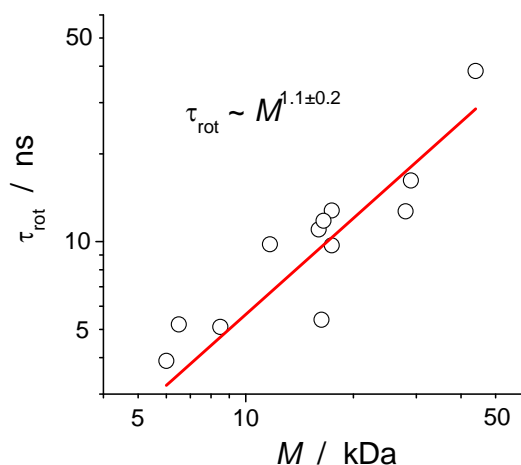
$$\tau_{\text{rot}}(35 \text{ mg/ml}) : \tau_{\text{rot}}(85 \text{ mg/ml}) : \tau_{\text{rot}}(113 \text{ mg/ml}) : \tau_{\text{rot}}(185 \text{ mg/ml}) = 1.00 : 1.67 : 2.20 : 6.00,$$

which corresponds to the relative increase in viscosity (or the translational diffusion slow-down). Obviously, this results in a strong mismatch between the fitting curves and the experimental data.



**Figure S10.**  $T_2$  and  $T_{1\rho}$  relaxation times with their best fit results (red lines) by use of a fixed ratio of the rotational correlation time  $\tau_{\text{rot}}$  following the retardation of translational diffusion.

### Rotational correlation time vs. molecular mass: statistics of literature data.



**Figure S11.** Rotational correlation time as a function of the protein  $M$  for 12 different proteins. All correlation times were recalculated for the temperature 20 °C, the details being presented in Table S3. It is clearly seen that the slope can be defined only with a large uncertainty.

**Table S3.** Collection of rotational correlation times for 12 different proteins. In all cases the correlation time was defined from the  $^{15}\text{N}$   $T_1/T_2$  ratio. Since the temperatures of the measurements were different, all the correlation times were recalculated to 20 °C using the Arrhenius dependence and an activation energy of 20 kJ/mol. Fig. S8 was plotted using the numbers in bold (2nd and 5th columns).

Protein	MM, kDa	$T$ / °C	$\tau_{\text{rot}}$ / ns	$\tau_{\text{rot}}$ (20°C)	Reference
Interleukin 1 $\beta$	<b>17.4</b>	36	8.3	<b>12.8</b>	Clore GM, Driscoll PC, Wingfield PT, Gronenborn AM (1990), <i>Biochemistry</i> 29: 7387–7401.
calbindin D9k	<b>8.5</b>	27	4.2	<b>5.1</b>	Kordel J, Skelton NJ, Akke M, Palmer AG, Chazin WJ (1992), <i>Biochemistry</i> 31: 4856–4866.
Bacillus-subtilis glucose permease-IIA domain	<b>17.4</b>	35	6.5	<b>9.7</b>	Stone MJ et al. (1992), <i>Biochemistry</i> 31: 4394–4406.
Thioredoxin	<b>11.7</b>	35	6.55	<b>9.8</b>	Stone MJ, Chandrasekhar K, Holmgren A, Wright PE, Dyson HJ (1993), <i>Biochemistry</i> 32: 426–435.
Interleukin-8	<b>16</b>	27	9.1	<b>11.0</b>	Grasberger BL, Gronenborn AM, Clore GM (1993), <i>J Mol Biol</i> 230: 364–372.
Igg binding domain	<b>6</b>	26	3.3	<b>3.9</b>	Barchi JJ, Jr., Grasberger B, Gronenborn AM, Clore GM (1994), <i>Prot Sci</i> 3: 15–21.
Ribonuclease HI	<b>16.5</b>	27	9.7	<b>11.8</b>	Mandel AM, Akke M, Palmer AG (1995), <i>J. Mol Biol</i> 246: 144–163.
Savinase	<b>28</b>	30	9.7	<b>12.7</b>	Remerowski ML, Pepermans HAM, Hilbers CW, Van De Ven FJM (1996), <i>Eur J Biochem</i> 235: 629–640.

SH3 domain	<b>6.5</b>	7	7.6	<b>5.2</b>	Chevelkov V, Zhuravleva AV, Xue Y, Reif B, Skrynnikov NR (2007), <i>J Am Chem Soc</i> 129: 12594–12595.
Flavodoxin	<b>16.3</b>	27	4.5	<b>5.4</b>	Hrovat A, Blümel M, Löhr F, Mayhew SG, Rüterjans H (1997), <i>J Biomol NMR</i> 10: 53–62.
Ectodomain of SIV gp41	<b>44</b>	45	20	<b>38.5</b>	Caffrey M, Kaufman J, Stahl SJ, Wingfield PT, Gronenborn AM, Clore GM (1998), <i>J Magn Reson</i> 135: 368–372.
TEM-1 $\beta$ -lactamase	<b>29</b>	30	12.4	<b>16.2</b>	Savard PY, Gagne SM (2006), <i>Biochemistry</i> 45: 11414-11424.

## References.

- (1) Mainz, A., S. Jehle, B.J. van Rossum, H. Oschkinat, and B. Reif (2009). Large protein complexes with extreme rotational correlation times investigated in solution by magic-angle-spinning NMR spectroscopy. *J Am Chem Soc* 131: 15968-15969. (Supporting Information)
- (2) Cotts R.M., M.J.R. Hoch, T. Sun, and J.R. Markert (1989). Pulsed Field Gradient Stimulated Echo Methods for Improved NMR Diffusion Measurements in Heterogeneous Systems. *J Magn Reson* 83: 252-266.
- (3) Chen A., C.S. Johnson Jr., M. Lin, and M.J. Shapiro (1998). Chemical Exchange in Diffusion NMR Experiments. *J Am Chem Soc* 120: 9094-9095.
- (4) Andrec M, and J.H. Prestegard (1997). Quantitation of chemical exchange rates using pulsed-field-gradient diffusion measurements. *J Biomol NMR* 9: 136-150.
- (5) Krushelnitsky, A.G., and V.D. Fedotov (1993). Overall and internal protein dynamics in solution studied by the nonselective proton relaxation. *J Biomol Struct Dyn* 11:121-141.
- (6) Chang, S.-L., and N. Tjandra (2005). Temperature dependence of protein backbone motion from carbonyl  $^{13}\text{C}$  and amide  $^{15}\text{N}$  NMR relaxation. *J Magn Reson* 174: 43-53.
- (7) Johnson, E, A.G. Palmer, and M. Rance (2007). Temperature dependence of the NMR generalized order parameter. *Prot Str Func Bioinf* 66: 796-803.
- (8) Blears, D.J and S.S.Danyluk (1968). Proton wide-line nuclear magnetic resonance spectra of hydrated proteins. *Biochim Biophys Acta* 154: 17-27.
- (9) Krushelnitsky, A.G., V.D. Fedotov, J. Spevacek, and J. Straka (1996). Dynamic structure of proteins in solid state.  $^1\text{H}$  and  $^{13}\text{C}$  NMR relaxation study. *J Biomol Struct Dyn* 14: 211-224.



## Electronic Supporting Material

### The "long tail" of the protein tumbling correlation function: observation by $^1\text{H}$ NMR relaxometry in a wide frequency and concentration range.

Matthias Roos<sup>1</sup>, Marius Hofmann<sup>2</sup>, Susanne Link<sup>1</sup>, Maria Ott<sup>1</sup>, Jochen Balbach<sup>1</sup>, Ernst Rössler<sup>2</sup>, Kay Saalwächter,<sup>1\*</sup> Alexey Krushelnitsky<sup>1\*</sup>

<sup>1</sup> Martin-Luther-Universität Halle-Wittenberg, Germany;

<sup>2</sup> University of Bayreuth, Germany

#### *Rescaling of $R_{1\rho}$ data for visual comparison with $R_1$ data*

For visual comparison only,  $R_{1\rho}$  was rescaled by a factor of 10/3 to be in accordance with the absolute values of the  $R_1$  data. This scaling factor directly follows from the approximation

$J(\omega_0) \approx J(2\omega_0)$ , providing

$$R_1 \approx \frac{10}{3} K_{\text{HH}} J(2\omega_0) \quad (\text{S1})$$

$$R_{1\rho} \approx \frac{K_{\text{HH}}}{3} [3J(2\omega_{\text{SL}}) + 7J(2\omega_0)] \quad (\text{S2})$$

$$R_2 \approx \frac{K_{\text{HH}}}{3} [3J(0) + 7J(2\omega_0)] . \quad (\text{S3})$$

Making use of  $J(2\omega_{\text{SL}}) \gg J(2\omega_0)$  in the case of the  $R_{1\rho}$  measurements, in which  $\omega_{\text{SL}}/2\pi \leq 40\text{kHz}$  and  $\omega_0/2\pi = 400\text{MHz}$ , it follows that

$$R_1 \approx \frac{10}{3} K_{\text{HH}} J(2\omega_0) \quad (\text{S4})$$

$$\frac{10}{3} R_{1\rho} \approx \frac{10}{3} K_{\text{HH}} J(2\omega_{\text{SL}}) . \quad (\text{S5})$$

Thus, the two quantities that map out  $J(\omega)$  can be compared visually in the same plot.

In contrast, our  $R_2$  measurements were performed at a Larmor frequency of 20 MHz, which does not allow for safely neglecting the contribution of  $J(2\omega_0)$  to  $R_2$ . A correct rescaling of

$R_2$  thus requires knowledge on both  $J(\omega_0)$  and  $J(2\omega_0)$ , which is impossible without further data analysis. Of course, data fitting was performed using the exact equations for  $R_1$ ,  $R_{1\rho}$  and  $R_2$ .

### ***Reliability of the FC-NMR $R_1$ measurements at low fields***

The shortest relaxation delay in our field-cycling measurements was 0.7 ms ("dead time"). In the presence of a distribution of  $R_1$ s, the initial slope of the field-cycling relaxation decay may thus not reflect the true initial slope at  $t=0$ , as illustrated by Fig. S1.

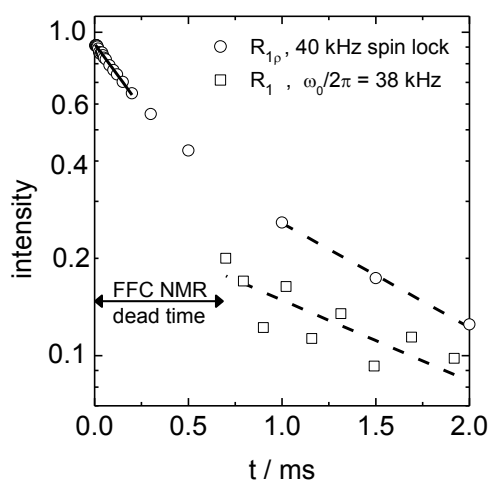


Figure S1. BSA ( $c=213$  mg/ml,  $T=20$  °C)  $R_{1\rho}$  and  $R_1$  relaxation decays after the water component subtraction. As seen from the  $R_{1\rho}$  decay, the experimentally observed relaxation rate (solid and dashed line) may significantly change during the dead time of the FC NMR measurements. This issue is of relevance as soon as the relaxation time of the fast relaxing protein component becomes comparable to the dead time.

The relaxation decays were only considered as being reliable if a bi-exponential fit to the data provided - within the fitting uncertainty - the same initial slope as the fitting result based upon a log-normal distribution, see Fig. S2. In the case of bi-exponential analysis, the relaxation decay was fitted according to the equation

$$A(t) = P_a \exp(-R_a t) + P_b \exp(-R_b t) \quad , \quad (S6)$$

where  $P_{a,b}$  and  $R_{a,b}$  are the intensities and relaxation rates of the components  $a$  and  $b$ . Then, the mean relaxation rate was defined as

$$\langle R \rangle = \frac{P_a R_a + P_b R_b}{P_a + P_b} \quad . \quad (S7)$$

If the two evaluation methods gave different results, the distribution of  $R_1$ s within the undetected part of the decay (initial 0.7 ms), and thus the estimate of  $\langle R_1 \rangle$ , was considered uncertain. This situation occurs in the case of BSA at low relaxation fields, potentially due to the presence of oligomers, as these cause increased relaxation rates and an increased curvature of the relaxation decays.

For the sake of clarity in notation, the brackets of  $\langle R_1 \rangle$  will henceforth be skipped, as is also the case in the main paper.

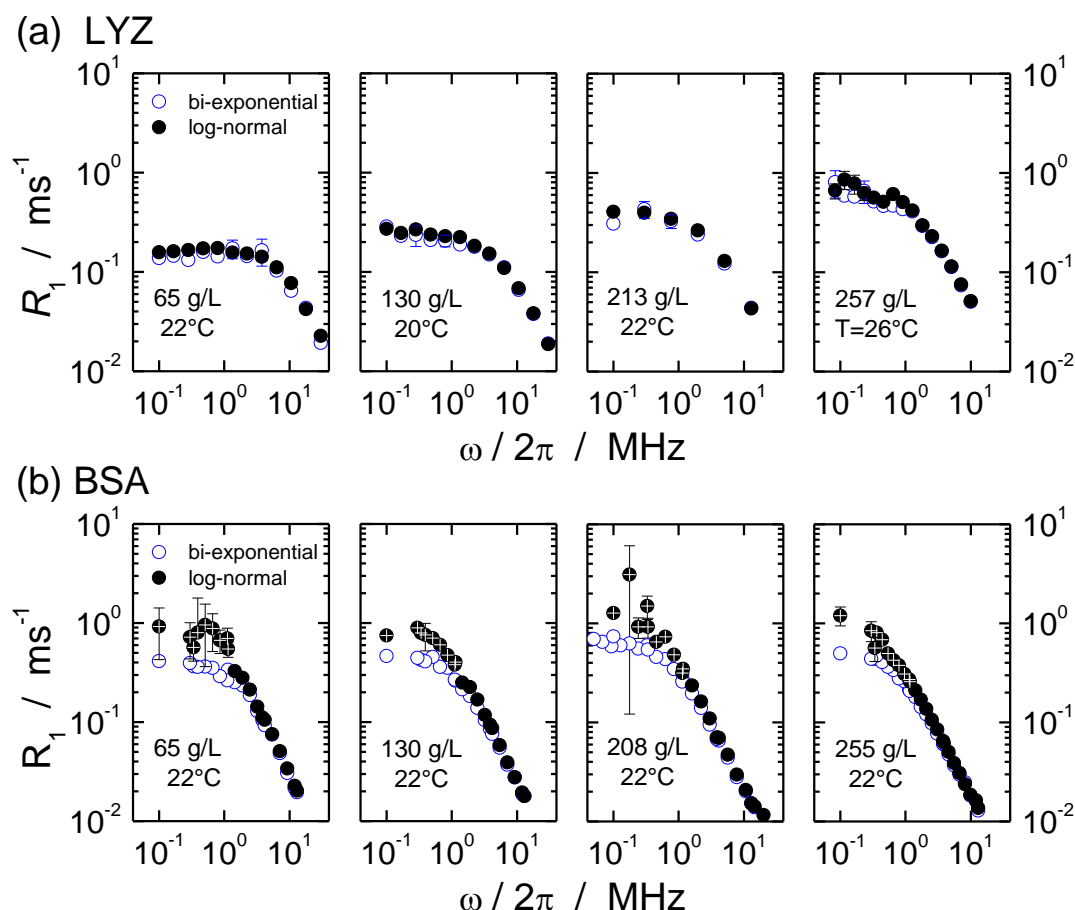


Figure S2. Relaxation rates of (a) LYZ and (b) BSA as provided by bi-exponential (open symbols) and log-normal-distribution fitting to the raw data. If not indicated by error bars, the fitting error was of the same order of magnitude as or smaller than the symbol size. Data points marked with a cross were not used for further data analysis.

An additional argument supporting the accuracy of the FC-NMR measurements at low field in LYZ solution is provided by the comparison of  $R_1$  with  $R_{1\rho}$  measurements, as the latter provide a safe, dead-time free estimate of the relaxation rate at low frequencies in absence of any relevant delays during the pulse sequence. At 65 mg/ml and 130 mg/ml protein concentration, the FC-NMR measurements provide a plateau value of  $R_1$  that at all temperatures matches the corresponding value of  $10/3 R_{1\rho}$  (see Fig. S3). The same holds for the samples of higher protein concentrations (213 mg/ml, 257 mg/ml) at temperatures above 15°C. At lower temperatures of the highly concentrated samples, however, no plateau of  $R_1$  is

observed, which we address to the increased relevance of the slow component of rotational diffusion. However, to be safe with respect to the influence of potentially biased data points in highly concentrated LYZ solutions, we excluded FC-NMR values of a mean relaxation rate above  $0.5 \text{ ms}^{-1}$  (at and below this value,  $R_1$  reaches a plateau, in agreement with  $R_{1\rho}$ ) from data fitting, and observed very similar results compared to the fit of all data without exceptions, see Table S1.

