Development and applications of NMR relaxation methods to study protein and ligand dynamics.

Wernersson, Sven

2022

Document Version:
Publisher's PDF, also known as Version of record

Development and applications of NMR relaxation methods to study protein and ligand dynamics

Development and applications of NMR relaxation methods to study protein and ligand dynamics

by Sven Wernersson

LUND
UNIVERSITY

Thesis for the degree of Doctor of Philosophy
Thesis advisors: Professor Mikael Akke and Assistant Professor Pär Söderhjelm
Faculty opponent: Professor Vladislav Orekhov, Gothenburgh University

To be presented, with the permission of the Faculty of Engineering of Lund University in lecture hall KC:C at 13.00, 20 May, 2022

Organization	Document name
LUND UNIVERSITY Division of Biophysical Chemistry Box 124 SE–221 00 LUND Sweden	**DOCTORAL DISSERTATION** Date of disputation 2022-05-20 Sponsoring organization
Author(s) Sven Wernersson	

Title and subtitle

Development and applications of NMR relaxation methods to study protein and ligand dynamics

Abstract

This thesis involves the development and application of Nuclear Magnetic Resonance (NMR) relaxation experiments to investigate protein dynamics, a property that is critical to protein function. The dynamics of protein BRD4 is studied and found to indicate possible transient dimerization of the second bromodomain. Ways to directly infer the dynamics of protein ligands is evaluated using a combination of NMR relaxation experiments and molecular dynamics (MD) simulations. The impact of the presence of DMSO on protein – ligand interactions is studied using NMR relaxation dispersion and isothermal calorimetry (ITC), and DMSO is shown to have a significant impact. Ways to speed up NMR relaxation measurements using a combination of reduced dimensionality and non uniform sampling (NUS) is investigated, for both transverse auto relaxation rates and transverse, longitudinal cross correlated relaxation rates.

Key words

Nuclear Magnetic Resonance, NMR, NMR relaxation, Galectin-3, BRD4, Method development, Protein dynamics

Classification system and/or index terms (if any)

Supplementary bibliographical information	Language
	English
ISSN and key title	ISBN
	978-91-7422-882-3 (print) 978-91-7422-883-0 (pdf)
Recipient's notes	Number of pages: 194 / Price
	Security classification

Development and applications of NMR relaxation methods to study protein and ligand dynamics

by Sven Wernersson

LUND
UNIVERSITY

Cover illustration front: Quack? by Camille Garcia.

Faculty of Engineering, Division of Biophysical Chemistry

ISBN: 978-91-7422-882-3 (print)
ISBN: 978-91-7422-883-0 (pdf)

Printed in Sweden by Media-Tryck, Lund University, Lund 2022

Plans are worthless, but planning is everything.
Dwight D. Eisenhower

Contents

List of publications

This thesis is based on the following publications:

I **Bromodomain Interactions with Acetylated Histone 4 Peptides in the BRD4 Tandem Domain: Effects on Domain Dynamics and Internal Flexibility**
Sven Wernersson*, Romel Bobby*, Liz Flavell, Alexander G. Milbradt, Geoffrey A. Holdgate, Kevin J. Embrey, Mikael Akke
Manuscript

II **Conformational Dynamics of Designed Drug-like Ligands Bound to Proteins Determined by 19F NMR Relaxation and Molecular Dynamics Simulations**
Sven Wernersson, Olof Stenström, Simon Birgersson, Magdalena Riad, Ulrich Weininger, Göran Carlström, Ulf J. Nilsson, Mikael Akke
Manuscript

III **The cosolvent dimethylsulfoxide influences protein–ligand binding kinetics via solvent viscosity effects revealing the success rate of complex formation following diffusive protein–ligand encounter**
Sven Wernersson*, Simon Birgersson*, Mikael Akke
Manuscript

IV **Rapid measurement of heteronuclear transverse relaxation rates using non-uniformly sampled $R_{1\rho}$ accordion experiments.**
Sven Wernersson, Göran Carlström, Andreas Jakobsson, Mikael Akke
Magnetic Resonance, 2(2), 571-587, 2021. https://doi.org/10.5194/mr-2-571-2021

V **An accordion NUS experiment measuring transverse cross-correlated dipole/CSA relaxation**
Sven Wernersson, Göran Carlström, Andreas Jakobsson, Mikael Akke
Manuscript

*These authors contributed equally to the article.

All papers are reproduced with permission of their respective publishers.

List of publications not included in thesis

I **Entropy–Entropy Compensation between the Protein, Ligand, and Solvent Degrees of Freedom Fine-Tunes Affinity in Ligand Binding to Galectin-3C**
Johan Wallerstein, Vilhelm Ekberg, Majda Misini Ignjatović, Rohit Kumar, Octav Caldararu, Kristoffer Peterson, **Sven Wernersson**, Ulrika Brath, Hakon Leffler, Esko Oksanen, Derek T. Logan, Ulf J. Nilsson, Ulf Ryde, and Mikael Akke
JACS Au 1.4 (2021): 484-500

II **Backbone ^{1}H, ^{13}C, and ^{15}N resonance assignments of BoMan26A, a -mannanase of the glycoside hydrolase family 26 from the human gut bacterium Bacteroides ovatus**
Sven Wernersson, Viktoria Bågenholm, Cecilia Persson, Santosh Kumar Upadhyay, Henrik Stålbrand and Mikael Akke
Biomolecular NMR assignments 13.1 (2019): 213-218

III **Structural model of dodecameric heat-shock protein Hsp21: Flexible N-terminal arms interact with client proteins while C-terminal tails maintain the dodecamer and chaperone activity**
Gudrun Rutsdottir, Johan Härmark, Yoran Weide, Hans Hebert, Morten I. Rasmussen, **Sven Wernersson**, Michal Respondek, Mikael Akke, Peter Højrup, Philip J.B. Koeck, Christopher A.G. Söderberg, Cecilia Emanuelsson
Journal of Biological Chemistry 292.19 (2017): 8103-8121

My contribution to the papers

Paper I: Bromodomain Interactions with Acetylated Histone 4 Peptides in the BRD4 Tandem Domain: Effects on Domain Dynamics and Internal Flexibility

I did the analysis of all data with the help of MA. I significantly contributed to writing the manuscript and created most of the figures.

Paper II: Conformational Dynamics of Designed Drug-like Ligands Bound to Proteins Determined by 19F NMR Relaxation and Molecular Dynamics Simulations

I processed and analysed all flourine relaxation data, performed the analysis of the MD trajectories and the model-free fitting. I contributed significantly to the writing of the paper, and created several figures.

Paper III: The cosolvent dimethylsulfoxide influences protein–ligand binding kinetics via solvent viscosity effects revealing the success rate of complex formation following diffusive protein–ligand encounter

I performed and analysed all NMR experiments. I contributed significantly to writing the paper and created several figures.

Paper IV: Rapid measurement of heteronuclear transverse relaxation rates using non-uniformly sampled $R_{1\rho}$ accordion experiments.

I helped design the experiments with MA and GC. I ran and analysed all experiments with the help of GC. I wrote the first draft of the paper and created most of the figures.

Paper V: An accordion NUS experiment measuring transverse cross-correlated dipole/CSA relaxation

I designed the experiments. I ran and analysed all experiments with the help of GC. I wrote the first draft of the paper and created most of the figures.