## Experimental and fitting results

### (i) LYZ

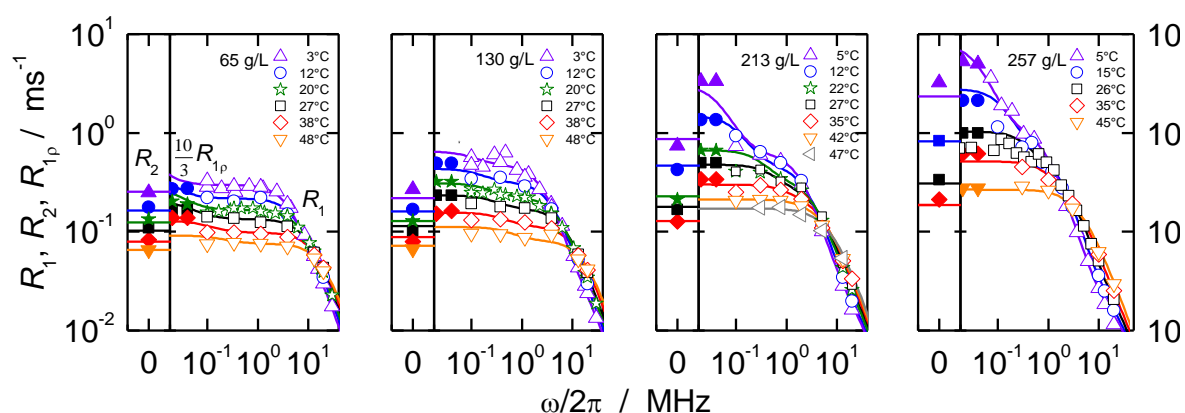


Figure S3. Dispersion profiles of LYZ at several temperatures. The shown data points include the full FC- $R_1$  data set, a part of it being shown in Fig. 5 of the main paper.  $10/3R_{1\rho}$  data points are shown as solid symbols. Additionally, we here include data points for  $R_1$  measured with a saturation recovery sequence on a Bruker minispec mq20 at the resonance frequency 20 MHz. In order not to overload the figure, error bars were skipped, the measuring uncertainty of the relaxation rates is comparable to the symbol size. A part of these data plotted as a function of temperature is shown in Fig. S5a.

Table S1. Fit result of LYZ. The values in row A are the best-fit results including all data points measured. For row B potentially biased data points ( $\langle R_1 \rangle > 0.5 \text{ ms}^{-1}$ ) were excluded (see explanation above).

	$c$ g/L	Brownian rotation		“Long tail” of rot. diffusion			Internal dynamics			$K_{\text{HH}}^{(\text{av})}$ $10^9 \text{ s}^2$
		$\tau_{\text{rot}}$ ns	$E_{\text{rot}}$ kJ/mol	$S_{\text{rot}}^2$	$\tau_{\text{S}}$ ns	$E_{\text{S}}$ kJ/mol	$\tau_{\text{int}}$ ns	$S_{\text{int}}^2$	$E_{\text{int}}$ kJ/mol	
A	65	<b><math>10.5 \pm 0.2</math></b>	<b><math>25 \pm 1</math></b>	$\leq 0.003$	$\geq 3000$	<b><math>54 \pm 2</math></b>	<b><math>2.3 \pm 0.1</math></b>	<b><math>0.72 \pm 0.01</math></b>	<b><math>6 \pm 3</math></b>	<b><math>6.1 \pm 0.1</math></b>
	130	<b><math>14.6 \pm 0.2</math></b>	<b><math>31 \pm 1</math></b>	<b><math>0.014 \pm 0.002</math></b>	<b><math>410 \pm 40</math></b>	<b><math>29 \pm 1</math></b>				
	213	<b><math>23.9 \pm 0.3</math></b>	<b><math>27 \pm 1</math></b>	<b><math>0.09 \pm 0.01</math></b>	<b><math>330 \pm 10</math></b>	<b><math>73 \pm 1</math></b>				
	257	<b><math>39.9 \pm 0.6</math></b>	<b><math>35 \pm 2</math></b>	<b><math>0.23 \pm 0.01</math></b>	<b><math>380 \pm 10</math></b>	<b><math>76 \pm 1</math></b>				
B	65	$10.5 \pm 0.1$	$24 \pm 1$	$\leq 0.003$	$\geq 3000$	$53 \pm 2$	$2.2 \pm 0.1$	$0.73 \pm 0.01$	$9 \pm 1$	$6.1 \pm 0.1$
	130	$13.9 \pm 0.2$	$29 \pm 1$	$0.016 \pm 0.001$	$390 \pm 30$	$33 \pm 1$				
	213	$23.3 \pm 0.4$	$30 \pm 1$	$0.13 \pm 0.01$	$240 \pm 10$	$65 \pm 1$				
	257	$38.1 \pm 0.7$	$36 \pm 1$	$0.23 \pm 0.01$	$410 \pm 10$	$73 \pm 1$				

(ii) BSA

Commercially available BSA contains a significant fraction of oligomers in solution. If these are neglected upon fitting the data, the procedure will give a mean rotational correlation time that incorporates an uncertain weighting of monomers vs. oligomers. However, since  $\tau_S \gg \tau_{rot}$  a spread of the rotational correlation does not challenge the sensitivity of the fit to the slow component:  $R_2$  and  $R_{1\rho}$  values are most sensitive to the latter, whereas relaxation at higher Larmor frequencies serves to fix  $\tau_{rot}$ .

Taking into account the oligomer fraction increases the number of fitting parameters, but provides more reasonable results. To reduce the complexity of the fitting model (otherwise, the fitting is too uncertain and ambiguous), we assumed the oligomer fraction  $P_M$  to stay the same over the entire temperature range ( $4^\circ\text{C} \leq T \leq 30^\circ\text{C}$ ), and fixed the rotational activation energy of oligomers and monomers at all concentrations to 20 kJ/mol, that is, the activation energy of water flow that governs the protein rotational motion (in the case of LYZ and  $\alpha$ B-crystallin (Roos et al. 2015) we obtained similar values of  $E_{rot}$ ). We divided the rotational correlation time into two components,  $\tau_{rot}$  and  $\tau_{rot2}$ , that correspond to the monomer and oligomer fractions, respectively. In practice, the second term of  $J(\omega)$ , see Eq. (7) of the main article, was replaced by

$$S_{int}^2 (1 - S_{rot}^2) \left[ \frac{P_M \tau_{rot}}{1 + (\omega \tau_{rot})^2} + \frac{(1 - P_M) \tau_{rot2}}{1 + (\omega \tau_{rot2})^2} \right]. \quad (\text{S8})$$

Figs. S4 and S5 displays the corresponding fitting results. According to the data analysis, oligomers comprise more than half of the protein mass. This outcome is in accordance with our size-exclusion chromatography results obtained for dilute BSA solutions (cf. Materials and Methods), and further agrees well with the findings of ref. (Atmeh et al. 2007). Table S2 presents the fitting results for the two options of the data analysis (single  $\tau_{rot}$  and  $\tau_{rot}/\tau_{rot2}$  components of the overall tumbling correlation function).

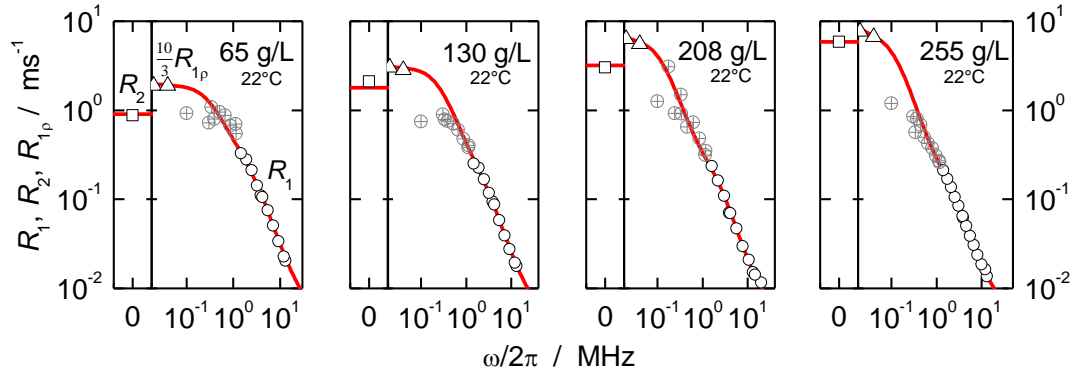


Figure S4. Dispersion profiles of the BSA relaxation data (the same as shown in Fig. 5b of the main paper). Fitting curves (solid red lines) correspond to the analysis that takes into account oligomers (Eq. S8). Crossed symbols are uncertain data points that were not included in data fitting (see explanation in the text). Other BSA data are shown as function of temperature in Fig. S5b.

Table S2. Fitting results of BSA, excluding uncertain data points. Row A is from fits excluding oligomers, in which case  $\tau_{\text{rot}}$  represents an average over monomers and oligomers. Row B is the best fit result including oligomers, where “M” and “O” denote monomers and oligomers, respectively. The oligomer fractions  $P_M$  (see Eq. S8) are  $0.45 \pm 0.05$ ,  $0.60 \pm 0.05$ ,  $0.60 \pm 0.05$  and  $0.70 \pm 0.08$  for the concentrations 65, 130, 208 and 255 g/L, respectively.

	$c$ g/L	<u>Brownian rotation</u>		<u>“Long tail” of rot. diffusion</u>			<u>Internal dynamics</u>			$K_{\text{HH}}^{(\text{av})}$ $10^9 \text{ s}^2$
		$\tau_{\text{rot}}$ ns	$E_{\text{rot}}$ kJ/mol	$S_{\text{rot}}^2$	$\tau_S$ $\mu\text{s}$	$E_S$ kJ/mol	$\tau_{\text{int}}$ ns	$S_{\text{int}}^2$	$E_{\text{int}}$ kJ/ mol	
A	65	$107 \pm 1$	$17 \pm 1$	$\leq 0.002$	$\geq 20$	$40 \pm 3$	$2.0 \pm 0.1$	$0.79 \pm 0.01$	$0.2 \pm 0.1$	$7.8 \pm 0.1$
	130	$158 \pm 1$	$18 \pm 1$	$0.006 \pm 0.001$	$23 \pm 4$	$39 \pm 2$				
	208	$203 \pm 3$	$23 \pm 1$	$0.10 \pm 0.01$	$3 \pm 1$	$45 \pm 1$				
	255	$263 \pm 4$	$19 \pm 1$	$0.12 \pm 0.01$	$6 \pm 1$	$45 \pm 2$				
B	65	M: $36 \pm 2$ O: $260 \pm 20$	$20$ (*)	$\leq 0.002$	$\geq 30$	$29 \pm 5$	$1.4 \pm 0.1$	$0.74 \pm 0.01$	$5 \pm 4$	$6.0 \pm 0.1$
	130	M: $35 \pm 2$ O: $330 \pm 20$	$20$ (*)	$0.006 \pm 0.002$	$29 \pm 7$	$37 \pm 2$				
	208	M: $45 \pm 2$ O: $730 \pm 30$	$20$ (*)	$0.02 \pm 0.01$	$15 \pm 2$	$62 \pm 2$				
	255	M: $42 \pm 2$ O: $760 \pm 23$	$20$ (*)	$0.05 \pm 0.01$	$15 \pm 2$	$52 \pm 2$				

(\*) To stabilize these fits, the activation energy of  $\tau_{\text{rot}}$  was fixed to 20 kJ/mol at all concentrations.

(iii) Relaxation rates as a function of temperature

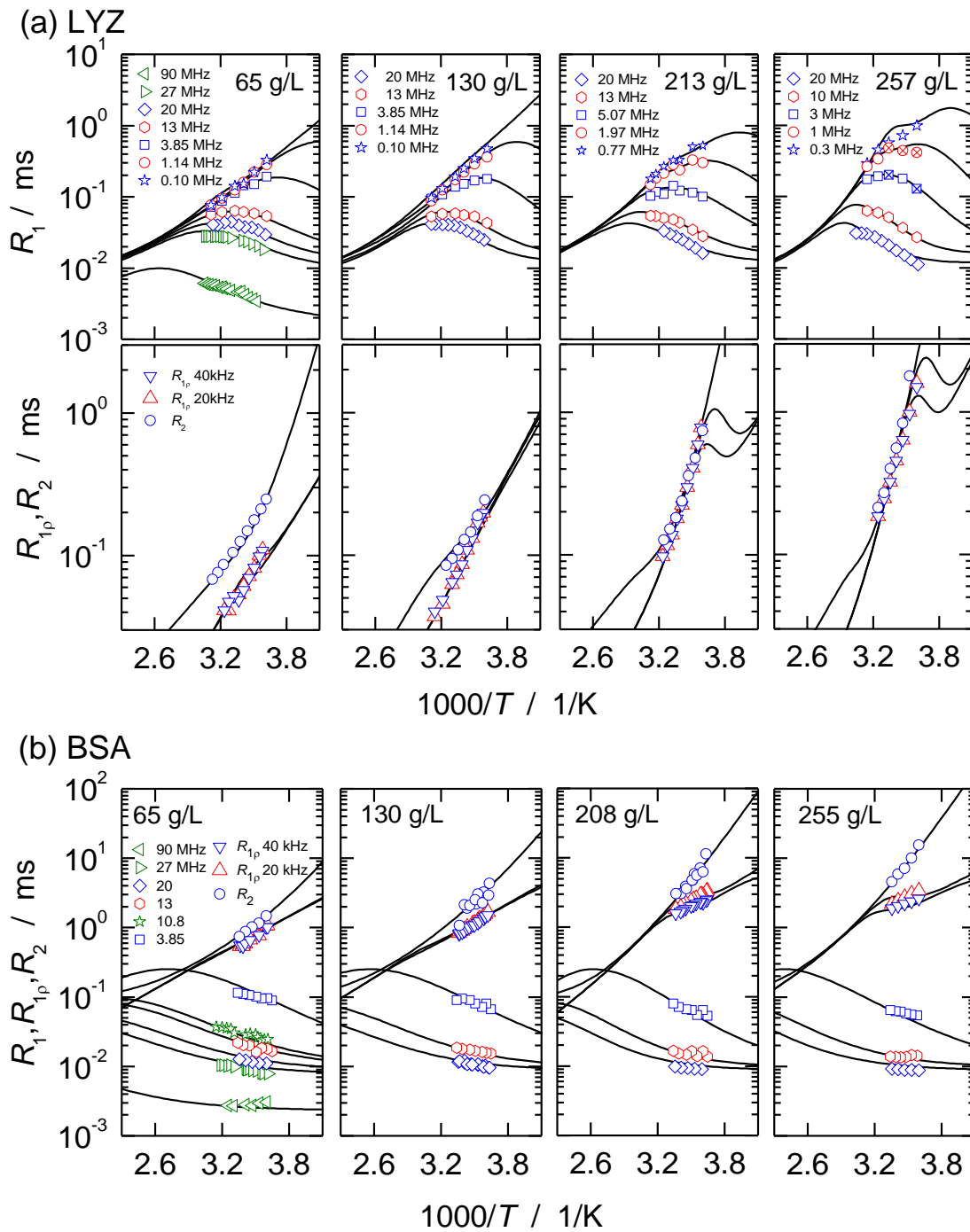


Figure S5. Temperature dependencies of  $R_1$ ,  $R_{1p}$  and  $R_2$  relaxation times together with the best fit results (solid lines).  $R_1$  values at 90, 27 and 10.8 MHz were taken from ref. (Krushelnisky 2006). Two symbols marked with a cross (LYZ 257 g/L) were interpolated from measurements at different frequencies but at the same temperature, and are only shown for visual presentation, but were not included in the fitting.

### Fitting LYZ data assuming no "long tail" component

The necessity of inclusion of the "long tail" in the fitting model was justified in our previous papers, see the Introduction. Here, by the example of LYZ data, we demonstrate once more that neglecting the "long tail" provides a systematic inconsistency between experimental points and fitting curves. Fig. S6 presents the LYZ relaxation data in frequency and temperature domains along with the fitting curves assuming  $S_{rot}^2=0$ .

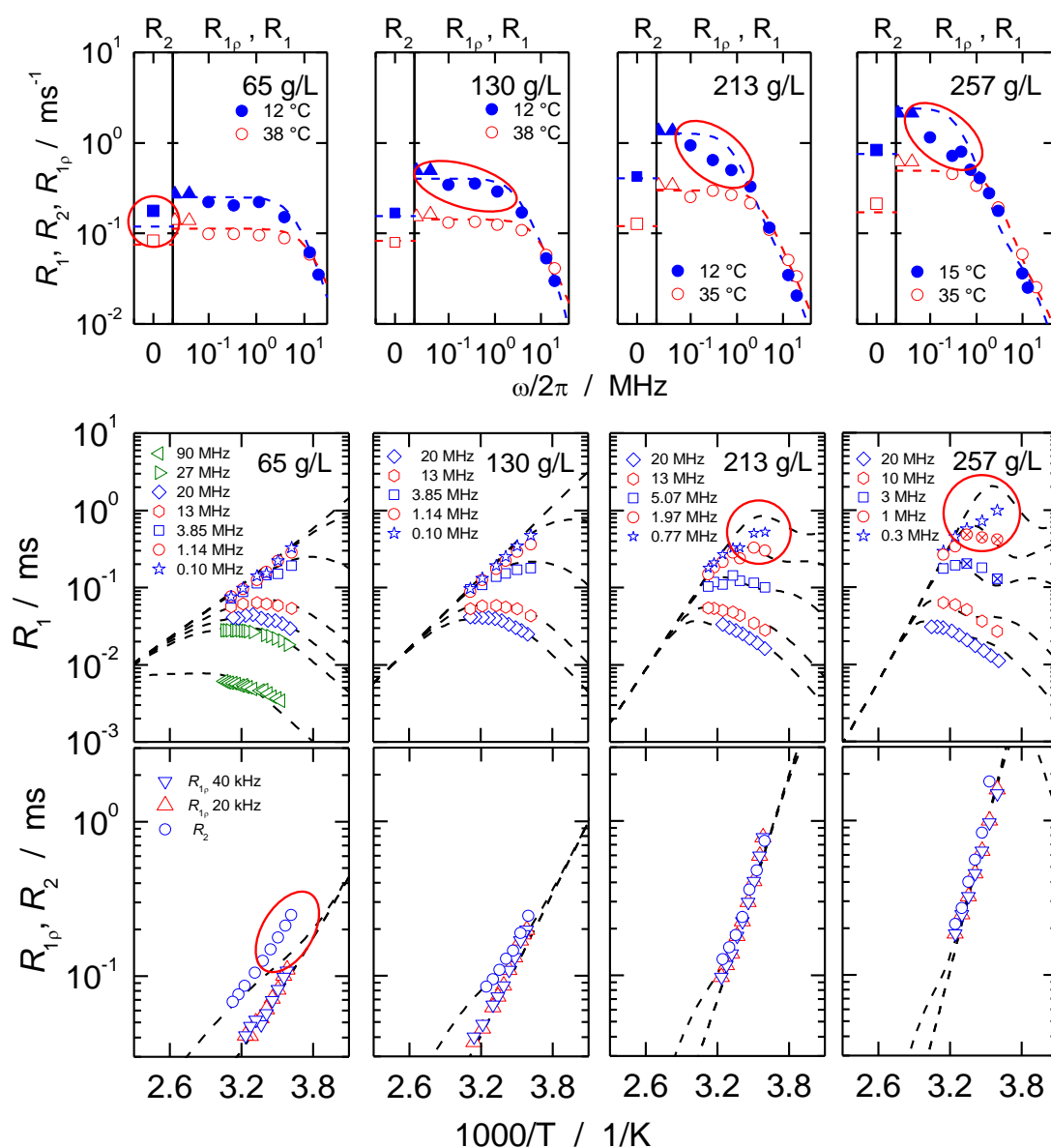


Figure S6. Experimental data and the fitting curves for LYZ solutions. Fitting was performed assuming  $S_{rot}^2=0$ . The regions of systematic inconsistency between the experimental and fitting relaxation rates values are indicated by red circles/ovals.