Acknowledgements

I would like to give a big thanks to my supervisor **Mikael** for giving me the opportunity to do an internship back in 2015, and the journey since. I would also like to thank my co-supervisor **Pär** for his genuine interest and support. **Uli** – I couldn't not imagine better help for anyone starting in a lab then what you provided. **Göran**, you have been extremely helpful with everything related to practical NMR.

To the rest of the my group: **Olof**, **Johan**, **Kristine**, **Kristofer**. Thank you for great discussions and company! A special thanks to **David Fushman**, for allowing me to visit your lab several times. It was a pleasure. There are many others at BPC and CMPS that provided good company: **Bhakat**, **Filip**, **Mads** and **Magnus** deserves special mention! Also, thank you **Pathfinder gang** (you know who you are)! It has been far more fun than I could have imagined.

My **parents** and my **brothers**, thank you for always being there.

Camille – Life has not been the same since you came to Sweden. I love you.

Popular summary

This thesis involves the study of proteins. These molecules exist in many variants, and perform most of our bodily functions. This includes signaling (for example the presence of capsaicin, which makes food spicy), molecular transport (including oxygen), catalysis (including food digestion) and part of our body structure (such as muscles). So it is perhaps not too strange that an overwhelming fraction of the medicines used and developed today target different proteins. The building blocks of proteins are 20 different amino acids. These amino acids can be linked in any conceivable combination and length to chains. Each protein is defined by their unique chain of amino acids. During the creation of each chain, it spontaneously folds due to the electrochemical forces that are always present between amino acids. The result of this folding process is called the protein structure, which help proteins to perform their tasks. To fully understand protein function, knowledge of the protein structure needs to be complemented with knowledge of protein dynamics, the inherent flexibility, and movements of the structure. The dynamic properties of proteins is the focus of this thesis, and the pharmaceutical industry its likeliest area of application.

In the thesis I studied three different proteins. Galectin-3 is a protein that binds to sugar molecules, which can be found both inside and outside of cells. Problems with this protein are connected to a variety of diseases including cancer, fibrosis, and various forms of dementia. The protein BRD4 helps to control protein production in cells. This gives it a large impact on cancer progression and several types of cardiovascular diseases. When it comes to Ubiquitin, one of its main purposes is to signal which proteins have served their purpose and are to be degraded and recycled. In the thesis I study Ubiquitin only indirectly, as I use it as a model system given the wealth of previously published data to compare my results with.

The main experimental method that I used is called Nuclear Magnetic Resonance spectroscopy (NMR), which share the same underlying technology as Magnetic Resonance Imaging (MRI) used in hospitals. My work also utilises results from other experimental methods, as different methods often complement each other. Important methods include Isothermal Calorimetry (ITC), which allows for measurements of the energy difference when two molecules bind to each other. Molecular Dynamics (MD) simulations, are a form of computer calculations that studies protein dynamics and interactions in atomistic detail. X-ray crystallography, uses high energy X-ray beams to discover protein structures.

A total of five articles are presented in this thesis. The topic of the first article is the protein BRD4. It is a very large protein, with an amino acid chain of up to 1362 residues long. BRD4 is considered to have a large potential as a target for cancer treatments, but its length makes it hard to study. In the article we look at the impact on the protein dynamics from restricting studies of BRD4 to protein fragments of different lengths. The second article is a study on the interaction between Galectin-3 and a small molecule (similar to a pharmaceutical) from the perspective of the small molecule. This is accomplished by directly measuring on the fluorine atoms located on the small molecule. The third article is a study of the impact on interaction between Galectin-3 and a small molecule in the presence of Dimethyl Sulfoxide (DMSO). The largest effect of DMSO is an increased viscosity, which makes the interaction slower. The fourth and fifth articles presents a method for faster measurements of protein dynamics using NMR, which is especially useful for unstable and low-concentration samples.

My hope is that this thesis work helps us gain an improved understanding on how different parts of proteins affect each other and on how proteins interact with small molecules (especially regarding BRD4 and Galectin-3). In addition, I hope that this work has provided future scientists additional tools for continued studies of proteins and their dynamics.

Populärvetenskaplig sammanfattning

I denna avhandling studeras proteiner. Dessa molekyler finns i många varianter, och tillsammans utför de stora delar av kroppens funktioner. Detta inkluderar olika former av signalering (till exempel närvaro av kapsaicin som ger mat hetta), molekyltransport (som till exempel syre), katalys (av till exempel matnedbrytning) och delar av kroppens struktur (som till exempel muskler). Det är därmed kanske inte så konstigt att de flesta läkemedel som används och är under utveckling idag är utformade för att påverka proteiner. Proteiners byggstenar är 20 olika aminosyror. Dessa molekyler kan sammanlänkas i varje tänkbar kombination och längd till kedjor. Varje protein definieras av sin unika kedja. Efter att varje kedja skapats så börjar den spontant att vikas ihop, på grund av de elektrokemiska krafter som är närvarande mellan aminosyror. Resultatet av vikningen kallas för proteinets struktur, vilken hjälper proteinet att uppfylla dess uppgifter. För att fullt förstå proteiners funktion så behöver man utöver strukturen även studera dess dynamik, d.v.s. flexibiliteten och rörligheten i strukturen. Det är proteiners dynamiska egenskaper som är i fokus för denna avhandling, där läkemedelsutveckling är den troligaste tillämpningen av min forskning.

I avhandlingen studerar jag tre olika proteiner. Galektin-3 är ett protein som binder till sockermolekyler, vilka återfinns både inuti och utanpå celler. Problem med detta protein är kopplat till sjukdomar som cancer, fibros och diverse demenssjukdomar. Proteinet BRD4 hjälper till att styra kroppens produktion av proteiner. Detta protein har stor påverkan vid cancer och flera former av hjärt - och kärlsjukdomar. Proteinet Ubiquitin används bland annat för att signalera vilka proteiner som slutfört sina uppgifter och därmed ska brytas ner för att ge material till andra proteiner. Detta protein studerar jag indirekt, främst använder jag det för metodutveckling då det finns tidigare data att jämföra mina resultat med.

Den huvudsakliga metod jag använt för att studera proteindynamik kallas kärnmagnetisk resonansspektroskopi (NMR, Nuclear Magnetic Resonance), och är i grunden samma teknik som används för magnetröntgenkameror på sjukhus. Jag använder även resultat från andra mätmetoder, då de ofta kompletterar NMR spektroskopi. Här bör nämnas isotermisk titreringskalorimetri (ITC, Isothermal Titration Calorimetry), vilket används för att mäta energin-förändringen vid inbindningen av två molekyler till varandra. Molekyldynamik-simuleringar (MD), är en form av datorberäkningar som i hög detaljnivå simulerar proteiner och deras interaktioner. Röntgenkristallografi utreder molekylers struktur med hjälp av röntgen-strålar.

I avhandlingen presenteras fem artiklar. Den första artikeln handlar om proteinet BRD4 som nämndes ovan. Det är ett väldigt stort protein, med en kedja upp till 1362 sammanlänkade aminosyror. BRD4 har stor potential för cancer-behandling, men dess längd gör det svårt att studera. I artikeln studerar vi konsekvenserna för proteindynamiken av att begränsa sig till proteinfragment av olika längd. Den andra artikeln är en studie om interaktionen mellan Galektin-3 och en småmolekyl (som liknar läkemedel), ur småmolekylens perspektiv. Detta genom att mäta direkt på floratomerna som sitter på småmolekylen. Den tredje artikeln studerar interaktionen mellan Galektin-3 och en småmolekyl, i närvaro av olika mängder DMSO (dimetylsulfoxid). Största effekten av DMSO är en högre viskositet, som gör protein-småmolekyl interaktionen långsammare. Den fjärde och femte artikeln presenterar en metod för att snabbare mäta proteindynamik med hjälp av NMR, vilket är viktigt för instabila och lågkoncentrerade prover.

Förhoppningen är att detta arbete hjälper oss få en utökad förståelse av hur olika delar av proteiner påverkar varandra och hur proteiner interagerar med småmolekyler (särskilt för BRD4 och Galektin-3). Det är även min förhoppning att detta arbete gett framtida forskare mer verktyg för fortsatta studier av proteiner och deras dynamik.

CHAPTER 1

Introduction

This is a thesis about methods in Nuclear Magnetic Resonance (NMR) spectroscopy. The focus is on describing NMR theory, different NMR methods and in the articles the application of all the above. At the same time, this is a thesis on the study of proteins. Studying and improving our understanding of proteins and their interactions is the ultimate goal. Therefore, to start we need to look a bit at proteins and why they are important. This will be exemplified by a deeper look of the proteins Galectin-3 and BRD4, which have been studied in articles presented in the thesis. Finally, the last section describes the layout of the thesis.

1.1 Proteins and their importance

Proteins perform a multitude of activities for essentially all living beings [1]. This includes signaling (such as degradation marker Ubiquitin [2] and the capsaicin receptor [3]), molecule transport (including the oxygen carrier Hemoglobin [4] and the water transporter Aquaporin [5]), catalysis (such as the digestive enzyme pepsin [6]) and body structure (including the muscle protein Titin [7]).

In a 2011 review, representatives from the Bayer Healthcare Company defined drug targets as "*A protein, peptide or nucleic acid with activity that can be modulated by a drug, which can consist of a small molecular*

weight chemical compound (SMOL) or a biologic (BIOL), such as an antibody or a recombinant protein" [8]. In addition, the authors also stress the need for detailed knowledge about the function of the drug target and its interaction with the proposed drug. Given the many important functions provided by proteins it is perhaps no wonder that they are dominating as drug targets in the pharmaceutical industry [9, 10]. Another class of drug targets are nucleic acids (DNA and RNA) [11], which we historically have lacked detailed knowledge about [12]. Research is ongoing on nucleic acids as drug targets [13], including by NMR spectroscopy methods [14, 15].

An introduction to protein structure and function can be found in most basic biochemistry textbooks, such as *Introduction to Protein Structure* by Branden and Tooze [16]. Proteins are made from combinations of 20 different types of amino acids, bound to each other in sequence, like a chain. Each amino acid has a nitrogen amide, an aliphatic carbon (C_α), and a carboxyl group bound together in that order. These repeated segments are what forms the protein backbone. Also linked to each C_α is a side-chain group, which differs depending on the amino acid. Amino acid chains self-interact to form local (secondary) structures, the three most common of which are highlighted on the protein Ubiquitin in figure 1.1. These are: α-helices (red), β-sheets (blue), and loops (green). The sum of these secondary structure elements form the overall (tertiary) structure of the protein. The tertiary structure is sometimes divided into domains, defined by structural or functional properties. In some cases, several proteins can assemble together into so called quaternary structures. The structure of proteins can be determined using X-ray crystallography (see chapter 5), NMR spectroscopy [17], and more recently using the emerging techniques of cryogenic Electron Microscopy (cryo-EM) [18] and predictive methods [19].

A class of proteins called Intrinsically Disordered Proteins (IDPs) does not always conform to a single or indeed any secondary structure, as they interchange between states of varying disorder [20]. This flexibility enables IDPs to interact with different targets [21]. The overall importance of protein flexibility or dynamics [22] is increasingly recognised [23]. Protein dynamics involve both fast localized fluctuations [24] and slower conformational changes [25], and is critical for many protein functions [26], including gene expression [27], molecular recognition [28] and allosteric regulation [29]. Protein dynamics can be studied using NMR spectroscopy, and many experiments have been developed for this purpose [30–32]. Methods for studying protein dynamics using NMR spectroscopy is the aim of this thesis work.

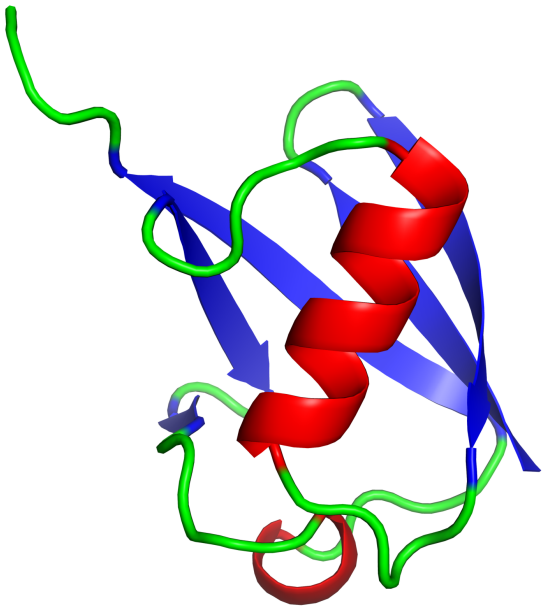


Figure 1.1: X-ray crystal structure of Ubiquitin (pdb: 1UBQ). Secondary structure elements are color-coded as α-helices (red), β-sheets (blue), and loops (green). The image was rendered using PyMOL [33].

Protein function will be exemplified in the following subsections, which describe the three proteins studied in this thesis, namely Galectin-3, Bromodomain-containing protein 4 (BRD4), and Ubiquitin.

1.1.1 Galectin-3

Galectins are a family of sugar binding proteins defined by their common Carbohydrate Recognition Domain (CRD), containing an evolutionary conserved binding site. The overall structure of the CRD is a concave form of the structural motif called a β-sandwich, with the binding site on the surface of the focal point (see figure 1.2) [34]. The main binding target

Figure 1.2: Neutron crystal structure of the Galectin-3 carbohydrate binding domain (GAL3C) (pdb: 6F2Q). The carbohydrate binding site is located at the focal point of the concave β-sheet in front of the figure. The image is rendered using the PyMOL software package [33].

of galectins are sugars on the surface of glycosylated (sugar modified) proteins and lipids [35], so called glycoproteins and glycolipids [34]. In humans, galectins have been found to be involved with many functions, including "autophagy, cell migration, immune response, inflammation, intracellular transporting, and signaling" [36]. This broad range of functions is related to that different cell types in the body expresses different forms of glycosylation, in turn leading to different galectin function. Galectins are involved in several diseases including cancer, fibrosis and heart disease [37].

A total of 15 galectins have been found in mammals [35] since their first discovery in 1975 [38]. They differ in whether they contain 1 or 2 CRDs, with the only galectin to have any other domain being galectin-3, which in addition to a single CRD contains an intrinsically disordered domain [35]. Galectins are synthesised in the cytosol of cells [34], but can be transported to the outside of the cell via a so far poorly understood mechanism [36]. Galectin-3 (Gal-3) is expressed throughout all cell types in the body [39]. While typically a monomer, it can oligomerise via ligand interactions or by itself at relatively high concentrations ($\sim$100 mM), a process that is mediated via its intrinsically disordered region [34]. Inside cells, Gal-3

participates in the regulation of the growth and death (apoptosis) of the cell [37]. Gal-3 and galectin-4 also regulate intravesicular (between cell compartments) glycoprotein traffic [34]. Outside the cell, Gal-3 has been noted to accumulate around ruptured bacterial cells, which is believed to signal autophagosomes, part of the body's degradation system [36].