As seen in that graph, at low concentration  $R_2$  values cannot be reasonably fitted. The correlation time  $\tau_S$  at this concentration is large, and thus it affects only  $R_2$ ; all other



relaxation data can be described by the simplest model quite well. Upon increasing concentration,  $\tau_s$  becomes shorter (see the discussion of this phenomenon in the main paper), and thus the region of the fitting inconsistency shifts towards FC  $R_1$ 's in the frequency range 0.5 - 1 MHz. The higher the concentration, the larger the order parameter  $S_{\text{rot}}^2$  and, hence, the larger the fitting discrepancy, which is seen in Fig. S6. Finally, we note that the root mean square deviation of the fitting (see Eq. 11 of the main paper) with and without the "long tail" is 0.12 and 0.2, respectively.

### *Simulation of high-field $^{15}\text{N}$ NMR relaxation data.*

The  $^{15}\text{N}$  relaxation parameters were simulated according to the standard formulae for the heteronuclear dipole-dipole and CSA relaxation mechanisms with a typical values for the N-H pair in the protein backbone (see e.g. Daragan and Mayo 1997). Below we show the simulated  $R_1$ ,  $R_2$  and NOE's for three frequencies as a function of the product  $S_{\text{rot}}^2 \tau_s$  for the "rigid" and "mobile" residues:

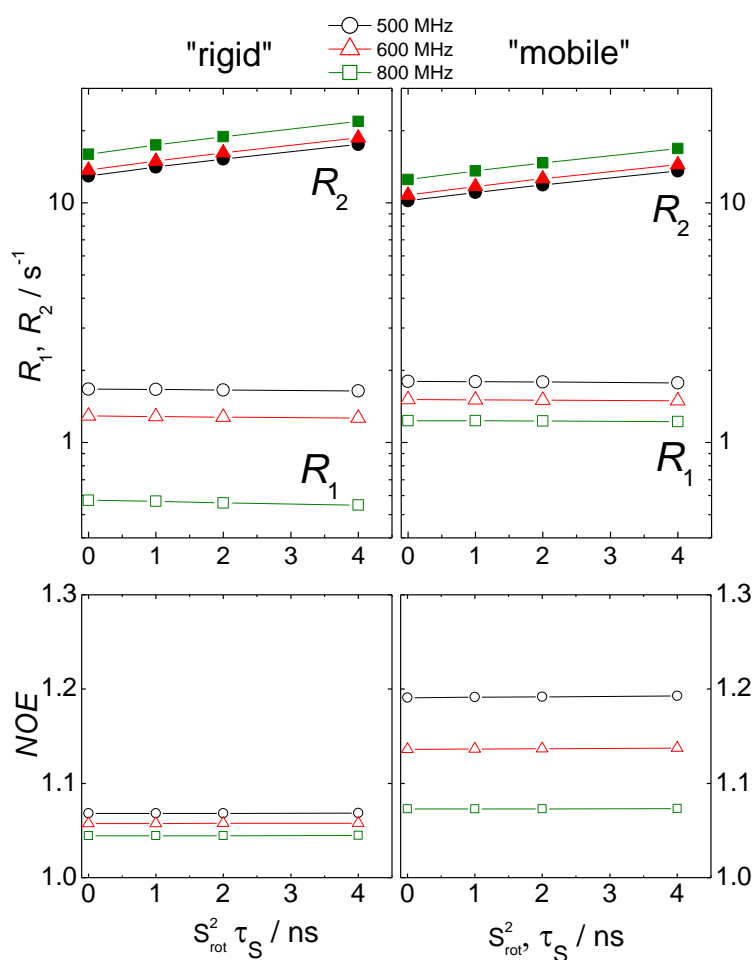


Figure S7. Simulated  $R_1$ ,  $R_2$  and NOE's as a function of the product  $S_{\text{rot}}^2 \tau_s$  for "rigid" and "mobile" residues.

## ***References***

Atmeh RF, Arafa IM, Al-Khateeb M (2007) Albumin aggregates: hydrodynamic shape and physico-chemical properties. *Jordan J Chem* 2:169-182.

Daragan VA, Mayo KH (1997) Motional model analyses of protein and peptide dynamics using C-13 and N-15 NMR relaxation. *Prog NMR Spect* 31:63-105.

Krushelnitsky A (2006) Intermolecular electrostatic interactions and Brownian tumbling in protein solutions. *Phys Chem Chem Phys* 8:2117–2128.

Roos M, Link S, Balbach J, Krushelnitsky A, Saalwächter K (2015) NMR-detected Brownian dynamics of  $\alpha$ B-crystallin over a wide range of concentrations. *Biophys J* 108:98-106.

## Supporting Information

for

### Coupling and Decoupling of Rotational and Translational Diffusion of Proteins under Crowding Conditions

Matthias Roos,<sup>1,§</sup> Maria Ott,<sup>1,§</sup> Marius Hofmann,<sup>2</sup> Susanne Link,<sup>1</sup> Ernst Rössler,<sup>2</sup>

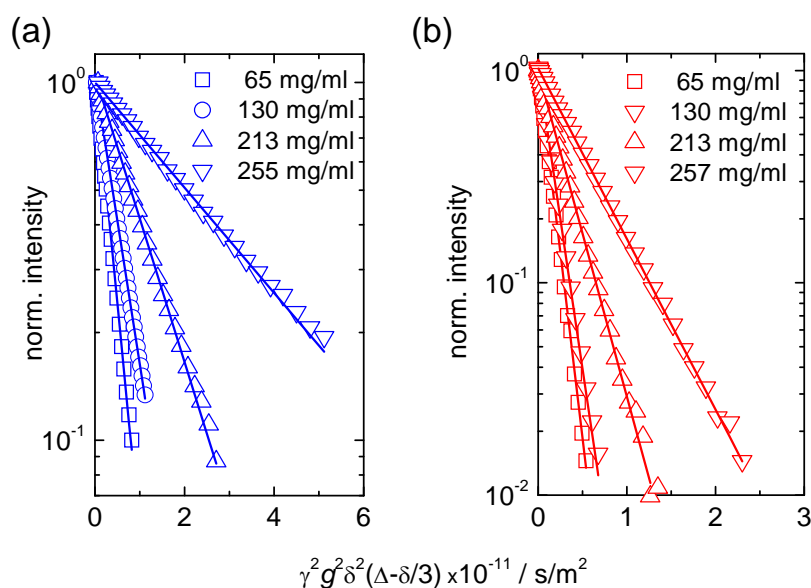
Jochen Balbach,<sup>1</sup> Alexey Krushelnitsky,<sup>1,\*</sup> Kay Saalwächter<sup>1,\*</sup>

<sup>1</sup> Martin-Luther-Universität Halle-Wittenberg, Institut für Physik, 06120 Halle (Saale), Germany;

<sup>2</sup> Universität Bayreuth, Experimentalphysik II, 95440 Bayreuth, Germany.

#### SI1: PFG NMR and sub-ensemble effects

Translational diffusion coefficients were measured via the pulsed-gradient stimulated echo (PGSTE) technique, using bipolar gradient pulses [1]. This means that for spatial encoding as well as decoding, positive and negative gradient pulses of duration  $\delta/2$  are applied during the first and second half, respectively, of a Hahn echo of total duration  $\tau$ . Thus, during both encoding/decoding periods, the NMR signal is reduced by transverse relaxation as described by  $I(2\tau) \propto \exp[-2\tau/T_2]$ . This may lead to sub-ensemble selection in a polydisperse system. Our data for BSA and HEWL, see Fig. S1, are well represented by exponential single-component decay functions. Since commercial BSA is known to contain oligomers [2], this deserves some further comments.

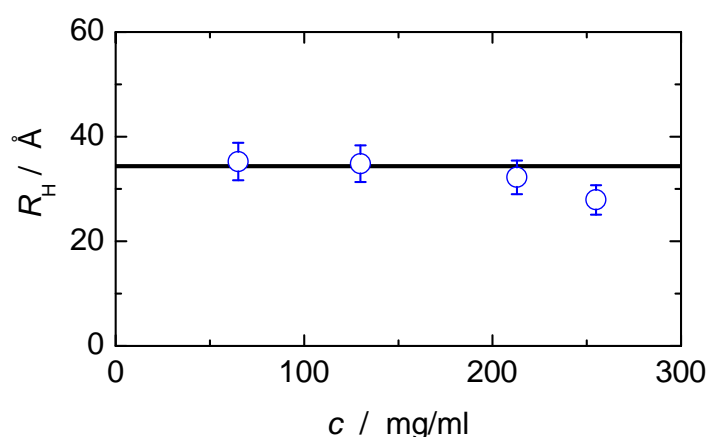


**Figure S1:** PFG NMR diffusion decays for (a) BSA and (b) HEWL at 20°C. Note the singly exponential nature of the decays, reflecting a single diffusion coefficient.

<sup>§</sup> equal contributions

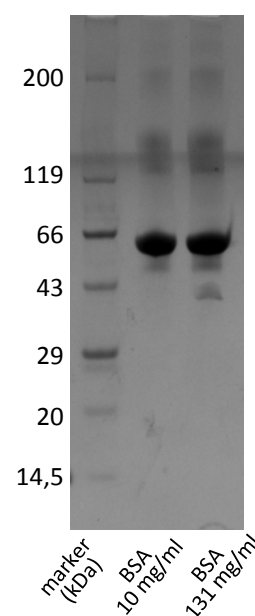
\* corresponding authors

**Figure S2:** Hydrodynamic radii of BSA as estimated from the macro-viscosity and the long-time translational diffusion coefficients based upon the GSE relation, eq. (1) of the main paper. Note that the strong concentration dependencies of the measured data at high concentration impart larger errors for slight concentration mismatches. The horizontal line marks the value expected for BSA monomers (34.4 Å).



Hydrodynamic radii for BSA obtained from eq. (1) of the main paper (see Fig. S2) match at all concentrations the size of the monomer, demonstrating that there seems to be no significant signal contribution from the oligomers. Yet, as shown previously by size-exclusion chromatography and blue native (BN) PAGE on our BSA sample [2], BSA contains a significant portion of oligomers (dimers up to hexamers) [3,4]. These oligomers even remained on an SDS-PAGE (Fig. S3), i.e., under denaturing conditions. This indicates a high stability of the oligomers once the oligomers are formed. Indeed, the dimerization of BSA likely happens due to covalent binding among free sulfhydryl groups of cysteine [5], providing long-time stable oligomers.

To explain this apparent discrepancy, we note that the inverse of the transverse relaxation time scales with rotational diffusion roughly as  $1/T_2 \propto (1 - a_s) \tau_r + a_s \tau_s$ , where  $\tau_r$  is the tumbling time, and  $a_s$  and  $\tau_s$  characterize the slow component of rotational tumbling, see ref. [2]. Consequently, the slower the rotational dynamics of the particle, the faster the NMR signal decays during the encoding/decoding periods. This effect is of particular relevance if the transverse relaxation time  $T_2$  is shorter than the duration of the gradient encoding periods. This is obviously the case for the oligomers in our BSA solutions. Consequently, our translational diffusion data for BSA reflect monomer diffusion within a polydisperse surrounding.



**Figure S3:** SDS-PAGE of BSA, reflecting monomers, dimers, and higher order oligomers. BSA solutions were diluted down to about 1 mg/ml directly before application of the SDS-PAGE.

For HEWL, there is no further consequence, as it is purely monomeric [2]. The  $\alpha$ Bc oligomers, on the other hand, contain at their C-terminus relatively long trains of unstructured residues with high mobility, competing the effect of slow overall rotation of the protein assembly. The finally detected NMR diffusion signal mainly results from these unstructured residues independently of the actual rotational correlation time of the overall assemblies. Thus, in this case, PFG NMR provides the full unbiased ensemble average of the  $\alpha$ Bc assemblies [6].

## SI2: Polarized fluorescence correlation spectroscopy (FCS)

### *Fluorescence microscope set up*

FCS experiments were conducted using a home-built confocal microscope. The excitation light was provided by a fiber laser (FemtoFiber pro TVIS, Toptica) with pulse lengths below 1 ps and a repetition rate of 80 MHz. Two cube polarizers (PBSH, CVI Melles Griot) were used to build a polarization-dependent optical delay of 6 ns to realize a pulse train of orthogonal, linearly polarized excitation pulses with a total frequency of 160 MHz. The excitation light was fed into a single-mode fiber (LMA-8, NKT Photonics) for spatial filtering. After collimation the light was reflected by a dichroic mirror and focused by a microscope objective (Zeiss C-Apochromat, 63x/1.25, WI) into the sample solution. Fluorescent light collected by the microscope objective was split by a polarizing beam splitter cube and focused onto two single-photon avalanche diodes (SPCM-AQRH-14-TR, Excelitas). The active area of the single-photon counting modules served as the confocal pinhole. The resulting effective focus volume as detected by diffusion measurements of unbound Alexa647 dye molecules was about 1.2 fL at excitation powers of 100  $\mu$ W and 638 nm. Pulses from the detectors were fed into a TCSPC board (TimeHarp260, Picoquant) operating in the time-tagged time-resolved mode with 25 ps time resolution.

### *Samples and sample preparation*

FCS experiments of concentrated BSA solutions were conducted by adding small amounts of commercially available, fluorescently labeled BSA (A34785, Thermo Fischer Scientific) to the unlabeled, concentrated solutions. The labeled BSA molecules carried 4-5 AlexaFluor647 dyes per molecule with a fluorescence lifetime of 1.2 ns. The total protein volume fractions were determined by SAXS experiments (see SI3). For FCS measurements of  $\alpha$ Bc we took advantage of the serendipitous fluorescence of some  $\alpha$ Bc molecules in the solutions as prepared for NMR or SAXS measurements, with a fluorescence lifetime of 0.86 ns (see Fig. S4). All FCS measurements were performed at a wavelength of 638 nm with excitation powers of 100  $\mu$ W and 175  $\mu$ W for BSA and  $\alpha$ Bc, respectively.

### Cross-correlation analysis

All detected fluorescence photons were sorted with respect to the polarization of the excitation pulses and according to their relative polarization (see Fig. S4). While orthogonally polarized photons were discarded, the time series of the collinear polarized photons of the two excitation polarizations were cross-correlated, leading to an anti-correlated build-up term in the correlation function on the time scale of rotational motion and a correlated decay on longer times due to translational motion (see Fig. 1 in the main text).

At small lag times  $\tau$ , the characteristic time scale of rotation can be determined by a mono-exponential fit

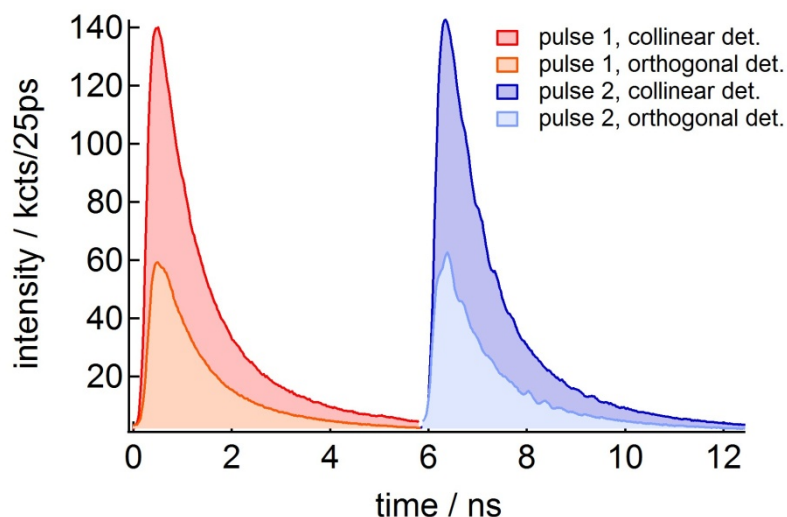
$$G(\tau) = y_0 - A \exp \{-(\tau - t_0)/\tau_r\}, \quad (\text{S1})$$

where  $y_0$  and  $A$  are some amplitude factors,  $t_0$  is a system-dependent time offset, and  $\tau_r$  is the characteristic time of rotation. Following the fitting procedure as described in ref. [7], combined correlation functions,  $G^{2 \times 1}(-\tau)$  and  $G^{1 \times 2}(\tau)$ , as displayed in Fig. S5, were used to improve the fitting accuracy. The superscripts relate to the photons detected after pulse 1 and 2 and denote the order of cross-correlation. In order to minimize the statistical noise of  $G(\tau)$  at small  $\tau$  we collected up to  $10^9$  photons for each measurement. The noise at high protein concentrations and for  $\alpha\text{Bc}$  is due to high uncorrelated background scattering and the low quantum yield of the  $\alpha\text{Bc}$  fluorescence.

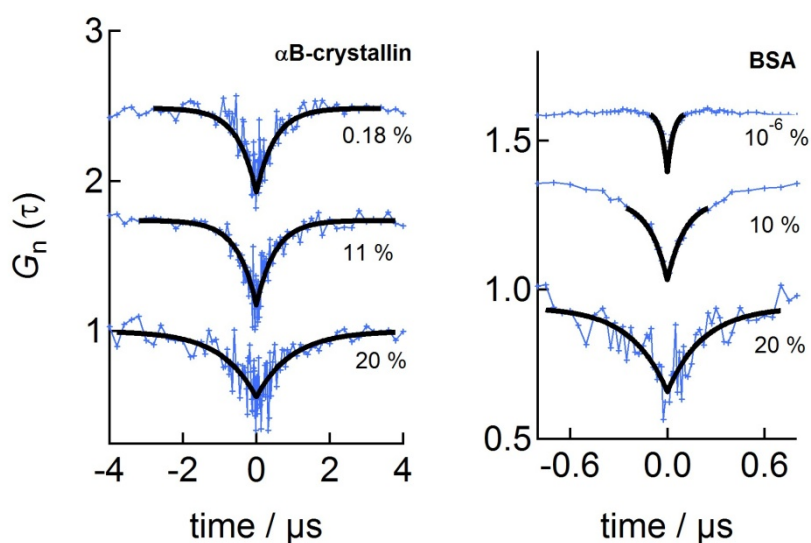
The characteristic decay time of the translational diffusion was determined at long times by the fitting function

$$G(\tau) = \frac{1}{N} \left(1 + \left(\frac{\tau}{\tau_D}\right)^\alpha\right)^{-1} \left(1 + S^2 \left(\frac{\tau}{\tau_D}\right)^\alpha\right)^{-1/2}, \quad (\text{S2})$$

where  $S$  and  $\alpha$  are system parameters describing the shape of the focal volume,  $N$  is the average number of labeled molecules in the focal volume, and  $\tau_D$  the average dwell time in the focus, being related to the inverse translational diffusion coefficient.  $G(\tau)$  was normalized with respect to the average number of molecules in the focus volume,  $G_n(t) = G(\tau) \times N$ . In our study we used  $\tau_D$  as a relative measure without the attempt to determine absolute diffusion coefficients. Control experiments with unbound dye molecules in different glycerol-water solutions could exclude an impact of the slightly increased refractive index at higher protein concentration on the determination of  $\tau_D$  for the investigated concentration range. The refractive index at 20 % protein concentration is about 1.38.



**Figure S4:** Time-correlated single-photon counting histograms allow one to sort each detected fluorescence photon with respect to the polarization of the excitation pulse (red/blue) and with respect to its relative polarization compared to the polarization of the excitation pulse (collinear vs. orthogonal). For cross-correlation analysis, only the photons collinear to the excitation polarization were used. The data set exemplarily displays the fluorescence emission of  $\alpha$ Bc.



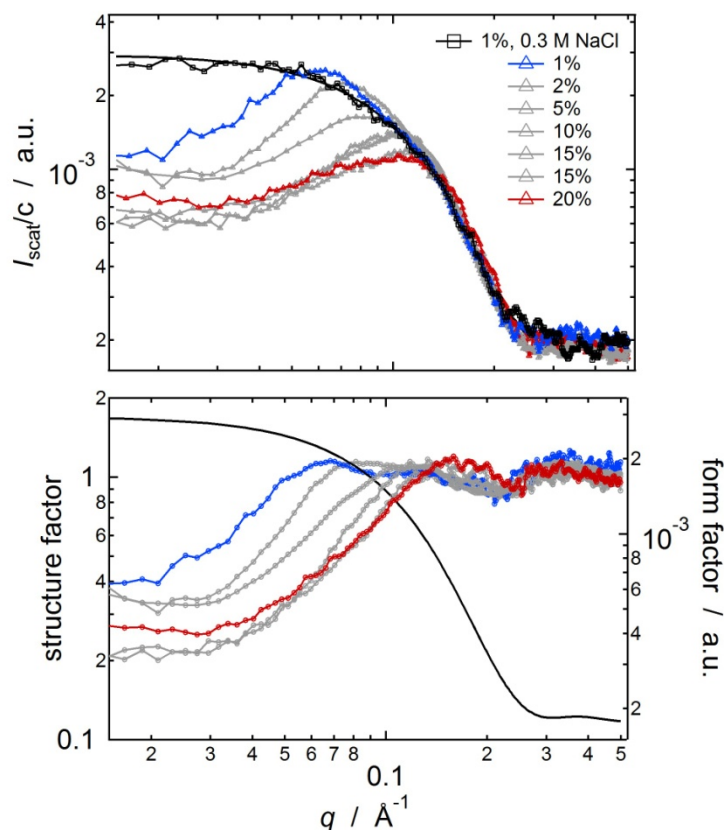
**Figure S5:** Determination of the rotational diffusion time by FCS: Normalized correlation function  $G_n(\tau)$  (blue line and symbols) of  $\alpha$ Bc (left) and BSA (right) at different volume fractions. The mono-exponential fitting functions (black lines) display a concentration dependence which is more pronounced for BSA than for  $\alpha$ Bc. Data sets were shifted vertically for clarity.

### SI3: X-ray scattering experiments

SAXS experiments of protein solutions were conducted using mark tubes made from borosilicate glass (diameter 1 mm, wall thickness 10  $\mu\text{m}$ , Hilgenberg GmbH) and using a collimated X-ray beam of a beam size of 500-600  $\mu\text{m}$ . The X-ray beam was generated by an X-ray source of rotating anode type with Cu target from Rigaku operated at 2.4 kW, and was combined with a confocal optics from Osmic to provide monochromatic Cu K $\alpha$  radiation. The X-ray beam was collimated by a system of three pinholes. The flight path was fully evacuated, and the scattered radiation was detected by a Bruker Hi-Star multiwire proportional chamber. The data were collected as frames of 1024  $\times$  1024 pixels at one ( $\alpha\text{B}$ -crystalline) or two (HEWL, BSA) distances to the detector. The resulting accessible  $q$ -range was  $0.012 \text{ \AA}^{-1} < q < 0.14 \text{ \AA}^{-1}$  or  $0.012 \text{ \AA}^{-1} < q < 0.5 \text{ \AA}^{-1}$ , respectively, which was calibrated using reflections of silver behenate.

After subtraction of the scattering background the scattering intensities were normalized according to the protein concentration, see Fig. S6a. The form factors were determined at low concentrations (0.5-2 vol%) and for HEWL and BSA with added salt (0.3-0.5 M NaCl) for electrostatic screening. The experimentally determined form factor intensities were analyzed using a program provided by NIST [8]: The form factor of BSA could be fitted to the scattering intensity of an oblate ellipsoid [9] ( $a=1.69\pm 0.02$ ,  $b=c=4.56\pm 0.02$  nm), HEWL to a prolate ellipsoid [10] ( $a=b=1.39\pm 0.01$  nm,  $c=2.44\pm 0.02$  nm) and  $\alpha\text{B}$ -crystalline to a Schulz distribution of spheres [11] with an average radius of  $5.86\pm 0.04$  nm and a polydispersity of  $0.15\pm 0.01$ . The structure factors were determined by division of the measured, background-corrected and concentration-normalized scattering intensity  $I_{\text{scat}}/c$  by the concentration-normalized form factor intensity, and were likewise analyzed using the program available from NIST [8] via applying the implemented fitting protocols for a hard-sphere, screened-Coulomb and two-Yukawa interaction potential. The position of the maximum of the structure factor did not depend upon whether the measured intensity or the fitting curve for the form factor was used for the calculation. An additional correction factor was finally used to address small concentration and/or volume uncertainties in a way that the structure factor at high  $q$ -values was about 1. Fig. S6b shows the determined structure factors exemplary for HEWL at different concentrations.





**Figure S6:** (a) Scattering intensities of HEWL at different volume fractions and experimental conditions. (b) The ellipsoidal fit to the form factor ( $a=b=1.49\pm 0.01$  nm,  $c=2.44\pm 0.01$  nm, black line) at a volume fraction of 1% and after addition of salt (0.3M NaCl) was used to derive the concentration dependent structure factors (colored and grey).

## SI4: Polydispersity and anisotropy effects in NMR relaxometry

### *Effect of polydispersity (BSA)*

In our NMR relaxometry study [2], we analyzed the NMR data of BSA in two ways: (i) Using two components of mass fractions  $P_M$  and  $P_O$ ,  $P_M + P_O = 1$ , accounting for monomers and the dispersion of oligomers, respectively, and (ii) a single component reflecting the overall slow-down of rotational diffusion of the polydisperse sample.

Applying the model accounting for monomers and oligomers, we defined the mean tumbling time  $\tau_r^*$  in analogy to eq. (4) of the main paper as

$$\frac{1}{\tau_r^*} = \frac{P_M}{\tau_{r,M}} + \frac{P_O}{\tau_{r,O}}, \quad (S3)$$

where  $\tau_{r,M}$  is the (standard) rotational correlation of monomers, and  $\tau_{r,O}$  reflects an average correlation time of the oligomers. Together with the apparent slow component of the rotational diffusion [2], the overall apparent inverse mean tumbling time, again in analogy to eq. (4) of the main article, reads

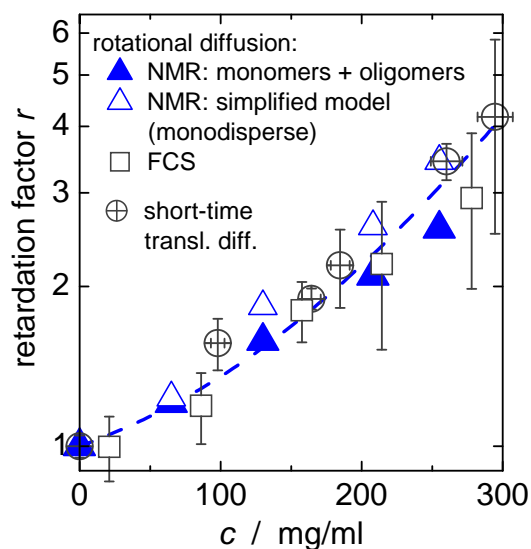
$$\frac{1}{\tau_{r,ini}} = \frac{1-a_s}{\tau_r^*} + \frac{a_s}{\tau_s} \quad (S4)$$

The quantity defined in this way corresponds to the initial slope of the rotational correlation function, and is safely determined during data fitting. Yet, when neglecting the contribution of oligomers ( $P_M = 1$ ) during data analysis, the effective weighting of oligomers relative to monomers is uncertain, and biased absolute values are obtained. Both ways of analyzing the NMR relaxometry data, however, provide the same concentration dependence, see Fig. S7. The concentration dependence is crucial when judging on coupling or decoupling of rotational diffusion from translation and viscosity.

Regarding HEWL, neither analysis of the NMR relaxometry data, nor size-exclusion chromatography indicated the presence of oligomers. HEWL data analysis on rotational diffusion thus solely relies on monomers plus the apparent long tail of rotational diffusion. The NMR relaxometry data of  $\alpha$ B-crystallin report on the average rotational correlation time of the  $\alpha$ B-crystallin assemblies [6].

Since HEWL is the only truly monodisperse protein in our study, one might tentatively take this as the origin of its specific concentration behavior in comparison to BSA and  $\alpha$ B-crystallin. However, it has been shown that both mono- and polydisperse spherical colloids reveal a clear decoupling between rotational and translational diffusion with increasing concentration [12], such that polydispersity cannot explain our main findings.

**Figure S7:** Normalized mean apparent tumbling times of BSA [2] as obtained by (i) accounting for monomers and oligomers (filled triangles), and (ii) fitting the data with a single component (open triangles). For comparison, the concentration dependencies of the FCS rotational correlation times (squares) and that of short-time translational diffusion coefficients (circles) reported in ref. [13] are shown as well. The dashed line is a guide to the eye. In the main article, the fitting result of the model including oligomers is used, as this fitting model is physically more correct



### Effect of shape anisotropy

In general the shape of a protein differs from that of a symmetric, spherical object. In our NMR relaxometry study [2] providing the data presented in the main article we assumed a spherically symmetric diffusion tensor, i.e., describing the main tumbling motion of the protein (the fast dominating component of the overall rotational correlation function, i.e. "normal" Brownian tumbling) by a single correlation time. As described below, this simplification can be justified.

As the simplest approximation accounting for a non-spherical shape, the protein may be described by an axially symmetric ellipsoid, going along with two intrinsic relaxation times,  $\tau_a$  and  $\tau_b$  for rotation about the principal axes  $a$  and  $b$ , respectively. For a prolate ellipsoid of axis ratio  $p < 1$ , the ratio  $\Gamma$  of the corresponding rotational correlation times reads [14]

$$\Gamma = \frac{\tau_a}{\tau_b} = \frac{f_b}{2} \left( \frac{1}{f_a} + \frac{1}{f_b} \right), \quad (\text{S5})$$

where

$$f_a = \frac{4(1-p^2)}{3(2-p^2s)} \quad \text{and} \quad f_b = \frac{4(1-p^4)}{3p^2(s(2-p^2)-2)}, \quad (\text{S6})$$

and

$$s = 2(1-p^2)^{-1/2} \ln \left[ \frac{1+(1-p^2)^{1/2}}{p} \right]. \quad (\text{S7})$$

The spectral density as resolved by NMR relaxometry then reads [15]

$$J(\omega) = c_1 J(\omega; \tau_1) + c_2 J(\omega; \tau_2) + c_3 J(\omega; \tau_3), \quad (\text{S8})$$

and is composed of three mixed correlation times with regards to  $\tau_a$  and  $\tau_b$ ,

$$\tau_1 = \tau_a, \quad \tau_2 = 6\tau_a/(\Gamma+5), \quad \tau_3 = 3\tau_a/(2\Gamma+1), \quad (\text{S9})$$

with the amplitudes  $c_j$  given by

$$c_1 = (3\cos^2\beta - 1)^2/4, \quad c_2 = 3\cos^2\beta \sin^2\beta, \quad c_3 = (3\sin^4\beta)/4. \quad (\text{S10})$$

Here,  $\beta$  is the angle of the inter-nuclear vector with the principal axis of the diffusion tensor. Performing high-field NMR studies, in which spectral resolution allows discriminating among different spin pairs, the impact of an ellipsoidal shape can indeed be resolved [16]. However, as we rely on a low-field NMR signal, only an integral proton signal can be analyzed. Under such circumstances, the effect of a non-spherical shape of the protein is not as evident as for high-field NMR measurements. Here, due the multitude of different orientations of the inter-nuclear vector, we assume an isotropic distribution of  $\beta$ . Then,

$$\langle c_j \rangle = \frac{1}{2} \int_0^\pi c_j \sin \beta \, d\beta = \begin{cases} 1/5 & \text{for } j=1 \\ 2/5 & \text{for } j=2 \\ 2/5 & \text{for } j=3 \end{cases} \quad (\text{S11})$$

and

$$J(\omega) = \sum_{j=1}^3 \langle c_j \rangle J(\omega; \tau_j) . \quad (\text{S12})$$

As done in the main article, we define the mean rotational correlation time  $\langle 1/\tau_r \rangle^{-1}$  by the initial slope of the rotational auto-correlation function, resulting in

$$\langle 1/\tau_r \rangle = 0.2 \tau_1^{-1} + 0.4 \tau_2^{-1} + 0.4 \tau_3^{-1} = (\Gamma + 2)/3 \tau_a^{-1} . \quad (\text{S13})$$

As  $\alpha$ Bc oligomers are near-spherical in shape, their data analysis can safely be performed using the simpler model of an isotropic object. For BSA and HEWL the axis ratio is  $p = 0.3$  and  $p = 0.6$ , respectively. For such scenario,  $\tau_1, \tau_2$  and  $\tau_3$  are still of the same order of magnitude (cf. Table S1). Since NMR relaxometry is sensitive to  $\omega\tau_r$  on a logarithmic scale, the effect of shape anisotropy is thus negligible.

To demonstrate the negligible contribution of an ellipsoidal shape on our NMR relaxometry data analysis, relaxation times of HEWL were fitted both by accounting for its ellipsoidal shape, and by using the simpler model of a spherical object. As expected, both models provide within the experimental uncertainty the same results; see Table S2. For BSA, we refrained from such an analysis since no information on the shape (and distribution of sizes) of the oligomers are available.

Given the increased extent of excluded-volume effects at high protein concentration, one may wonder about the impact of an ellipsoidal shape of the protein. Under crowding, rotation about the short principal axis can be considered to be more retarded than rotation about the long axis. In particular, the slow component of rotational diffusion might be considered to originate from this kind of rotational anisotropy, i.e., it may be possible to model it via an apparently large shape anisotropy.