Gal-3 has been indicated as a contributor to the symptoms of Alzheimer's, Parkinson's and Huntington's disease, where it triggers an inflammatory response in the brain [40]. In addition, it has shown to be involved in fibrosis [40] and several cancers, including "bladder cancer, colon cancer, dendritic, macrophages, melanoma, ovarian cancer, stromal, and syncytiotrophoblast cells" [36]. In this thesis, Gal-3 - ligand interactions are studied in papers II and III. Furthermore, it is used as a model system in paper IV.

1.1.2 Bromodomain-containing protein 4

Bromodomain-containing protein 4 (BRD4) function is connected to histones, the proteins responsible for packing DNA into chromosomes. DNA is wound around repeated octamers of histones, forming a complex called *chromatin* [1]. The local shape of chromatin helps control gene expression, as the DNA of genes active in transcription, replication, and repair are more exposed to the cell [41]. This is in part regulated via epigenetic modifications to chromatin (epigenetic signaling [41]), that is to say "heritable changes in the pattern of post-translational modifications" [42]. This includes methylation of the DNA itself and different modifications to the histone proteins, *e.g.* histone (de)acetylation via Histone Acetyl-Transferases (HATs) and Histone DeACetylases (HDACs) [42]. Issues with epigenetic signalling is a known contributor to many cases of cancer and inflammation [43].

BRD4 belongs to the Bromodomain and Extra-Terminal (BET) protein family, which read epigenetic modifications on chromatin and can signal chromatin- modifying enzymes. In addition to BRD4, the BET protein family contains BRD2, BRD3 and BRDT [41]. Specifically, BET-members read acetylated lysine sites on histones (and some other proteins including NF-κB and the PTEFb complex [44]) via their bromodomains. The bromodomain is a conserved sequence of around 110 amino acids, structured as a bundle of four α-helices connected via loop regions [43] (see figure 1.3). BRD4 contains two Bromodomains (BD1, BD2), and has several other features including its eponymous extra terminal domain, several clustered phosphorylation sites and a C-terminal domain along its 1362 amino acid

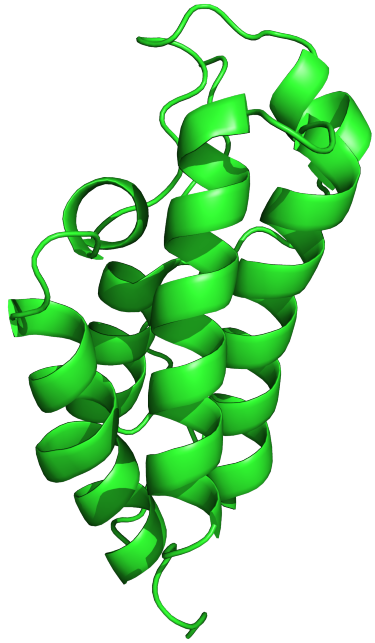


Figure 1.3: X-ray crystal structure of the BromoDomain 1 (BD1) of BRomoDomain-containing protein 4 (BRD4) (pdb: 4CLB). The image is rendered using the PyMOL software package [33].

long sequence. In addition to its epigenetic reader function, BRD4 is involved in several protein-protein interactions and has both kinase and HAT activity [44]. BRD4 is an pharmaceutical target of interest, as it has been linked to several cancers, inflammatory diseases and viral infections [42].

BRD4 is involved throughout the cell cycle, partially as a 'memory', helping the cell to remember previous expression cycles. It has also been found to specifically regulate genes involved in cell differentiation and cell cycle control [44]. The PTEFb complex mentioned above is an important component of DNA transcription, as it helps promote RNA polymerase activity. BRD4 has been found to interact with PTEFb both via Bromodomain 2 (BD2) and the C-terminal domain. This interaction has been found to be central for initiating transcription after division of cell nuclei [41]. While an incomplete list, the functions mentioned here exemplify some of the many links BRD4 has with different forms of cancer, including breast and brain cancer, leukaemia and cancer metastasis [44]. For more information, see for example White *et. al.* [43] and Hajmirza *et. al.* [41]. A central role of BRD4 in inflammation is the regulation of Nuclear Factor-κB (NF-κB), a transcription factor that in turn is the master regulator of cellular inflammatory

response. BRD4 binds to NF-κB after the latter has been acetylated on lysine position 310. Inhibition of BRD4 has been found to reduce transcription of pro-inflammatory genes [42]. The interactions between BRD4 and NF-κB has also been linked to heart failure [45]. For more information about the role of BRD4 in cardiovascular diseases including atherosclerosis, pulmonary arterial hypertension, and essential hypertension, see Lin *et al.* [45]. BRD4 and its ligand interactions are studied in paper I, which also provides additional biological context.

1.1.3 Ubiquitin

Ubiquitin (see figure 1.1) is used as a model system in paper V, and is not directly studied in this thesis. Briefly, as the name implies Ubiquitin is found throughout all cells, where it is involved in protein degradation [1] and stem cell differentiation [46]. It is a relatively small protein containing all three major secondary structure elements, as highlighted in figure 1.1.

1.2 Layout of the thesis

NMR spectroscopy includes many methods for different purposes such as imaging [47], structure determination [48], dynamics measurements [30], metabolic studies [49,50], and *in vivo* studies of proteins [51,52]. The focus of the next three chapters is the specific sub-field of this thesis, namely protein dynamics measurements using NMR relaxation experiments. Chapter 2 is a basic statistical mechanics description of nuclear spins. Following that, chapter 3 focuses on NMR relaxation theory. Finally chapter 4 goes through the actual NMR relaxation experiments that I have performed.

Chapter 5 describes other important methods that appears in the papers. Generally I have not performed any experiments based on those methods, but I utilised results created from them via collaborators. These methods are solid state NMR spectroscopy, Isothermal Calorimetry (ITC), X-ray crystallography, and Molecular Dynamics (MD) simulations. Chapter 6 is a guide to the articles presented in the thesis. After the guide to the papers comes the the papers themselves. The intent is that the proceeding sections will have made the articles at least somewhat understandable.

CHAPTER 2

Basic NMR

This chapter gives a brief introduction to Nuclear Magnetic Resonance (NMR) and NMR Spectroscopy in solution. This is by necessity an incomplete overview, which assumes basic knowledge of statistical and quantum mechanics, and omits many fundamental concepts in NMR such as the secular approximation [53, 54]. Given the basic and general nature of this chapter specific citations are usually not given. The majority of this and the following chapter are based on textbooks [55–57].