**Table S1:** Impact of the ellipsoidal shape on the rotational correlation times entering the spectral density in NMR relaxometry.

sample	$\Gamma$	$\tau_1/\tau_a$	$\tau_2/\tau_a$	$\tau_3/\tau_a$	$\langle 1/\tau_r \rangle^{-1}/\tau_a$
BSA (monomer)	2.31	1	0.82	0.53	0.70
HEWL	1.26	1	0.96	0.85	0.92

**Table S2:** Fitting results for HEWL assuming a spherical object (**S**) as compared to fitting the same data using an ellipsoidal representation (**E**) of the protein. Data for (**S**) was directly taken from ref. [2]; data for (**E**) was obtained by a new analysis of the NMR relaxation times reported in ref. [2].

	tumbling motion			"long tail" of tumbling			internal dynamics			
	$c$ g/L	$\tau_r$ ns	$E_r$ kJ/mol	$a_s$	$\tau_s$ ns	$E_S$ kJ/mol	$\tau_{\text{int}}$ ns	$S_{\text{int}}^2$	$E_{\text{int}}$ kJ/mol	$K_{\text{HH}}^{(\text{av})}$ $10^9 \text{ s}^2$
<b>S</b>	65	10.5±0.2	25±1	≤0.003	≥3000	54±2				
	130	14.6±0.2	31±1	0.014±0.002	410±40	29±1	2.3	0.72	6±3	6.1
	213	23.9±0.3	27±1	0.09±0.01	330±10	73±1	±0.1	±0.01		±0.1
	257	39.9±0.6	35±2	0.23±0.01	380±10	76±1				
<b>E</b>	65	10.5±0.1	24±1	≤0.003	≥3000	52±2				
	130	16.6±0.1	31±1	0.013±0.001	450±30	28±1	2.3	0.72	9±1	6.1
	213	23.6±0.3	27±1	0.093±0.009	330±10	73±1	±0.1	±0.01		±0.1
	257	39.3±0.6	37±1	0.22±0.01	390±10	75±1				

To investigate this effect, we allowed for an adjustable axis ratio during the fitting procedure, with no additional account of the "long tail" (i.e.,  $a_s = 0$ ). Such fitting provides very large apparent axis ratios; see Table S3. Also note that the slow component of rotational diffusion already affects NMR relaxometry data analysis at 65 mg/ml HEWL, see ref. [2] and Table S2. Fitting the NMR relaxation times at this concentration using a freely adjustable apparent axis ratio, one observes an apparent axis ratio as large as 1:12 despite of the less pronounced excluded volume effect at this rather low concentration. Moreover, to compensate for the corresponding slow rotation about the short axis, the fit provides a  $\tau_r$  underestimated by 50% compared to the expected minimal value of  $\tau_r(c \rightarrow 0) \approx 9 \text{ ns}$  as estimated from  $R_h = 19 \text{ \AA}$  [17]. Given the unphysical fitting results at this low reference concentration, we refrained from considering this model, and rely on modeling the "slow tail" of rotational diffusion as described in the main text. It is also noted that the overall fitting quality in terms of the normalized  $\chi^2$  is noticeably inferior (17% as compared 13% for the long-tail model).

**Table S3:** Fitting results for HEWL when using an adjustable axis ratio (during data fitting, the value of  $\Gamma$  was adjusted), without additional parametrization of the slow component.

$c$ g/L	tumbling motion		apparent ellipsoidal shape		internal dynamics				
	$\langle 1/\tau_r \rangle^{-1}$ ns	$E_r$ kJ/mol	$\Gamma$	$\Rightarrow$ axis ratio $p$	$\tau_{\text{int}}$ ns	$S_{\text{int}}^2$	$E_{\text{int}}$ kJ/mol	$K_{\text{HH}}^{(\text{av})}$ $10^9 \text{ s}^2$	
65	6.0±0.1	30±1	13.0±0.4	~ <b>1:12</b>					
130	9.7±0.2	34±1	11.7±0.4	~ <b>1:11</b>	2.1	0.74	2	5.9	
213	20.1±0.4	50±1	22.2±0.6	~ <b>1:16</b>	±0.1	±0.01	±1	±0.1	
257	40.6±0.8	62±1	28.6±0.6	~ <b>1:19</b>					

## SI5: Size polydispersity effect on rotational diffusion of $\alpha$ B-crystallin

Given a polydisperse sample, FCS and NMR provide different kinds of averages over the ensemble. This effect leads to intrinsically different absolute values for the mean rotational correlation time.

In FCS, rotational dynamics is resolved via the initial slope of the cross-correlation function  $G_n(t)$ . Consequently, FCS provides an average over *inverse* individual rotational correlation times, see also Fig. 1 of the main article and eq. (S3),

$$\text{FCS: } \langle 1/\tau_r \rangle^{-1} \propto \langle 1/R_H^3 \rangle^{-1}, \quad (\text{S14})$$

which, at the same time, conforms to eq. (4) of the main article. Thus, FCS provides a *harmonic* average and probes the *mean inverse* hydrodynamic volume of the protein.

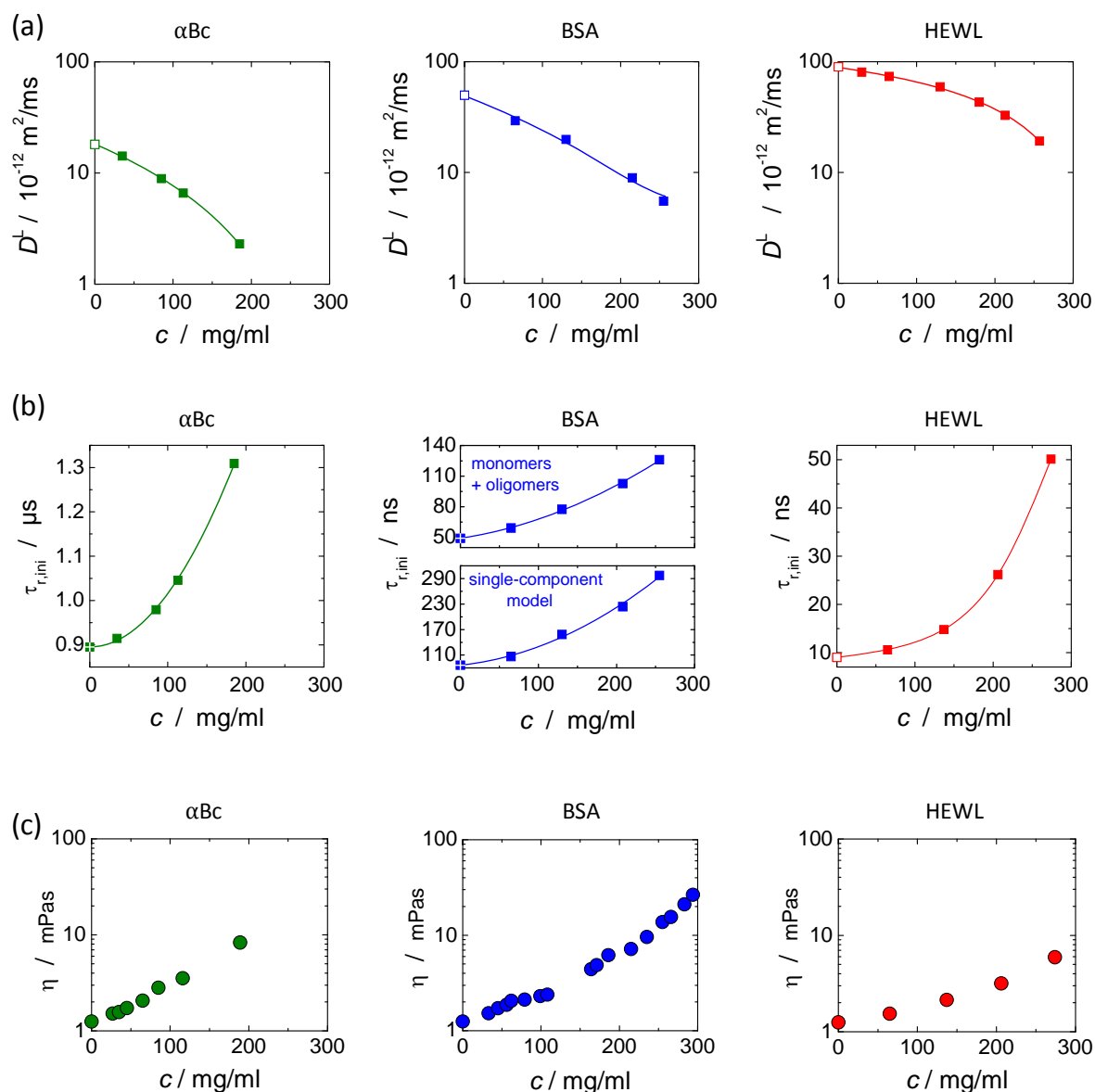
Rotational dynamics as probed by NMR relaxation times is determined by the spectral density of motion, roughly  $R_{1,2,1\rho} \propto J(\omega) \propto \tau_r / (1 + (\omega\tau_r)^2) \sim \tau_r$ . Thus, NMR relaxometry provides *not* the *harmonic* average of the individual rotational correlation times but a quantity close to the *arithmetic* average over the individual correlation times themselves, unless the distribution of correlation times (going along with a multi-component data analysis) is not considered explicitly in the fit (as done for BSA, see SI4). Thus, we have

$$\text{NMR: } \langle \tau_r \rangle \propto \langle R_H^3 \rangle. \quad (\text{S15})$$

NMR relaxometry thus probes, approximately, the *mean* hydrodynamic volume.

Since  $\langle R_H^{-3} \rangle^{-1} \leq \langle R_H^3 \rangle$ , where  $\langle R_H^n \rangle = \int_0^\infty p(R_H) R_H^n dR_H$  and  $p(R_H)$  is the distribution of hydrodynamic radii  $R_H$ , FCS naturally provides a smaller value of the mean rotational correlation time as compared to NMR relaxometry. Specifically, assuming a log-normal size distribution of  $\alpha$ Bc, with  $R_h$  between 60 and 110 Å (full width at half maximum of about 50 Å), the (small) discrepancy between the FCS- and NMR-determined correlation times can be explained. Note that for BSA, our NMR relaxometry data allowed us to distinguish between monomers and oligomers (see SI4), such that a harmonic average could be re-introduced, cf. eq. (S3).

## SI6: Absolute values of NMR-based translational and rotational diffusion coefficients, and extrapolation to zero concentration



**Figure S8:** Absolute translational diffusion coefficients (a) and mean tumbling times (b) (solid symbols) at 20°C as measured by pulsed-field gradient NMR and NMR relaxometry [2], respectively, together with their calculated (open symbols) or extrapolated (crossed filled symbols) value at zero concentration. Polynomial fits of second order (solid lines) can be used for extrapolation. Translational diffusion coefficients at zero concentration were estimated by the SE relation using the solvent ( $\text{D}_2\text{O}$ ) viscosity,  $\eta_0 = 1.25 \text{ mPas}$ , and hydrodynamic radii of 95 Å [6], 34.4 Å (Monomer) [18] and 19 Å [17] for  $\alpha$ Bc, BSA and HEWL, respectively. The rotational correlation time of HEWL at  $c=0$  was not extrapolated but calculated to be 9 ns using the SED relation. Absolute values of the viscosity at 20°C are shown in (c).

## SI7: Volume fractions and inter-particle distances

	sample	$\phi$ (steric)	$R_{cc}/\text{\AA}$	$R_{cc}/2R_H$
<b><math>\alpha</math>Bc</b>	(35 $\pm$ 3) mg/ml	0.060 $\pm$ 0.005	256 $\pm$ 22 <sup>(a)</sup>	1.35 $\pm$ 0.12
	(85 $\pm$ 3) mg/ml	0.145 $\pm$ 0.005	218 $\pm$ 8 <sup>(a)</sup>	1.15 $\pm$ 0.05
	(113 $\pm$ 5) mg/ml	0.192 $\pm$ 0.009	210 $\pm$ 10 <sup>(a)</sup>	1.10 $\pm$ 0.06
	(185 $\pm$ 5) mg/ml	0.315 $\pm$ 0.009	199 $\pm$ 6 <sup>(a)</sup>	1.05 $\pm$ 0.04
<b>BSA</b>	(65 $\pm$ 5) mg/ml	0.048 $\pm$ 0.004	104 $\pm$ 19 <sup>(b)</sup>	1.51 $\pm$ 0.28
	(130 $\pm$ 5) mg/ml	0.096 $\pm$ 0.004	87 $\pm$ 13 <sup>(b)</sup>	1.26 $\pm$ 0.19
	(208 $\pm$ 5) mg/ml	0.153 $\pm$ 0.004	77 $\pm$ 10 <sup>(b)</sup>	1.12 $\pm$ 0.15
	(255 $\pm$ 5) mg/ml	0.187 $\pm$ 0.004	73 $\pm$ 9 <sup>(b)</sup>	1.06 $\pm$ 0.13
<b>HEWL</b>	(65 $\pm$ 5) mg/ml	0.046 $\pm$ 0.003	64 $\pm$ 3 <sup>(b)</sup>	1.68 $\pm$ 0.08
	(130 $\pm$ 5) mg/ml	0.091 $\pm$ 0.003	49 $\pm$ 3 <sup>(b)</sup>	1.29 $\pm$ 0.08
	(213 $\pm$ 5) mg/ml	0.150 $\pm$ 0.003	42 $\pm$ 3 <sup>(b)</sup>	1.11 $\pm$ 0.08
	(257 $\pm$ 5) mg/ml	0.180 $\pm$ 0.003	40 $\pm$ 2 <sup>(b)</sup>	1.05 $\pm$ 0.05

**Table S4:** Steric volume fractions  $\phi$  and inter-molecular (center-to-center) distances  $R_{cc}$ , the latter being also normalized to the hydrodynamic size  $R_H$  of the molecule.  $\phi$  is calculated from the known specific volumes  $\mathcal{V}$  (1.7 ml/g [19], 0.735 ml/g [13] and 0.702 ml/g [20] for  $\alpha$ Bc, BSA and HEWL, respectively), and the center-to center distance was either calculated <sup>(a)</sup> from the volume fraction in case of hard-sphere behavior [21], or directly estimated <sup>(b)</sup> from the maximum of the SAXS structure factor,  $R_{cc} = 2\pi/q_{max}$ . The hydrodynamic radii are 95  $\text{\AA}$  [6], 34.4  $\text{\AA}$  (Monomer) [18] and 19  $\text{\AA}$  [17] for  $\alpha$ Bc, BSA and HEWL, respectively.



## References

- [1] Cotts, R.M.; Hoch, M.J.R.; Sun, T.; Markert, J.T. *J. Magn. Reson.* **1989**, *83*, 252
- [2] Roos, M.; Hofmann M.; Link, S.; Ott, M.; Balbach, J.; Rössler, E.; Saalwächter, K.; Krushelnitsky, A. *J. Biomol. NMR* **2015**, *63*, 403
- [3] Squire, P.G.; Moser, P.; O’Konski, C.T. *Biochemistry* **1968**, *7*, 4261
- [4] Atmeh, R.F.; Arafa, I.M.; Al-Khateeb, M. *Jordan J. Chem.* **2007**, *2*, 169
- [5] Langer, K.; Anhorn, M.G.; Steinhäuser, I.; Dreis, S.; Celebi, D.; Schrickel, N.; Faust, S.; Vodel, V. *Int. J. Pharm.* **2008**, *347*, 109
- [6] Roos, M.; Link, S.; Balbach, J.; Krushelnitsky, A.; Saalwächter, K. *Biophys. J.* **2015**, *108*, 98
- [7] Loman, A.; Gregor, I.; Stutz, C.; Mund, M.; Enderlein, J. *Photochem. Photobiol. Sci.* **2010**, *9*, 627
- [8] Kline, S.R. *J. Appl. Cryst.* **2006**, *39*, 895
- [9] Heinen, M.; Zanini, F.; Roosen-Runge, F.; Fedunova, D.; Zhang, F.; Hennig, M.; Seydel, T.; Schweins, R.; Sztucki, M.; Antalik, M.; Schreiber, F.; Nägele, G. *Soft Matter* **2012**, *8*, 1404
- [10] Abramo, M.C.; Caccamo, C.; Costa, D.; Pellicane, G.; Ruberto, R.; Wanderlingh, U. *J. Chem. Phys.* **2012**, *136*, 035103
- [11] Jehle, S.; Vollmar, B.S.; Bardiaux, B.; Dove, K.K.; Rajagopal, P.; Gonen, T.; Oschkinat, H.; Klevit, R.E. *Proc. Nat. Acad. Sci. USA* **2011**, *108*, 6409
- [12] Zhang, H.; Nägele, G. *J. Chem. Phys.* **2002**, *117*, 5908
- [13] Roosen-Runge, F.; Hennig, M.; Zhang, F.; Jacobs, R.M.J.; Sztucki, M.; Schober, H.; Seydel, T.; Schreiber, F. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 11815
- [14] Cantor, C.R.; Schimmel, P.R. Biophysical chemistry, part II: techniques for the study of biological structure and function. **1980**, *W. H. Freeman and Company*
- [15] Woessner, D.E. *J. Chem. Phys.* **1962**, *37*, 647
- [16] Birchall, A.J.; Lane, A.N. *Eur. Biophys. J.* **1990**, *19*, 73
- [17] Parmar, A.S.; Muschol, M. *Biophys. J.* **2009**, *97*, 590
- [18] Axelsson, I. *J. Chrom. A* **1978**, *152*, 21
- [19] Finet, S.; Tardieu, A.J. *Cryst. Growth* **2001**, *232*, 40
- [20] Chalikian, T.V.; Totrov, M.; Abagyan, R.; Breslauer, K.J. *J. Mol. Biol.* **1996**, *260*, 588
- [21] Torquato, S. *Phys. Rev. E* **1995**, *51*, 3170

## S.6. Bibliography

- T. Ando and J. Skolnick. Crowding and hydrodynamic interactions likely dominate in vivo macromolecular motion. *Proc. Nat. Acad. Sci. USA*, 107:18457–18462, 2010.
- B. Antalek. Using Pulsed Gradient Spin Echo NMR for chemical mixture analysis: How to obtain optimum results. *Concepts Magn. Reson.*, 14:225–258, 2002.
- R.C. Augusteyn. Macromolecular structure of the eye lens. *Prog. Polym. Sci.*, 23:375–413, 1998.
- J. Balbo, P. Mereghetti, D.-P. Herten, and R.C. Wade. The shape of protein crowders is a major determinant of protein diffusion. *Biophys. J.*, 104:1576–1584, 2013.
- A.J. Banchio, J. Bergenholtz, and G. Nägele. Rheology and dynamics of colloidal suspensions. *Phys. Rev. Lett.*, 82:1792–1795, 1999a.
- A.J. Banchio, G. Nägele, and J. Bergenholtz. Viscoelasticity and generalized Stokes-Einstein relations of colloidal dispersions. *J. Chem. Phys.*, 75:958–961, 1999b.
- D.S. Banks and C. Fradin. Anomalous diffusion of proteins due to molecular crowding. *Biophys. J.*, 89:2960–2971, 2005.
- P. Bartlett, S.I. Henderson, and S.J. Mitchell. Measurement of the hydrodynamic forces between two polymer-coated spheres. *Phil. Trans. R. Soc. Lond. A*, 359:883–895, 2001.
- A.V. Barzykin and A.I. Shushin. Effect of anisotropic reactivity on the rate of diffusion-controlled reactions: Comparative analysis of the models of patches and hemispheres. *Biophys. J.*, 80:2062–2073, 2001.
- P.A. Beckmann. Spectral densities and nuclear spin relaxation in solids. *Phys. Rep.*, 171:85–128, 1988.
- I. Bertini, Y.K. Gupta, C. Luchinat, G. Parigi, C. Schlörb, and H. Schwalbe. NMR spectroscopic detection of protein protons and longitudinal relaxation rates between 0.01 and 50 MHz. *Angew. Chem. Int. Ed.*, 44:2223–2225, 2005.
- S.P. Bhat and C.N. Nagineni.  $\alpha$ B subunit of lens-specific protein  $\alpha$ -crystallin is present in other ocular and non-ocular tissues. *Biochem. Biophys. Res. Commun.*, 158:319–325, 1989.
- T. Blochowicz, C. Tschirwitz, S. Benkhof, and E.A. Rössler. Susceptibility functions for slow relaxation processes in supercooled liquids and the search for universal relaxation patterns. *J. Chem. Phys.*, 118:7544–7555, 2003.

- H. Bloemendal, W. de Jong, R. Jaenicke, N.H. Lubsen, C. Slingsby, and A. Tardieu. Ageing and vision: structure, stability and function of lens crystallins. *Prog. Biophys. Mol. Biol.*, 86:407–485, 2004.
- G. Bodenhausen and D.J. Ruben. Natural abundance nitrogen-15 NMR by enhanced heteronuclear spectroscopy. *Chem. Phys. Lett.*, 69:185–189, 1980.
- A.J. Boersma, I.S. Zuhorn, and B Poolman. A sensor for quantification of macromolecular crowding in living cells. *Nature Methods*, 12:227–229, 2015.
- N. Braun, M. Zacharias, J. Peschek, A. Kastenmueller, J. Zou, M. Hanzlik, M. Haslbeck, J. Rappsilber, J. Buchner, and S. Weinkauf. Multiple molecular architectures of the eye lens chaperone  $\alpha$ B-crystallin elucidated by a triple hybrid approach. *Proc. Nat. Acad. Sci. USA*, 108:20491–20496, 2011.
- D. Bray and C. Thomas. The actin content of fibroblasts. *Biochem. J.*, 147:221–228, 1975.
- G.J. Brownsey, T.R. Noel, R. Parker, and S.G. Ring. The glass transition behavior of the globular protein bovine serum albumin. *Biophys. J.*, 85:3943–3950, 2003.
- A. Bujacz. Structures of bovine, equine and leporine serum albumin. *Acta Cryst.*, D68: 1278–1289, 2012.
- H.Y. Carr and E.M. Purcell. Effects of diffusion on free precession in nuclear magnetic resonance experiments. *Phys. Rev.*, 94:630–638, 1954.
- I. Chang and H. Sillescu. Heterogeneity at the glass transition: translational and rotational self-diffusion. *J. Phys. Chem. B*, 101:8794–8801, 1997.
- M.S. Cheung, D. Klimov, and D. Thirumalai. Molecular crowding enhances native state stability and refolding rates of globular proteins. *Proc. Natl. Acad. Sci. USA*, 102:4753–4758, 2005.
- B. Cichocki and B.U. Felderhof. Diffusion of Brownian particles with hydrodynamic interaction and hard core repulsion. *J. Chem. Phys.*, 94:556–562, 1991.
- B. Cichocki, M.L. Ekiel-Jeżewska, and E. Wajnryb. Lubrication corrections for three-particle contribution to short-time self-diffusion coefficients in colloidal dispersions. *J. Chem. Phys.*, 111:3265–3273, 1999.
- G.M. Clore, P.C. Driscoll, P.T. Wingfield, and A.M. Gronenborn. Analysis of the backbone dynamics of interleukin-1 $\beta$  using two-dimensional inverse detected heteronuclear  $^{15}\text{N}$ - $^1\text{H}$  NMR spectroscopy. *Biochemistry*, 29:7387–7401, 1990a.

- G.M. Clore, A. Szabo, A. Bax, L.E. Kay, P.C. Driscoll, and A.M. Gronenborn. Deviations from the simple two-parameter model-free approach to the interpretation of nitrogen-15 nuclear magnetic relaxation of proteins. *J. Am. Chem. Soc.*, 112:4989–4991, 1990b.
- R.M. Cotts, M.J.R. Hoch, T. Sun, and J.T. Markert. Pulsed field gradient stimulated echo methods for improved NMR diffusion measurements in heterogeneous systems. *J. Magn. Reson.*, 83:252–266, 1989.
- J. Danielsson, X. Mu, L. Lang, H. Wang, A. Binolfi, F.-X. Theillet, B. Bekei, D.T. Logan, Ph. Selenko, H. Wennerström, and M. Oliveberg. Thermodynamics of protein destabilization in live cells. *Proc. Nat. Acad. Sci. USA*, 112:12402–12407, 2015.
- M.M. Dedmon, Ch.N. Patel, G.B. Young, and G.J. Pielak. FlgM gains structure in living cells. *Proc. Nat. Acad. Sci. USA*, 99:12681–12684, 2002.
- M. Delaye and A. Tardieu. Short-range order of crystallin proteins accounts for eye lens transparency. *Nature*, 302:415–417, 1983.
- B.K. Derham and J.J. Harding.  $\alpha$ -crystallin as a molecular chaperone. *Prog. Retin. Eye Res.*, 18:463–509, 1999.
- R. Diamond. Real-space refinement of the structure of hen egg-white lysozyme. *J. Mol. Biol.*, 82:371–391, 1974.
- J.A. Dix and A.S. Verkman. Crowding effects on diffusion in solutions and cells. *Annu. Rev. Biophys.*, 37:247–263, 2008.
- M. Długosz and J.M. Antosiewicz. Evaluation of proteins' rotational diffusion coefficients from simulations of their free Brownian motion in volume-occupied environments. *J. Chem. Theory Comput.*, 10:481–491, 2014.
- M. Długosz and J. Trylska. Diffusion in crowded biological environments: applications of Brownian dynamics. *BMC Biophysics*, 4:1–9, 2011.
- B. Doliwa and A. Heuer. Cage effect, local anisotropies, and dynamic heterogeneities at the glass transition: A computer study of hard spheres. *Phys. Rev. Lett.*, 80:4915–4918, 1998.
- J.L. Dote, D. Kivelson, and R.N. Schwartz. A molecular quasi-hydrodynamic free-space model for molecular rotational relaxation in liquids. *J. Phys. Chem.*, 85:2169–2180, 1981.
- D. Drenckhahn and T.D. Pollard. Elongation of actin filaments is a diffusion-limited reaction at the barbed end and is accelerated by inert macromolecules. *J. Biol. Chem.*, 261:12754–12758, 1986.

- R.A. Dubin, E.F. Wawrousek, and J. Piatigorsky. Expression of the murine alpha B-crystallin gene is not restricted to the lens. *Mol. Cell Biol.*, 9:1083–1091, 1989.
- A.K. Dunker, I. Silman, V.N. Uversky, and J.L. Sussman. Function and structure of inherently disordered proteins. *Curr. Opin. Struct. Biol.*, 18:756–764, 2008.
- H.J. Dyson and P.E. Wright. Intrinsically unstructured proteins and their functions. *Nat. Rev. Mol. Cell Biol.*, 6:197–208, 2005.
- H. Ecroyd and J.A. Carver. Crystallin proteins and amyloid fibrils. *Cell. Mol. Life Sci.*, 66: 62–81, 2009.
- K.V. Edmond, M.T. Elsesser, G.L Hunter, D.J Pine, and E.R. Weeks. Decoupling of rotational and translational diffusion in supercooled colloidal fluids. *Proc. Nat. Acad. Sci. USA*, 109: 17891–17896, 2012.
- J.T. Edward. Molecular volumes and the Stokes-Einstein equation. *J. Chem. Educ.*, 47: 261–270, 1970.
- A. Einstein. Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen. *Ann. Phys.*, 322:549–560, 1905.
- A.H. Elcock. Models of macromolecular crowding effects and the need for quantitative comparisons with experiment. *Curr. Opin. Struct. Biol.*, 20:196–206, 2010.
- R.J. Ellis. Macromolecular crowding: an important but neglected aspect of the intracellular environment. *Curr. Opin. Struct. Biol.*, 11:114–119, 2001a.
- R.J. Ellis. Macromolecular crowding: obvious but underappreciated. *Trends Biochem. Sci.*, 26:597–604, 2001b.
- R.J. Ellis and A.P. Minton. Join the crowd. *Nature*, 425:27–28, 2003.
- R.J. Ellis and A.P. Minton. Protein aggregation in crowded environments. *Biol. Chem.*, 387: 485–497, 2006.
- E. Ermakova, A.G. Krushelnitsky, and V.D. Fedotov. Brownian dynamics simulation of electrostatically interacting proteins. *Mol. Phys.*, 100:2849–2855, 2002.
- I.V. Ermolina, A.G. Krushelnitsky, I.N. Ivoylov, Y.D. Feldman, and V.D. Fedotov. Investigation of molecular motions and interprotein interactions by NMR and TDDS. *Appl. Magn. Reson.*, 5:265–283, 1993.

- P.P. Fagerholm, B.T. Philipson, and B. Lindström. Normal human lens - the distribution of proteins. *Exp. Eye. Res.*, 33:615–620, 1981.
- M. Feig and Y. Sugita. Variable interactions between protein crowders and biomolecular solutes are important in understanding cellular crowding. *J. Phys. Chem. B*, 116:599–605, 2012.
- R.B. Fenwick and H.J. Dyson. Classic analysis of biopolymer dynamics is model free. *Biophys. J.*, 110:3–6, 2016.
- F. Fujara, D. Kruk, and A.F. Privalov. Solid state field-cycling NMR relaxometry: Instrumental improvements and new applications. *Prog. Nucl. Magn. Reson. Spectrosc.*, 82:39–69, 2014.
- S.J. Gibbs and C.S. Johnson Jr. A PFG NMR experiment for accurate diffusion and flow studies in the presence of eddy currents. *J. Magn. Reson.*, 93:395–402, 1991.
- D. Gnutt and S. Ebbinghaus. The macromolecular crowding effect - from *in vitro* into the cell. *Biol. Chem.*, 397:37–44, 2016.
- D. Gnutt, M. Gao, O. Brylski, M. Heyden, and S. Ebbinghaus. Excluded-volume effects in living cells. *Angew. Chem. Int. Ed.*, 54:2548–2551, 2015.
- M.L. Grant. Nonuniform charge effects in protein-protein interactions. *J. Phys. Chem. B*, 105:2858–2863, 2001.
- M. Grimaldo, F. Roosen-Runge, F. Zhang, T. Seydel, and F. Schreiber. Diffusion and dynamics of  $\gamma$ -globulin in crowded aqueous solutions. *J. Phys. Chem. B*, 118:7203–7209, 2014.
- S. Gupta, J. Stellbring, E. Zaccarelli, C.N. Likos, M. Camargo, P. Holmqvist, J. Allgaier, L. Willner, and D. Richter. Validity of the Stokes-Einstein relation in soft colloids up to the glass transition. *Phys. Rev. Lett.*, 115:128302, 2015.
- S. Gupta, R. Biehl, C. Sill, J. Allgaier, M. Sharp, M. Ohl, and D. Richter. Protein entrapment in polymeric mesh: Diffusion in crowded environment with fast process on short scales. *Macromolecules*, 49:1941–1949, 2016.
- E.L. Hahn. Spin echoes. *Phys. Rev.*, 80:580–594, 1950.
- D. Hall and A.P. Minton. Macromolecular crowding: qualitative and semiquantitative successes, quantitative challenges. *Biochim. Biophys. Acta*, 1649:127–139, 2003.

- J. Han and Herzfeld. Macromolecular diffusion in crowded solutions. *Biophys. J.*, 65: 1155–1161, 1993.
- Y. Han, A. M. Alsayed, M. Nobili, J. Zhang, T. C. Lubensky, and A. G. Yodh. Brownian motion of an ellipsoid. *Science*, 314:626–630, 2006.
- B.P. Hills, S.F. Takacs, and P.S. Belton. The effects of proteins on the proton N.M.R. transverse relaxation times of water. I. Native bovine serum albumin. *Mol. Phys.*, 67:903–918, 1989.
- H. Hofmann, A. Soranno, A. Borgia, K. Gast, D. Nettels, and B. Schuler. Polymer scaling laws of unfolded and intrinsically disordered proteins quantified with single-molecule spectroscopy. *Proc. Nat. Acad. Sci. USA*, 109:16155–16160, 2012.
- J. Hong and L.M. Gierasch. Macromolecular crowding remodels the energy landscape of a protein by favoring a more compact unfolded state. *J. Am. Chem. Soc.*, 132:10445–10452, 2010.
- J. Horwitz. The function of alpha-crystallin in vision. *Semin. Cell Dev. Biol.*, 11:53–60, 2000.
- J. Horwitz.  $\alpha$ -crystallin. *Exp. Eye. Res.*, 76:145–153, 2003.
- C.-M. Hu and R. Zwanzig. Rotational friction coefficients for spheroids with the slipping boundary condition. *J. Chem. Phys.*, 60:4354–4357, 1974.
- G.L. Hunter and E.R. Weeks. The physics of the colloidal glass transition. *Rep. Prog. Phys.*, 75:066501, 2012.
- S. Ikeda and K. Nishinari. Intermolecular forces in bovine serum albumin solutions exhibiting solidlike mechanical behaviors. *Biomacromolecules*, 1:757–763, 2000.
- S. Ikeda and K. Nishinari. Solid-like mechanical behaviors of ovalbumin aqueous solutions. *Int. J. Biol. Macromol.*, 28:315–320, 2001a.
- S. Ikeda and K. Nishinari. On solid-like rheological behaviors of globular protein solutions. *Food Hydrocolloids*, 15:401–406, 2001b.
- H. Inoue and T. Matsumoto. Viscoelastic characterization of solid-like structure in aqueous colloids of globular proteins. *Colloids Surf., A*, 109:89–96, 1996.
- T. Iwaki, A. Kume-Iwaki, R.K. Liem, and J.E. Goldman.  $\alpha$ B-crystallin is expressed in non-lenticular tissues and accumulates in Alexander's disease brain. *Cell*, 57:71–78, 1989.

- T. Iwaki, T. Wisniewski, I. Akiko, E. Corbin, N. Tomokane, J. Tateishi, and J.E. Goldman. Accumulation of alpha B-crystallin in central nervous system glia and neurons in pathologic conditions. *Am. J. Pathol.*, 140:345–356, 1992.
- A. Jerschow and N. Müller. Suppression of convection artifacts in stimulated-echo diffusion experiments. Double-Stimulated-Echo experiments. *J. Magn. Reson.*, 125:372–375, 1997.
- C.S. Johnson Jr. Diffusion ordered nuclear magnetic resonance spectroscopy: principles and applications. *Prog. Nucl. Magn. Reson. Spectrosc.*, 34:203–256, 1999.
- M.T. Johnston and R.H. Ewoldt. Precision rheometry: Surface tension effects on low-torque measurements in rotational rheometers. *J. Rheol.*, 57:1515–1532, 2013.
- G.P. Jones. Spin-lattice relaxation in rotating frame: weak-collision case. *Phys. Rev.*, 148:332–335, 1966.
- R.B. Jones. Rotational diffusion of a tracer colloid particle: 1. short-time orientational correlations. *Physica A*, 150:339–356, 1988.
- E. Jordan, F. Roosen-Runge, S. Leibfarth, F. Zhang, M. Sztucki, A. Hildebrandt, O. Kohlbacher, and F. Schreiber. Competing salt effects on phase behavior of protein solutions: Tailoring of protein interaction by the binding of multivalent ions and charge screening. *J. Phys. Chem. B*, 118:11365–11374, 2014.
- W.K. Kegel and A. van Blaaderen. Direct observation of dynamical heterogeneities in colloidal hard-sphere suspensions. *Science*, 287:290–293, 2000.
- J.S. Kim and A. Yethiraj. Effect of macromolecular crowding on reaction rates: A computational and theoretical study. *Biophys. J.*, 96:1333–1340, 2009.
- M. Kim, S.M. Anthony, S.C. Bae, and S. Granick. Colloidal rotation near the colloidal glass transition. *J. Chem. Phys.*, 135:054905, 2011.
- R. Kimmich. Field cycling in NMR relaxation spectroscopy: Applications in biological, chemical and polymer physics. *Bull. Magn. Reson.*, 1:195–218, 1980.
- R. Kimmich and E. Anoardo. Field-cycling NMR relaxometry. *Prog. Nucl. Magn. Reson. Spectrosc.*, 44:257–320, 2004.
- S. Kirschke. Untersuchung der Diffusion von Proteinen mittels Isotopen-gefilteter PFG-NMR. Bachelor's thesis, Martin-Luther-Universität Halle-Wittenberg, Institut für Physik, 2013.