2.1 The Nuclear Spin

Atomic nuclei with a non-zero spin quantum number possess an intrinsic nuclear magnetic moment and nuclear spin. This is determined by the combination of protons and neutrons in the nucleus. This thesis work is focused on so called spin-1/2 nuclei (notably ^{1}H, ^{15}N and ^{13}C), which have spin quantum numbers of [+1/2, -1/2]. This is in contrast to integer spin nuclei, which won't be discussed here. The protein in samples used for NMR measurements are typically prepared ('labelled') to contain the nuclear isotopes of interest. This thesis work almost exclusively involves measurements on proteins with nearly uniform ^{1}H, ^{15}N labelling. NMR spectroscopy involves placing atomic nuclei in a static magnetic field B_0, which makes the nuclear spins and B_0 interact with an energy described by defining a Hamiltonian operator $\mathscr{H}$ in a right-handed coordinate system

(XYZ), with the direction of B_0 along the Z-axis (eq. 3.12 in [55]):

$$\mathscr{H} = -\gamma B_0 \hat{I}_z \tag{2.1}$$

where $\hat{I}_z$ is the longitudinal component of the spin angular momentum vector. The spin quantum numbers correspond to two equal (*degenerate*) but distinct energy levels for the nuclear spin. But when we put molecules in the external magnetic field, spins pointing along the field will have slightly lower energy. This is called the *Zeeman splitting* of the energy levels, and it is this energy difference that makes NMR measurements possible. Inside this external magnetic field, the populations (p_r) of the nuclear spin state (r) will follow the Boltzmann distribution:

$$p_r = \frac{e^{-\hbar\omega_r/k_B T}}{\sum_s e^{-\hbar\omega_s/k_B T}} \tag{2.2}$$

where $\hbar$ is the reduced Planck constant ($6.63 \cdot 10^{-34}\mathrm{m}^2 \cdot \mathrm{kg} \cdot s^{-1}$), k_B is the Boltzmann constant ($1.38 \cdot 10^{-23}\mathrm{m}^2 \cdot \mathrm{kg} \cdot \mathrm{s}^{-2} \cdot \mathrm{K}^{-1}$), and the index s goes over all quantum spin states (here, $\pm 1/2$). The energy levels (ω_s) are defined in frequency units as $\omega_s = \pm 1/2 \gamma B_0$, where γ is the gyromagnetic ratio between the magnetic moment and angular momentum of the nucleus. For protons ($\gamma \approx 268 \cdot 10^6$ rad$\cdot$s$^{-1}\cdot$T^{-1}) in a static field of 11.7 T this corresponds to energy difference $|\hbar\omega_r| \approx 3 \cdot 10^{-25}$ J [56]. This small energy difference, induced by the static magnetic field is what enables NMR measurements. It also makes NMR spectroscopy an inherently insensitive method, requiring measurements on the net or bulk nuclear spin magnetisation summed up over all molecules in the sample of interest. Given the low sensitivity, a typical 500 μl NMR sample contains relatively high protein concentrations of around 0.5 mM $= 1.5 \cdot 10^{17}$ protein molecules. Finally, nuclear spins are not restricted to the two energy levels, but exist in a superposition of the two, represented by Dirac notation as (eq. 2.198 in [57]):

$$\langle \Psi | = c_\alpha^* \langle \alpha | + c_\beta^* \langle \beta | \tag{2.3}$$

$$| \Psi \rangle = c_\alpha | \alpha \rangle + c_\beta | \beta \rangle \tag{2.4}$$

where $|\Psi\rangle$ and $\langle\Psi|$ are the wavefunction (function describing the nuclear spin state) and its complex conjugate, c_α and c_β are numerical coefficients,

and (α, β) are shorthand for the spin states (+1/2,-1/2) respectively. Decomposing wavefunctions into sums in this manner is a general feature of quantum mechanics.

2.1.1 Larmor frequency and the chemical shift

A spin 1/2 nuclei experiencing a static external magnetic field (B_0) will precess around this field, with its larmor frequency,

$$\omega_0 = -\gamma B_0 \tag{2.5}$$

This is equivalent to the energy of the interaction between nuclei and static field, as defined by equation 2.1. This precession will generate a Radio Frequency (RF) current in the NMR spectrometer, which is recorded. The current is generated via electromagnetic induction, a phenomena commonly used for stove tops and wind turbines. However, all nuclei experience various degrees of shielding from B_0 due to the electrons surrounding them. This shielding depends on the type and orientation of nearby nuclei. This leads to a shift of larmor frequency, called the chemical shift. This shift is in principle orientation dependent, but it is averaged in solution, which leads to an isotropic shielding constant:

$$\sigma = (\sigma_{11} + \sigma_{22} + \sigma_{33})/3 \tag{2.6}$$

where $(\sigma_{11}, \sigma_{22}, \sigma_{33})$ denote the principal axes of the chemical shielding tensor in the Mehring convention [58]. The effective larmor frequency is:

$$\omega = -\gamma(1 - \sigma)B_0 \tag{2.7}$$

For solution NMR measurements, the anisotropy of this shielding (Chemical shift anisotropy, CSA) is still present in NMR relaxation processes, which will be discussed in chapter 3.

2.1.2 Multiple spin systems

Molecules studied via NMR typically include two or more chemically bonded NMR-active nuclei, such as the ^{15}N-^{1}H spin pairs studied in this thesis.

Such nuclei have a scalar (J) coupling between them. This represents the *through-bond* mediated interaction between the nuclei, that is to say the interaction mediated through the shared electrons. For 2 coupled spin I and S, this leads to four different functions to describe the pure spin states:

$$\begin{bmatrix} \Psi_1 & \Psi_2 \\ \Psi_3 & \Psi_4 \end{bmatrix} = \begin{bmatrix} |\alpha\alpha\rangle & \cos(\theta)\,|\alpha\beta\rangle + \sin(\theta)\,|\beta\alpha\rangle \\ \cos(\theta)\,|\beta\alpha\rangle + \sin(\theta)\,|\alpha\beta\rangle & |\beta\beta\rangle \end{bmatrix}$$

Where (α, β) describes spin $\pm 1/2$ for spin I, S respectively. The angle θ is defined by:

$$\theta = \arctan\left(\frac{2\pi J_{IS}}{\omega_I - \omega_S}\right)/2 \tag{2.8}$$

Where J_{IS} is the scalar coupling constant and ω_I, ω_S are the larmor frequencies of spins I and S. It is clear that Ψ_2 and Ψ_3 are independent of each other when $\omega_I - \omega_S \gg 2\pi J_{IS}$, called the weak coupling condition. This is assumed to be valid throughout this thesis. These four spin states have different chemical shifts, which means that the NMR signals from spin I and S will be split into multiplets. This has led to the widespread use of *decoupling* elements during NMR experiments. Such elements are essentially a mixture of RF pulses designed to average the scalar couplings to zero, creating less cluttered spectra with merged peaks. Assuming a weak coupling, for an isolated spin pair of chemically bonded nuclei such as (^{1}H,^{15}N), the Hamiltonian operator for the interaction with B_0 can be described as (eq. 7.14 in [55]):

$$\mathscr{H} = \Omega_1 \hat{I}_z + \Omega_2 \hat{S}_z + 2\pi J_{12} \hat{I}_z \hat{S}_z \tag{2.9}$$

2.2 Bulk properties - Expectation values and the Spin Density Operator

According to the Copenhagen interpretation of quantum mechanics, when we make a measurement of the superposition of spin energy levels, this superposition will collapse to one of the energy levels. While we cannot predict the outcome from a single measurement of a nuclear spin, we can calculate the average *expectation value* of many measurements. For an

operator Q (such as the Hamiltonian) described by a single wavefunction, a general expectation value <Q> can be defined as (eqs. 2.38-39 in [57]):

$$< Q >= \langle \Psi | \, Q \, | \Psi \rangle = \sum_{nm} c_m^* c_n \, \langle m | \, Q \, | n \rangle \tag{2.10}$$

Where the second equality is a way to decompose any wavefunction, as was done in equations 2.3 and 2.4 above. For an NMR sample, it is impractical to describe the complete system as a single wavefunction, given the high degree of complexity involved. Instead, the description of the sample is divided into subensembles of molecules individually interacting with the surroundings (lattice). The expectation values of the system are then described as averages over the subensembles, (eq. 2.45 in [57]):

$$< \overline{Q} >= \langle \Psi | \, Q \, | \Psi \rangle = \sum_{nm} \overline{c_m^* c_n} \, \langle m | \, Q \, | n \rangle \tag{2.11}$$