- D. Kleshchanok, M. Heinen, G. Nägele, and P. Holmqvist. Dynamics of charged gibbsite platelets in the isotropic phase. *Soft Matter*, 8:1584–1592, 2012.
- S. Klumpp, M. Scott, S. Pedersen, and T. Hwa. Molecular crowding limits translation and cell growth. *Proc. Nat. Acad. Sci. USA*, 110:16754–16759, 2013.
- G.H. Koenderink and A.P. Philipse. Rotational and translational self-diffusion in colloidal sphere suspensions and the applicability of generalized Stokes-Einstein relations. *Langmuir*, 16:5631–5638, 2000.
- G.H. Koenderink, H. Zhang, M.P. Lettinga, G. Nägele, and A. Philipse. Rotational tracer diffusion in binary colloidal sphere mixtures. *Phys. Rev. E*, 64:022401, 2001.
- G.H. Koenderink, M.P. Lettinga, and A.P. Philipse. Rotational dynamics of charged colloidal spheres: Role of particle interactions. *J. Chem. Phys.*, 117:7751–7764, 2002.
- G.H. Koenderink, H. Zhang, D.G.A.L. Aarts, M.P. Lettinga, A.P. Philipse, and G. Nägele. On the validity of Stokes-Einstein-Debye relations for rotational diffusion in colloidal suspensions. *Faraday Discuss.*, 123:335–354, 2003.
- M.C. Konopka, I.A. Shkel, S. Cayley, M.T. Record, and J.C. Weisshaar. Crowding and confinement effects on protein diffusion in vivo. *J. Bacteriol.*, 188:6115–6123, 2006.
- B. Kresse, A.F. Privalov, and F. Fujara. NMR field-cycling at ultralow magnetic fields. *Solid State Nucl. Magn. Reson.*, 40:134–137, 2011.
- B. Kresse, A.F. Privalov, A. Herrmann, M. Hofmann, E.A. Rössler, and F. Fujara. Simultaneous measurement of very small magnetic fields and spin-lattice relaxation. *Solid State Nucl. Magn. Reson.*, 59-60:45–47, 2014.
- D. Kruk, A. Herrmann, and E. Rössler. Field-cycling NMR relaxometry of viscous liquids and polymers. *Prog. Nucl. Magn. Reson. Spectrosc.*, 63:33–64, 2012.
- A. Krushelnitsky and D. Reichert. Response of lysozyme internal dynamics to hydration probed by  $^{13}\text{C}$  and  $^1\text{H}$  solid-state NMR relaxation. *Appl. Magn. Reson.*, 27:501–518, 2004.
- A.G. Krushelnitsky. Intermolecular electrostatic interactions and Brownian tumbling in protein solutions. *Phys. Chem. Chem. Phys.*, 8:2117–2128, 2006.
- R. Kurbanov, T. Zinkevich, and A. Krushelnitsky. The nuclear magnetic resonance relaxation data analysis in solids: General  $R_1/R_{1\rho}$  equations and the model-free approach. *J. Chem. Phys.*, 135:184104, 2011.

- A. Kurut, B.A. Persson, T. Åkesson, J. Forsman, and M. Lund. Anisotropic interactions in protein mixtures: self assembly and phase behavior in aqueous solution. *J. Phys. Chem. Lett.*, 3:731–734, 2012.
- Y.Y. Kuttner, N.K. Kozer, E. Segal, G. Schreiber, and G. Haran. Separating the contribution of translational and rotational diffusion to protein association. *J. Am. Chem. Soc.*, 127: 15138–15144, 2005.
- P. Langevin. Sur la théorie du mouvement brownien. *C. R. Acad. Sci. (Paris)*, 146:530–533, 1908. Accessed via its translation into English; see D.S. Lemons and A. Gythiel. Paul Langevin's 1908 paper "on the theory of Brownian motion" ["Sur la théorie du mouvement brownien," *C. R. Acad. Sci. (Paris)* 146, 530-533 (1908)]. *Am. J. Phys.*, 65:1079-1081, 1997.
- C. Li, Y. Wang, and Pielak G.J. Translational and rotational diffusion of a small globular protein under crowded conditions. *J. Phys. Chem. B*, 113:13390–13392, 2009.
- P. Licinio and M. Delaye. Mutual and self-diffusion in concentrated  $\alpha$ -crystallin protein dispersion. A dynamic light-scattering study. *J. Phys. (Paris)*, 49:975–981, 1988.
- G. Lipari and A. Szabo. Effect of librational motion on fluorescence depolarization and nuclear magnetic resonance relaxation in macromolecules and membranes. *Biophys. J.*, 30:489–506, 1980.
- G. Lipari and A. Szabo. Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 1. Theory and range of validity. *J. Am. Chem. Soc.*, 104:4546–4559, 1982a.
- G. Lipari and A. Szabo. Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 2. Analysis of experimental results. *J. Am. Chem. Soc.*, 104:4559–4570, 1982b.
- Y. Liu, L. Porcar, J. Chen, W.-R. Chen, P. Falus, A. Faraone, E. Fratini, K. Hong, and P. Baglioni. Lysozyme protein solution with an intermediate range order structure. *J. Phys. Chem. B*, 115:7238–7247, 2011.
- J. Lowe, H. McDermott, I. Pike, I. Speudlove, M. Landon, and R. Mayer.  $\alpha$ B-crystallin expression in nonlenticular tissues and selective presence in ubiquitinated inclusion bodies in human disease. *J. Pathol.*, 166:61–68, 1992.
- K. Luby-Phelps. Cytoarchitecture and physical properties of cytoplasm: volume, viscosity, diffusion, intracellular surface area. *Int. Rev. Cytol.*, 192:189–221, 2000.
- C. Luchinat and G. Parigi. Collective relaxation of protein protons at very low magnetic field: A new window on protein dynamics and aggregation. *J. Am. Chem. Soc.*, 129:

- 1055–1064, 2007.
- Z. Luz and S. Meiboom. Nuclear magnetic resonance study of the protolysis of trimethylammonium ion in aqueous solution—order of the reaction with respect to solvent. *J. Chem. Phys.*, 39:366–370, 1963.
- P.D. Majors, J.L. Blackley, S.A. Altobelli, A. Caprihan, and E. Fukushima. Eddy current compensation by direct field detection and digital gradient modification. *J. Magn. Reson.*, 87:548–553, 1990.
- R.B. McClurg and C.F. Zukoski. The electrostatic interaction of rigid, globular proteins with arbitrary charge distributions. *J. Colloid Interface Sci.*, 208:529–542, 1998.
- S.R. McGuffee and A.H. Elcock. Atomically detailed simulations of concentrated protein solutions: The effects of salt, pH, point mutations, and protein concentration in simulations of 1000-molecule systems. *J. Am. Chem. Soc.*, 128:12098–12110, 2006.
- S.R. McGuffee and A.H. Elcock. Diffusion, crowding & protein stability in a dynamic molecular model of the bacterial cytoplasm. *PLoS Comput. Biol.*, 6:e1000694, 2010.
- M.G. McPhie and G. Nägele. Long-time self-diffusion of charged colloidal particles: electrokinetic and hydrodynamic interaction effects. *J. Chem. Phys.*, 127:034906, 2007.
- O. Medalia, I. Weber, Frangakis A.S., D. Nicastro, G. Gerisch, and W. Baumeister. Macromolecular architecture in eukaryotic cells visualized by cryoelectron tomography. *Science*, 298:1209–1213, 2002.
- E. Meirovitch, Y.E. Shapiro, A. Polimeno, and J.H. Freed. Protein dynamics from NMR: The slowly relaxing local structure analysis compared with Model-Free analysis. *J. Phys. Chem. A*, 110:8366–8396, 2006.
- E. Meirovitch, Y.E. Shapiro, A. Polimeno, and J.H. Freed. Structural dynamics of biomacromolecules by NMR: The slowly relaxing local structure approach. *Prog. Nucl. Magn. Reson. Spectrosc.*, 56:360–405, 2010.
- P. Mereghetti and R.C. Wade. Atomic detail Brownian dynamics simulations of concentrated protein solutions with a mean field treatment of hydrodynamic interactions. *J. Phys. Chem. B*, 116:8523–8533, 2012.
- P. Mereghetti, R.R. Gabdoulline, and R.C. Wade. Brownian dynamics simulation of protein solutions: structural and dynamical properties. *Biophys. J.*, 99:3782–9791, 2010.
- A.C. Miklos, M. Sarkar, Y. Wang, and G.J. Pielak. Protein crowding tunes protein stability. *J. Am. Chem. Soc.*, 133:7116–7120, 2011.

- A.P. Minton. Influence of excluded volume upon macromolecular structure and associations in 'crowded' media. *Curr. Opin. Biotechnol.*, 8:65–69, 1997.
- A.P. Minton. Implications of macromolecular crowding for protein assembly. *Curr. Opin. Struct. Biol.*, 10:34–39, 2000.
- A.P. Minton. The influence of macromolecular crowding and macromolecular confinement on biochemical reactions in physiological media. *J. Biol. Chem.*, 276:10577–10580, 2001.
- A.P. Minton. Influence of macromolecular crowding upon the stability and state of association of proteins: Predictions and observations. *J. Pharm. Sci.*, 94:1668–1675, 2005.
- A.P. Minton. How can biochemical reactions within cells differ from those in test tubes? *J. Cell Sci.*, 119:2863–2869, 2006.
- Sh. Mittal, R.K. Chowhan, and S. Rajendrakumar. Macromolecular crowding: Macromolecules friend or foe. *Biochim. Biophys. Acta*, 1850:1822–1831, 2015.
- G.A. Morris and R. Freeman. Enhancement of nuclear magnetic resonance signals by polarization transfer. *J. Am. Chem. Soc.*, 101:760–762, 1979.
- M.A. Mourão, J.B. Hakim, and S. Schnell. Connecting the dots: The effects of macromolecular crowding on cell physiology. *Biophys. J.*, 107:2761–2766, 2014.
- N. Muramatsu and A.P. Minton. Tracer diffusion of globular proteins in concentrated protein solutions. *Proc. Nat. Acad. Sci. USA*, 85:2984–2988, 1988.
- I.V. Nesmelova, D. Idiyatullin, and K.H. Mayo. Measuring protein self-diffusion in diffusion-protein mixtures using a pulsed gradient spin-echo technique with WATERGATE and isotope filtering. *J. Magn. Reson.*, 166:129–133, 2004.
- G. Nägele. Viscoelasticity and diffusional properties of colloidal model dispersions. *J. Phys.: Condens. Matter*, 15:S407–S414, 2003.
- T.J. O’Leary. Concentration dependence of protein diffusion. *Biophys. J.*, 52:137–139, 1987.
- F. Ould-Kaddour and D. Levesque. Molecular-dynamics investigation of tracer diffusion in a simple liquid: Test of the Stokes-Einstein law. *Phys. Rev. E*, 63:011205, 2000.
- D.J. Owen, P. Wigge, Y. Vallis, J.D. Moore, P.R. Evans, and H.T. McMahon. Crystal structure of the amphiphysin-2 SH3 domain and its role in the prevention of dynamin ring formation. *EMBO J.*, 17:5273–5285, 1998.

- A. Paciaroni, S. Cinelli, and G. Onori. Effect of the environment on the protein dynamical transition: A neutron scattering study. *Biophys. J.*, 83:1157–1164, 2002.
- B.R. Parry, I.V. Surovtsev, M.T. Cabeen, C.S. O’Hern, E.R. Dufresne, and C. Jacobs-Wagner. The bacterial cytoplasm has glass-like properties and is fluidized by metabolic activity. *Cell*, 156:183–194, 2014.
- D. Pascolini and S.P. Mariotti. Global estimates of visual impairment: 2010. *Br. J. Ophthalmol.*, 96:614–618, 2012.
- J. Peschek, N. Braun, T.M. Franzmann, Y. Georgalis, M. Haslbeck, S. Weinkauff, and J. Buchner. The eye lens chaperone  $\alpha$ -crystallin forms defined globular assemblies. *Proc. Nat. Acad. Sci. USA*, 106:13272–13277, 2009.
- W.S. Price. Pulsed-Field Gradient Nuclear Magnetic Resonance as a tool for studying translational diffusion: Part 1. Basic theory. *Concepts Magn. Reson.*, 9:299–336, 1997.
- W.S. Price. Pulsed-Field Gradient Nuclear Magnetic Resonance as a tool for studying translational diffusion: Part II. Experimental aspects. *Concepts Magn. Reson.*, 10:197–237, 1998.
- P.N. Pusey. Colloidal glasses. *J. Phys.: Condens. Matter*, 20:494202, 2008.
- P.N. Pusey and W. van Meegen. Observation of a glass transition in suspensions of spherical colloidal particles. *Phys. Rev. Lett.*, 59:2083–2086, 1987.
- G.B. Ralston. Effects of “crowding” in protein solutions. *J. Chem. Educ.*, 67:857–860, 1990.
- A.G. Redfield. Shuttling device for high-resolution measurements of relaxation and related phenomena in solution at low field, using a shared commercial 500 MHz NMR instrument. *Magn. Reson. Chem.*, 41:753–768, 2003.
- A.G. Redfield. High-resolution NMR field-cycling device for full-range relaxation and structural studies of biopolymers on a shared commercial instrument. *J. Biomol. NMR*, 52:159–177, 2012.
- K. Renkawek, W.W. de Jong, K.B. Merk, C.W.G.M. Frenken, F.P.A. van Workum, and G.J.C.G.M. Bosman.  $\alpha$ B-crystallin is present in reactive glia in Creutzfeldt-Jakob disease. *Acta. Neuropathol.*, 83:324–327, 1992.
- J. Riest and G. Nägele. Short-time dynamics in dispersions with competing short-range attraction and long-range repulsion. *Soft Matter*, 11:9273–9280, 2015.

- D. Ringe and G.A. Petsko. The 'glass transition' in protein dynamics: what it is, why it occurs, and how to exploit it. *Biophys. Chem.*, 105:667–680, 2003.
- M. Roos, M. Hofmann, S. Link, M. Ott, J. Balbach, E. Rössler, K. Saalwächter, and A. Krushelnitsky. The "long tail" of the protein tumbling correlation function: observation by  $^1\text{H}$  NMR relaxometry in a wide frequency and concentration range. *J. Biomol. NMR*, 63: 403–415, 2015a. *This paper is part of the dissertation at hand.*
- M. Roos, S. Link, J. Balbach, A. Krushelnitsky, and K. Saalwächter. NMR-detected Brownian dynamics of  $\alpha\text{B}$ -crystallin over a wide range of concentrations. *Biophys. J.*, 108:98–106, 2015b. *This paper is part of the dissertation at hand.*
- M. Roos, M. Ott, M. Hofmann, S. Link, E. Rössler, J. Balbach, A. Krushelnitsky, and K. Saalwächter. Coupling and decoupling of rotational and translational diffusion of proteins under crowding conditions. *J. Am. Chem. Soc.*, 138:10365–10372, 2016. *This paper is part of the dissertation at hand.*
- F. Roosen-Runge, M. Hennig, F. Zhang, R.M.J. Jacobs, M. Sztucki, H. Schober, T. Seydel, and F. Schreiber. Protein self-diffusion in crowded solutions. *Proc. Nat. Acad. Sci. USA*, 108: 11815–11820, 2011.
- D. Rosenbaum, P. C. Zamora, and C. F. Zukoski. Phase behavior of small attractive colloidal particles. *Phys. Rev. Lett.*, 76:150–153, 1996.
- M. Rothe. Diffusion in konzentrierten Proteinmischungen mittels PFG-NMR. Master's thesis, Martin-Luther-Universität Halle-Wittenberg, Institut für Physik, May 2015.
- M. Rothe, T. Gruber, S. Gröger, J. Balbach, K. Saalwächter, and M. Roos. Transient binding accounts for apparent violation of the generalized stokes-einstein relation in crowded protein solutions. *Phys. Chem. Chem. Phys.*, 18:18006–18014, 2016. *This paper is part of the dissertation at hand.*
- M. Rubinstein and R.H. Colby. *Polymer Physics*. Oxford Univ. Press, New York, 2003.
- L. Sapir and D. Harries. Is the depletion force entropic? Molecular crowding beyond steric interactions. *Curr. Opin. Colloid Interface Sci.*, 20:3–10, 2015.
- M.J. Saxton. Anomalous diffusion due to obstacles: A Monte Carlo study. *Biophys. J.*, 66: 394–401, 1994.
- A.P. Schlesinger, Y. Wang, X. Tadeo, O. Millet, and Pielak G.J. Macromolecular crowding fails to fold a globular protein in cells. *J. Am. Chem. Soc.*, 133:8082–8085, 2011.

- H.R. Schober and H.L. Peng. Heterogeneous diffusion, viscosity, and the Stokes-Einstein relation in binary liquids. *Phys. Rev. E*, 93:052607, 2016.
- H. Sedgwick, K. Kroy, A. Salonen, M.B. Robertson, S.U. Egelhaaf, and W.C.K. Poon. Non-equilibrium behavior of sticky colloidal particles: beads, clusters and gels. *Eur. Phys. J. E*, 16:77–80, 2005.
- P.N. Segre, S.P. Meeker, P.N. Pusey, and Poon W.C.K. Viscosity and structural relaxation in suspensions of hard-sphere colloids. *Phys. Rev. Lett.*, 75:958–961, 1995.
- M. Senske, L. Törk, B. Born, M. Havenith, C. Herrmann, and S. Ebbinghaus. Protein stabilization by macromolecular crowding through enthalpy rather than entropy. *J. Am. Chem. Soc.*, 136:9036–9041, 2014.
- M. Senske, A.E. Smith, and Pielak G.J. Protein stability in reverse micelles. *Angew. Chem. Int. Ed.*, 55:3586–3589, 2016.
- B. Sharma, A. Jaishankar, Y.-C. Wang, and G.H. McKinley. Rheology of globular proteins: apparent yield stress, high shear rate viscosity and interfacial viscoelasticity of bovine serum albumin solutions. *Soft Matter*, 7:5150–5160, 2011.
- M.C.R. Shastry and M.R. Eftink. Reversible thermal unfolding of ribonuclease T<sub>1</sub> in reverse micelles. *Biochemistry*, 35:4094–4101, 1996.
- D. Shoup, G. Lipari, and A. Szabo. Diffusion-controlled biomolecular reaction rates. *Biophys. J.*, 36:697–714, 1981.
- H. Sillescu. Heterogeneity at the glass transition: a review. *J. Non-Cryst. Solids*, 243:81–108, 1999.
- A.E. Smith, L.Z. Zhou, A.H. Gorenssek, M. Senske, and G.J. Pielak. In-cell thermodynamics and a new role for protein surfaces. *Proc. Nat. Acad. Sci. USA*, 113:1725–1730, 2016.
- M. Smoluchowski. Zur kinetischen Theorie der Brownschen Molekularbewegung und der Suspensionen. *Ann. Phys.*, 326:756–780, 1906.
- A. Soranno, I. Koenig, M.B. Borgia, H. Hofmann, F. Zosel, D. Nettels, and B. Schuler. Single-molecule spectroscopy reveals polymer effects of disordered proteins in crowded environments. *Proc. Nat. Acad. Sci. USA*, 111:4874–4879, 2015.
- E.O. Stejskal and J.E. Tanner. Spin diffusion measurements: Spin echoes in the presence of a time-dependent field gradient. *J. Chem. Phys.*, 42:288–292, 1965.

- F.H. Stillinger and J.A. Hodgdon. Translation-rotation paradox for diffusion in fragile glass-forming liquids. *Phys. Rev. E*, 50:2064–2068, 1994.
- G.R. Strobl. *The Physics of Polymers*. Springer, Berlin-Heidelberg, 2007.
- J.E. Tanner. Use of the Stimulated Echo in NMR diffusion studies. *J. Chem. Phys.*, 52: 2523–2526, 1970.
- A. Tardieu.  $\alpha$ -crystallin quaternary structure and interactive properties control eye lens transparency. *Int. J. Biol. Macromol.*, 22:211–217, 1998.
- D.G. Thomas. Transport characteristics of suspension: VIII.A note on the viscosity of Newtonian suspensions of uniform spherical particles. *J. Colloid Sci.*, 20:267–277, 1965.
- M.L. Tillett, M.A. Horsfield, L-Y. Lian, and T.J. Norwood. Protein-ligand interactions measured by  $^{15}\text{N}$ -filtered diffusion experiments. *J. Biomol. NMR*, 13:223–232, 1999.
- M. Tokuyama and I. Oppenheim. Dynamics of hard-sphere suspensions. *Phys. Rev. E*, 50: R16–R19, 1994.
- A.L. Tournier, J. Xu, and J.C. Smith. Translational hydration water dynamics drives the protein glass transition. *Biophys. J.*, 85:1871–1875, 2003.
- V. Tugarinov, Z. Liang, Y.E. Shapiro, J.H. Freed, and E. Meirovitch. A structural mode-coupling approach to  $^{15}\text{N}$  NMR relaxation in proteins. *J. Am. Chem. Soc.*, 123:3055–3063, 2001.
- A. van Blaaderen, J. Peetermans, G. Maret, and J.K.G Dhont. Long-time self-diffusion of spherical colloidal particles measured with fluorescence recovery after photobleaching. *J. Chem. Phys.*, 96:4591–4603, 1992.
- B. van den Berg, R.J. Ellis, and Ch.M. Dobson. Effects of macromolecular crowding on protein folding and aggregation. *EMBO J.*, 18:6927–6933, 1999. ISSN 0261-4189.
- A.S. Verkman. Solute and macromolecule diffusion in cellular aqueous compartments. *Trends Biochem. Sci.*, 27:27–33, 2002.
- D. Vitkup, D. Ringe, G.A. Petsko, and M. Karplus. Solvent mobility and the protein ‘glass’ transition. *Nat. Struct. Biol.*, 7:34–38, 2000.
- R. Walser, A.E. Mark, and W.F. van Gunsteren. On the validity of Stokes’ law at the molecular level. *Chem. Phys. Lett.*, 303:583–586, 1999.



- Y. Wang, C. Li, and G.J. Pielak. Effects of proteins on protein diffusion. *J. Am. Chem. Soc.*, 132:9392–9397, 2010.
- E.R. Weeks and D.A. Weitz. Subdiffusion and the cage effect studied near the colloidal glass transition. *Chem. Phys.*, 284:361–367, 2002a.
- E.R. Weeks and D.A. Weitz. Properties of cage rearrangements observed near the colloidal glass transition. *Phys. Rev. Lett.*, 89:095704, 2002b.
- C. Yigit, J. Heyda, and J. Dzubiella. Charged patchy particle models in explicit salt: Ion distributions, electrostatic potentials, and effective interactions. *J. Chem. Phys.*, 143:064904, 2015.
- F. Zhang, M.W.A. Skoda, R.M.J. Jacobs, S. Zorn, R.A. Martin, C.M. Martin, G.F. Clark, S. Weggler, A. Hildebrandt, O. Kohlbacher, and F. Schreiber. Reentrant condensation of proteins in solution induced by multivalent counterions. *Phys. Rev. Lett.*, 101:148101, 2008.
- F. Zhang, F. Roosen-Runge, A. Saute, M. Wolf, R.M.J. Jacobs, and F. Schreiber. Reentrant condensation, liquid–liquid phase separation and crystallization in protein solutions induced by multivalent metal ions. *Pure Appl. Chem.*, 86:191–202, 2014.
- H. Zhang and G. Nägele. Tracer-diffusion in binary colloidal hard-sphere suspensions. *J. Chem. Phys.*, 117:5908–5920, 2002.
- H.X. Zhou. Protein folding and binding in confined space and in crowded solutions. *J. Mol. Recognit.*, 17:368–375, 2004.
- H.X. Zhou, G.N. Rivas, and A.P. Minton. Macromolecular crowding and confinement: biochemical, biophysical, and potential physiological consequences. *Ann. Rev. Biophys.*, 37:375–397, 2008.
- S.B. Zimmerman and A.P. Minton. Macromolecular crowding: biochemical, biophysical, and physiological consequences. *Annu. Rev. Biophys. Biomol. Struct.*, 22:27–65, 1993.
- S.B. Zimmerman and S.O. Trach. Estimation of macromolecule concentrations and excluded volume effects for the cytoplasm of *Escherichia coli*. *J. Mol. Biol.*, 222:599–620, 1991.
- S. Zorilla, M.A. Hink, A. Visser, and M.P. Lillo. Translational and rotational motions of proteins in a protein crowded environment. *Biophys. Chem.*, 125:298–305, 2007.

## S.7. Curriculum vitae

### Personal data

---

Family name	Roos
First name	Matthias K.
Date of birth	March 27, 1988
Place of birth	Gelnhausen
Nationality	German

### Education

---

10/2012 – present	Doctoral studies Martin Luther University Halle-Wittenberg Institute of Physics (NMR group) Supervisor: Prof. Dr. Kay Saalwächter
10/2012 – 12/2012	including a graduate fellowship awarded by the SFB-TRR 102
10/2010 – 09/2012	Graduate studies in Physics (Master of Science) Martin Luther University Halle-Wittenberg Thesis: Nano-scale Roughness of Phase Boundaries in Heterogeneous Polymers As Studied by Spin-diffusion NMR
10/2007 – 09/2010	Undergraduate studies in Physics (Bachelor of Science) Martin Luther University Halle-Wittenberg Thesis: Untersuchung der lokalen und makroskopischen Spin-Diffusion in Flüssigkristallen und Polymeren
07/2007	University-entrance diploma Burg-Gymnasium Wettin, Wettin

Halle (Saale), September 2016

## List of publications (chronological order)

- [9\*] M. Roos, M. Ott, M. Hofmann, S. Link, E. Rössler, J. Balbach, A. Krushelnitsky, and K. Saalwächter. Coupling and decoupling of rotational and translational diffusion of proteins under crowding conditions. *J. Am. Chem. Soc.*, 138: 10365-10372 (2016)
- [8\*] M. Rothe, T. Gruber, S. Gröger, J. Balbach, K. Saalwächter, and M. Roos. Transient binding accounts for apparent violation of the generalized Stokes-Einstein relation in crowded protein solutions. *Phys. Chem. Chem. Phys.*, 18: 18006-18014 (2016)
- [7\*] M. Roos, M. Hofmann, S. Link, M. Ott, J. Balbach, E. Rössler, K. Saalwächter, and A. Krushelnitsky. The "long tail" of the protein tumbling correlation function: observation by <sup>1</sup>H NMR relaxometry in a wide frequency and concentration range. *J. Biomol. NMR*, 63: 403-415 (2015)
- [6] K. Schäler, M. Roos, P. Micke, Y. Golitsyn, A. Seidlitz, T. Thurn-Albrecht, H. Schneider, G. Hempel, and K. Saalwächter. Basic principles of static proton low-resolution spin diffusion NMR in nanophase-separated materials with mobility contrast. *Solid State Nucl. Magn. Reson.*, 72: 50-63 (2015)
- [5] M. Roos, P. Micke, K. Saalwächter, and G. Hempel. Moderate MAS enhances local <sup>1</sup>H spin exchange and spin diffusion. *J. Magn. Reson.*, 260: 28-37 (2015)
- [4\*] M. Roos, S. Link, J. Balbach, A. Krushelnitsky, and K. Saalwächter. NMR-detected Brownian dynamics of  $\alpha$ B-crystallin over a wide range of concentrations. *Biophys. J.*, 108: 98-106 (2015)
- [3] M. Roos, K. Schäler, A. Seidlitz, T. Thurn-Albrecht, and K. Saalwächter. NMR study of interphase structure in layered polymer morphologies with mobility contrast: disorder and confinement effects vs. dynamic heterogeneities. *Colloid Polym. Sci.*, 292: 1825-1839 (2014)
- [2] M. Roos, P. Micke, and G. Hempel. Monitoring nuclear spin-flip processes and measuring spin-diffusion constants via hole burning into the magnetization. *Chem. Phys. Lett.*, 536: 147-154 (2012)
- [1] W. Weissflog, U. Baumeister, M.-G. Tamba, G. Pelzl, H. Kresse, R. Friedemann, G. Hempel, R. Kurz, M. Roos, K. Merzweiler, A. Jáklí, C. Zhang, N. Diorio, R. Stannarius, A. Eremin, and U. Kornek. Unexpected liquid crystalline behaviour of three-ring bent-core mesogens: Bis(4-subst.-phenyl) 2-methyl-iso-phthalates. *Soft Matter*, 8: 2671-2685 (2012)

\* Publications marked with an asterisk are part of the thesis at hand.

## Scientific talks and posters (chronological order)

Listed are conferences and meetings at which a scientific talk or poster was presented by myself. Accounted for are only presentations relying on the results presented in this thesis.

- [10] SoftComp Annual Meeting 2016, June 6-9, 2016; Ancona, Italy (talk)
- [9] SFB-TRR 102 minisymposium, May 19<sup>th</sup>, 2016; Halle (Saale), Germany (talk)
- [8] Arbeitskreistreffen NMR 2016, March 11-14, 2016; Heldrungen, Germany (talk)
- [7] 80. Jahrestagung der DPG und DPG-Frühjahrstagung, March 6-11, 2016; Regensburg, Germany (talk)
- [6] 26<sup>th</sup> Faltertage, Oct. 23-25, 2015; Halle (Saale), Germany (talk)
- [5] 37<sup>th</sup> Annual Meeting of the GDCh Magnetic Resonance Division, Sept. 7-10, 2015; Darmstadt, Germany (talk)
- [4] 9<sup>th</sup> Conference on Fast Field Cycling NMR Relaxometry, July 27-30, 2015; Aberdeen, United Kingdom (talk)
- [3] 10<sup>th</sup> European Biophysics Congress (EBSA 2015), July 18-22, 2015; Dresden, Germany (poster)
- [2] Discussion Meeting on Aggregation and Clustering of Molecules, Oct. 30-31, 2014; Copenhagen, Denmark (poster)
- [1] Black Forest Focus 9 – Protein Dynamics: From Water Hydration to Crowding Effects, Sept. 25-29, 2013; Lenzkirch-Saig near Freiburg (Breisgau), Germany (poster)

... as well as several posters and talks presented at internal meetings of the cooperative research center SFB-TRR 102 and its graduate school

## **Statutory declaration – Eidesstattliche Erklärung**

Hiermit versichere ich, die vorliegende Arbeit selbständig und ohne fremde Hilfe verfasst und keine anderen als die von mir angegebenen Quellen und Hilfsmittel verwendet zu haben. Die den benutzten Werken wörtlich oder inhaltlich entnommenen Stellen habe ich als solche kenntlich gemacht.

Ich erkläre, keine anderweitigen Promotionsversuche unternommen und die vorliegende Dissertation weder in der jetzigen noch in einer anderen Fassung einer anderen wissenschaftlichen Einrichtung vorgelegt zu haben.

Halle (Saale), September 12, 2016

Matthias Roos

## Acknowledgment

At this point I would like to explicitly thank all the persons who supported me during the course of this research project, named in a *chronological order* as experienced *by myself* concerning the development of this thesis.

*I would like to declare my most sincere acknowledgment to*

PROF. KAY SAALWÄCHTER, for the opportunity to work on this very interesting project, and for his advice, support, and the insightful, motivating discussions. In addition I am thankful for the interesting conferences I was allowed to attend, and, in particular, for the mentoring and open and straightforward feedback;

DR. ALEXEY KRUSHELNITSKY, for the mentoring, support and advice, the introduction to NMR relaxometry, and, especially, for the fruitful and lively discussions, from which I benefited both scientifically and personally;

SUSANNE LINK, for providing the recombinantly expressed  $\alpha$ B-crystallin in milligram quantities which enabled the study of crowding effects on this protein, and for the general support in biophysical lab work;

PROF. JOCHEN BALBACH, for the scientific advice in bio-specific aspects and the support for the publications;

PROF. ERNST RÖSSLER, for the opportunity of performing the field-cycling NMR measurements, including insightful discussions on this field;

MARIUS HOFMANN, for having instructed me how to operate the field-cycling equipment and for having performed a long series of field-cycling NMR measurements on my behalf;

SEBASTIAN KIRSCHKE, for his work during his Bachelor's thesis that concerned the implementation and tests of the pulse sequence(s) for isotope-filtered PFG NMR;

DR. KLAUS SCHRÖTER and PROF. THOMAS THURN-ALBRECHT for the opportunity to perform the viscosity measurements using a rheometer, and for the technical support and discussions;

MAIK ROTHE, for his work during and also after his Master's thesis on the diffusion of protein mixtures;

DR. TOBIAS GRUBER, for providing the isotope-labeled SH3 domain and the discussions on this protein;

DR. MARIA OTT, for her support, help, the interesting and encouraging discussions, and the indispensable SAXS and FCS measurements;

FRANK LANGE, for his support with the PFG NMR equipment;

STEFAN GRÖGER, for the HSQC data and the pleasant working atmosphere;

SUSANNE MORGAN and DR. THOMAS MICHAEL, for administrative and organizational support within the framework of the SFB-TRR 102 and its integrated graduate school, particularly for the organization of soft-skills and scientific workshops that I willingly attended to broaden and improve my knowledge and soft-skills;

PROF. WOLFGANG BINDER, for being my mentor within the graduate school of the cooperative research center SFB-TRR 102;

and, last but not least, DR. GÜNTER HEMPEL, for the insightful and motivating discussions on NMR, beginning back to the time of my Bachelor's thesis.

Funding of this project is acknowledged to the cooperative research center SFB-TRR 102 that is funded by the GERMAN RESEARCH FOUNDATION (DEUTSCHE FORSCHUNGSGEMEINSCHAFT, DFG).

I'd also like to thank JEANETTE for her support "outside academia" – thanks for the great time, your help and the understanding.

For additional proofreading of the thesis I am indebted to MATTHIAS FISCHER and JOSHUA DAMRON. Thank you guys!

In this context, I finally would like to explicitly acknowledge the efforts, input, advices and refinements by KAY SAALWÄCHTER, ALEXEY KRUSHELNITSKY and all the CO-AUTHORS of the research articles in writing and publishing these papers.

\* \* \*