Note that the overbar is over the system coefficient terms $\sum_{nm} \overline{c_m^* c_n}$. These coefficients determine the difference between and evolution of the different subsystems, and form the basis of calculations in NMR spectroscopy. These calculations are typically performed using the matrix of averaged coefficients called the *density matrix*, which in turn is used to define the *density operator*, σ (eq. 2.46 in [57]):

$$\overline{c_m^* c_n} = \overline{\langle m | \, Q \, | n \rangle} = \langle m | \, \sigma \, | n \rangle = \sigma_{nm} \tag{2.12}$$

Specifically, the diagonal elements of the density matrix denotes populations of different spin states, and the off-diagonal elements represents coherent super positions of these populations, called coherences. The difference in spin quantum number between the constituent populations of a coherence is called the coherence order. The spin density operator allows the average of expectation value Q to be calculated as (eq. 2.46 in [57]):

$$< \overline{Q} >= \mathrm{Tr}\{\sigma Q\} \tag{2.13}$$

2.3 Manipulating The Spin Density Operator

The time evolution of the spin density operator can be calculated using the Liouville-von Neumann equation, defined as (eq. 2.53 in [57]):

$$\frac{d\sigma(t)}{dt} = -i\left[\mathscr{H}, \sigma\right] \tag{2.14}$$

Where the Hamiltonian $\mathscr{H}$ is assumed identical for all sub-ensembles, and the straight brackets denote a commutator operation. Typical NMR experiments consists of RF-pulses and delay periods, neither of which are time-independent. Ignoring the time-dependence of delay periods (the focus of next chapter), the spin evolution is typically calculated in a rotating frame of reference, by the following transformation (eqs. 2.59, 2.66 in [57]):

$$\sigma^r(t) = U\sigma U^{-1} \tag{2.15}$$

$$\mathscr{H}_e = U\mathscr{H}U^1 - iU\frac{dU^{-1}}{dt} \tag{2.16}$$

Where U is an unitary operator, *i.e.* fulfills the condition $U^*U = UU^* = E$, where E is the identity matrix. This is in many ways a natural transformation, when assuming that the rotating frame is has the same frequency as the RF-receiver. By convention, the rotating frame is defined as right-handed. This allows a general time-independent solution (eq. 2.67 in [57]) to the Liouville-von Neumann equation:

$$\sigma^r(t) = \exp\{-i\mathscr{H}t\}\sigma^r(0)\exp\{i\mathscr{H}t\} \tag{2.17}$$

From this point onward, the spin Hamiltonian and density operator will be defined in the rotating frame unless otherwise stated. The transformation of density operator σ^1 into σ^2 under the Hamiltonian $\mathscr{H}$ can then be represented as (eq. 2.207 in [57]):

$$\sigma^1 \xrightarrow{\mathscr{H}t} \sigma^2 \tag{2.18}$$

which is equivalent to (eq. 2.208 in [57]):

$$\sigma^2 = e^{-i\mathscr{H}t}\sigma^1 e^{i\mathscr{H}t} \tag{2.19}$$

2.4 Product Operators

The matrix form of the Dirac representation for the wavefunctions of spin states (α, β) is:

$$\langle\alpha| = \begin{bmatrix}1 & 0\end{bmatrix}, |\alpha\rangle = \begin{bmatrix}1\\0\end{bmatrix} \tag{2.20}$$

$$\langle\beta| = \begin{bmatrix}0 & 1\end{bmatrix}, |\beta\rangle = \begin{bmatrix}0\\1\end{bmatrix} \tag{2.21}$$

Similarly, product spin operators corresponding to the unique axes of the angular momentum vector in the rotating frame can be represented by the Pauli spin matrices (eq. 2.71 in [57]):

$$I_x = \frac{1}{2}\begin{bmatrix}0 & 1\\1 & 0\end{bmatrix}, \quad I_y = \frac{1}{2}\begin{bmatrix}0 & -i\\i & 0\end{bmatrix}, \quad I_z = \frac{1}{2}\begin{bmatrix}1 & 0\\0 & -1\end{bmatrix} \tag{2.22}$$

For multispin systems, these representations are made from the tensor products of single spins, for example:

$$|\alpha\beta\rangle = \begin{bmatrix}1\\0\end{bmatrix} \otimes \begin{bmatrix}0\\1\end{bmatrix} = \begin{bmatrix}0\\1\\0\\0\end{bmatrix} \tag{2.23}$$

$$I_{x,IS} = I_{x,I} \otimes E = \frac{1}{2}\begin{bmatrix}0 & 0 & 1 & 0\\0 & 0 & 0 & 1\\1 & 0 & 0 & 0\\0 & 0 & 1 & 0\end{bmatrix} \tag{2.24}$$

$$2I_xS_x = I_x \otimes S_x = \frac{1}{2}\begin{bmatrix}0 & 0 & 0 & 1\\0 & 0 & 1 & 0\\0 & 1 & 0 & 0\\1 & 0 & 0 & 0\end{bmatrix} \tag{2.25}$$

Where the I, IS subscript in eq. 2.24 denotes 1-spin and 2-spin systems, E is the identity matrix, and the 2 in eq. 2.25 is a normalisation factor. The spin Hamiltonian can be expressed in terms of these angular momentum operators, as noted by the definition in equation 2.1. The same is true for the density operator, which can be represented by orthogonal basis operators as (eq. 2.201 in [57]):

$$\sigma(t) = \sum_{k=1}^{K} b_k(t)\hat{\mathbf{B}}_k \tag{2.26}$$

Where K is the dimensionality of the spin system, proportional to the number of spins in the studied ensemble, and $b_k(t)$ are complex coefficients. This allows the evolution of the spin density operator defined by equation 2.19 to be calculated in terms of angular momentum operators. Further, if a given set of basis operators commute with each other on the format $[A, B] = iC$, then:

$$e^{-i\theta C} A e^{-i\theta C} = A\cos(\theta) + B\sin(\theta) \tag{2.27}$$

This condition holds true in the weak scalar coupling limit, and is the basis of the *product operator approach* [59], which will be demonstrated in the next section.

2.5 Measuring NMR

In this section, two simple 1D NMR experiments will be demonstrated. The pulse acquire experiment shows basic concepts for an isolated spin, and the spin echo experiment demonstrates properties of scalar coupled nuclei. But to understand these experiments, some necessary definitions are first needed.

All RF-pulses have a duration ('Pulse Width', φ), phase and power level. From the pulse width we can define the angular frequency of the pulse [60]:

$$\omega_1 = \frac{2\pi}{\varphi(2\pi)} \tag{2.28}$$

where $\varphi(2\pi)$ is the width of a RF-pulse at a given power level that leads to a full 2π rotation of the spin magnetisation. What we measure in NMR spectroscopy is the frequency offset Ω_I between the receiver and the nuclear spin. From this we can define a reduced static magnetic field $\Delta B_0 = -\Omega_I/\gamma$. When an RF-pulse is generated with frequency ω_1 close to the resonance frequency ω_I of the nuclear spin of interest, this spin will nutate around an effective field ω_{eff} equal to the vector sum of the reduced static field and the RF-pulse field ($\sqrt{\omega_1^2 + \Omega^2}$).

This shows that the applied RF-pulse needs to be large relative to the spin offset for the effective field to approximate the applied rf-field. The rf-pulses used in NMR spectroscopy are monochromatic in nature (single frequency), but acts as poly-chromatic pulses. This is due to the fundamental Heisenberg uncertainty relationship between the frequency and the inverse of the pulse length [60]. This effective field is tilted with an angle $\theta = \arctan\left(\frac{\omega_1}{\Omega}\right)$ away from the static magnetic field. For a typical ^{1}H-spectrum of a spectrometer operating at 600MHz ^{1}H-frequency, we have a maximum offset of $\approx$ 7 ppm = 4200 Hz $\Leftrightarrow$ 26389 rad/s. With a typical pulse width of $PW_{360} = 8 \cdot 4 = 32\mu$s, we have a pulse frequency of 31250 Hz $\Leftrightarrow$ 196350 rad/s. This means that signals at worst experience a field with an angle of 82° away from the static field, as opposed to 90° on-resonance. However, such signals also experience a stronger rf field strength ($(196350 + 26389)/196350 \approx 13\%$ greater), which partially compensates for this issue [60].

2.5.1 Pulse acquire

Pulse acquire refers to an NMR experiment consisting of three elements, A RF-pulse, an acquisition period, and a magnetic recovery period. The last of these is only needed if the experiment is repeated for an increase in the signal-to-noise ratio. For a pulse with the length equal to a rotation angle of $\pi/2$, we can represent the rotation as:

$$I_z \xrightarrow{(\tau_p \omega_1 I_y)} \cos(\tau_p \omega_1) I_z + \sin(\omega) I_x \xrightarrow{(\tau_p \omega_1 = \frac{\pi}{2})} I_x \tag{2.29}$$

where the rotation angle is defined by the product $\tau_p \omega_1$. Following the pulse, the spin magnetisation starts to precess with the larmor frequency. We then measure the projection of the magnetisation vector on the RF-receiver as the *free induction decay* (FID), described in section 2.6 below.

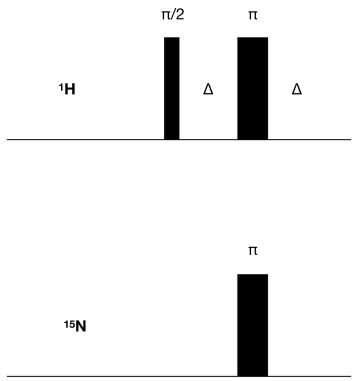


Figure 2.1: A basic NMR experiment incorporating a spin echo element. RF-pulses are shown as black rectangles, with the rotation angle noted above each pulse. The Δ delay periods are of length $1/4J_{HN}$.

For repeated experiments, we need to wait for the spins to return to equilibrium along B_0 before the next measurement. Finally, the pulse length should be calibrated for each type of nuclei and sample to know the corresponding rotation angle. Since what we measure is the projection of the transverse magnetisation, we can do this by following the signal intensity as a function of the pulse length.

2.5.2 The spin echo

A simple example of many concepts from this chapter is the heteronuclear spin echo. An experiment utilizing this concept is shown in figure 2.1. The experiment starts with a $\pi/2$ ^{1}H pulse to generate transverse magnetisation. This magnetisation evolves freely under delay Δ, is then inverted by a π-pulse and evolve for another Δ-delay. The effect of this sequence is the evolution of pure ^{1}H magnetisation into a mixed state of ^{1}H-^{15}N magnetisation. This is typically the starting point of most conventional multi-dimensional NMR experiments. The full product operator calculation for this sequence, for an $^1\mathrm{H}-^{15}\mathrm{N}$ spin pair is:

$$H_z \xrightarrow{\frac{\pi}{2}(H_y)} H_x \xrightarrow{\Omega\Delta(H_z)} \cos(\Omega\Delta)H_x + \sin(\Omega\Delta)H_y$$

$$\xrightarrow{2\pi J\Delta(2H_zN_z)}$$

$$\cos(\Omega\Delta)\cos(\pi J\Delta)H_x + \cos(\Omega\Delta)\sin(\pi J\Delta)2H_yN_z$$

$$+\sin(\Omega\Delta)\cos(\pi J\Delta)H_y - \sin(\Omega\Delta)\sin(\pi J\Delta)2H_xN_z$$

$$\xrightarrow{\pi(H_x+N_x)}$$

$$\cos(\Omega\Delta)\cos(\pi J\Delta)H_x + \cos(\Omega\Delta)\sin(\pi J\Delta)2H_yN_z$$
$$-\sin(\Omega\Delta)\cos(\pi J\Delta)H_y + \sin(\Omega\Delta)\sin(\pi J\Delta)2H_xN_z$$

$$\xrightarrow{\Omega\Delta(H_z)}$$

$$\cos(\Omega\Delta)\cos(\pi J\Delta)\cos(\Omega\Delta)H_x + \cos(\Omega\Delta)\cos(\pi J\Delta)\sin(\Omega\Delta)H_y$$
$$-\sin(\Omega\Delta)\cos(\pi J\Delta)\cos(\Omega\Delta)H_y + \sin(\Omega\Delta)\cos(\pi J\Delta)\sin(\Omega\Delta)H_x$$
$$+\sin(\Omega\Delta)\sin(\pi J\Delta)\sin(\Omega\Delta)2H_yN_z - \cos(\Omega\Delta)\sin(\pi J\Delta)\sin(\Omega\Delta)2H_yN_z$$
$$+\sin(\Omega\Delta)\sin(\pi J\Delta)\cos(\Omega\Delta)2H_xN_z + \sin(\Omega\Delta)\sin(\pi J\Delta)\sin(\Omega\Delta)2H_yN_z$$
$$= \cos(\pi J\Delta)H_x + \sin(\pi J\Delta)2H_yN_z$$

$$\xrightarrow{2\pi J\Delta(2H_zN_z)}$$

$$\cos(\pi J\Delta)\cos(\pi J\Delta)H_x + \cos(\pi J\Delta)\sin(\pi J\Delta)2H_yN_z$$
$$+\sin(\pi J\Delta)\cos(\pi J\Delta)2H_yN_z - \sin(\pi J\Delta)\sin(\pi J\Delta)H_x$$
$$= \cos(2\pi J\Delta)H_x + \sin(2\pi J\Delta)2H_yN_z \rightarrow \left[\Delta = \tfrac{1}{4J}\right] \rightarrow 2H_yN_z$$

2.6 The NMR signal

The measured FID or NMR signal from transverse single spin, single quantum magnetisation such as that represented by the I_x operator can be described as an complex, oscillating wave on the form (eq. 5.4 in [55]):

$$S(t) = S_0 e^{i\Omega t} e^{-R_2 t} \tag{2.30}$$

Where S_0 is the initial signal amplitude, Ω is the carrier offset (the frequency difference between the larmor frequency and the RF-receiver). The signal amplitude decreases exponentially in time, as the bulk magnetisation loses coherence and 'spreads out'. The rate of this process is determined by the transverse relaxation rate constant R_2 (properly introduced in the next chapter).

2.6.1 NMR signal processing

NMR signals are typically Fourier transformed to create a frequency spectrum. Signals of the form described by equation 2.30 generates a spectrum with peaks that have a real and an imaginary part with different lineshapes on the form (eq. 5.5 in [55]):

$$S(\omega) = \frac{S_0 R}{R^2 + (\omega - \Omega)^2} + i\frac{-S_0(\omega - \Omega)}{R^2 + (\omega - \Omega)^2} \tag{2.31}$$

The peak is centered at the offset Ω, with the amplitude $S(\omega)$ at a given frequency ω. The real part generates a well-phased (absorptive) Lorentzian lineshape, while the imaginary peak amplitude will be positive for frequencies higher than Ω, and negative for frequencies below. This means that the imaginary component generates a dephased (dispersive) Lorentzian lineshape. Typically, only the real absorptive peak is displayed, as it allows for more high resolution spectra. Also, at the center of the peak ($\omega = \Omega$), the height of the absorptive peak equals $1/R_2$, meaning that for a given signal intensity (peak volume), a signal with a higher R_2 rate will generate a broader peak.

2.6.2 NMR signals of multiple spin systems

Multiple spin systems will be represented in this section by the the previously discussed isolated ^{1}H-^{15}N spin pair. Only single quantum magnetisation generates observable NMR signals [56]. Operators such as I_x are referred to as in-phase magnetisation, and generates signals on the form described by equation 2.30, with the caveat that in the absence of decoupling they are split with distances equal to the J-coupling.

This in contrast to operators on the form $I_x S_z$, which generate signals with different signs for the split signals [56], referred to as anti-phase magnetisation. Interchange between in-phase and anti-phase magnetisation is mediated via the scalar coupling and cross-correlated relaxation, which is of importance for many applications mentioned throughout this thesis.

CHAPTER 3

NMR Relaxation Theory

For all NMR measurements the nuclear spin magnetisation is aligned transverse to B_0 by RF-pulses. After this point, the magnitude of the measurable magnetisation will decay over time. This is due to NMR relaxation, typically divided into two categories; (1) the realignment of nuclear spin magnetisation to a Boltzmann-distributed equilibrium along B_0, (2) The loss of coherence in the prepared magnetisation state. NMR relaxation is driven by fluctuations in the local magnetic fields surrounding a nuclear spin due to molecular motion. The two categories of relaxation are differentiated as; (1) spin-lattice relaxation, interchange between the nuclear spin and the surrounding solution (lattice), (2) spin-spin relaxation, interchange between local nuclear spins.

The simplest quantitative description of NMR relaxation comes from the Bloch equations [61]. Essentially, this assumes exponential relaxation driven by the phenomenological relaxation constants R_1, R_2. These correspond to the relaxation categories described above. It is therefore natural that the R_2 relaxation rate increases the linewidth of NMR signals, as it involves the loss of magnetisation coherence. Both R_1, R_2 describe auto-relaxation of a given nuclear spin. In addition, there are cross-correlated and cross-relaxation mechanisms that requires us to move beyond the Bloch equations. This will also allow us to couple the NMR relaxation rates to their underlying molecular motions in terms of spectral density distributions. This is mediated via different coupling mechanisms, most notably the dipolar and CSA couplings.

3.1 Bloch, Wangsness, Redfield NMR Relaxation Theory

There are several rigorous descriptions of NMR relaxation. Two of the more widely applied theories are Average Liouvillian theory [62] and Bloch, Wangsness, Redfield (BWR) theory [63, 64], presented here. This description is mainly based on chapter 5 of [57]. BWR is a semi-classical perturbation theory, meaning that the lattice is assumed to follow the laws of classical mechanics, and the Hamiltonian describing the spin - lattice interaction can be decomposed as (eq. 5.47 in [57]):

$$\mathscr{H}(t) = \mathscr{H}_0 + \mathscr{H}_1(t) \tag{3.1}$$

where $\mathscr{H}_0$ is a quantum mechanical Hamiltonian acting on the spin system, and $\mathscr{H}_1(t)$ describes a stochastic, time-dependent perturbation that couples the spin system to the lattice. The time-dependent Hamiltonian $\mathscr{H}_1(t)$ can be decomposed as (eq. 5.75 in [57]):

$$\mathscr{H}_1(t) = \sum_m \sum_{q=-k}^{k} (-1)^q F_{mk}^{-q}(t) \hat{\mathbf{A}}_{mk}^{q} \tag{3.2}$$

where F_k^{-q} are random time dependent spatial functions, A_k^q are tensor spin operators, q denotes coherence order, k spin tensor rank, and m different relaxation interactions. The tensor spin operator can in turn be defined by basis operators (generalisation of eq. 5.54 in [57]),

$$A_{mk}^q = \sum_p A_{mkp}^q \tag{3.3}$$

where the index p denotes degenerate spin operators with the same rank q. These basis operators need to fulfil the condition $[\mathscr{H}_0, A_{kp}^q] = \omega_p^q A_{kp}^q$, which is fulfilled by the standard Cartesian basis operators defined in the previous chapter. Under assumptions partially listed below, this allows the derivation of the following equation describing the time dependence of the density operator under the effects of NMR relaxation (eq. 5.71 in [57]):

$$db_r(t)/dt = \sum_s (-i\Omega_{rs} b_s(t) - \Gamma_{rs}[b_s(t) - b_{s0}]) \tag{3.4}$$

Where the $b_j(t)$ term is the spin density ($\sigma(t)$) operator of basis operator $\hat{\mathbf{B}}_r$, transformed into an *interaction frame*, defined by (eq. 5.74 in [57]):

$$b_j(t) = \langle \hat{\mathbf{B}}_j | \sigma(t) \rangle / \langle \hat{\mathbf{B}}_j | \hat{\mathbf{B}}_j \rangle \tag{3.5}$$

While not transparent from this expression, the purpose of this interaction frame is to remove the explicit dependence on the time-independent σ_0 (and equivalently $\mathscr{H}_0$) term of the spin density operator. This is analogous to the rotating frame transformation discussed in the last chapter. The term Ω_{rs} defines a characteristic frequency (eq. 5.72 in [57]):

$$\Omega_{rs} = \langle\ \hat{\mathbf{B}}_r | [\mathscr{H}_0, \hat{\mathbf{B}}_s] \rangle / \langle \hat{\mathbf{B}}_r | \hat{\mathbf{B}}_r \rangle \tag{3.6}$$

The relaxation rate constant Γ_{rs} for the relaxation between the operators $\hat{\mathbf{B}}_r$, $\hat{\mathbf{B}}_s$ is defined by a sum of different contributions (eqs. 5.76, 5.90 in [57]):

$$\begin{aligned}
\Gamma_{rs} =& \sum_m \Gamma_{rs}^m + \sum_{\substack{m,n \\ m \neq n}} \Gamma_{rs}^{mn} \\
=& \frac{1}{2} \sum_m \sum_q \sum_p \Big(\langle \hat{\mathbf{B}}_r | [\hat{\mathbf{A}}_{mkp}^{-q}, [\hat{\mathbf{A}}_{mkp}^{q}, \hat{\mathbf{B}}_s]] \rangle \Big/ \langle \hat{\mathbf{B}}_r | \hat{\mathbf{B}}_r \rangle \Big) (-1)^q J(\omega) \\
+& \frac{1}{2} \sum_{\substack{m,n \\ m \neq n}} \sum_q \sum_p \Big(\langle \hat{\mathbf{B}}_r | [\hat{\mathbf{A}}_{mkp}^{-q}, [\hat{\mathbf{A}}_{nkp}^{q}, \hat{\mathbf{B}}_s]] \rangle \Big/ \langle \hat{\mathbf{B}}_r | \hat{\mathbf{B}}_r \rangle \Big) (-1)^q J_{mn}(\omega)
\end{aligned} \tag{3.7}$$

where $J(\omega)$ and $J_{mn}(\omega)$ are the auto- and cross-correlated power spectral density functions, the subject of the next section. The spectral density includes the random spatial functions ($F(t)$) defined by equation 3.2. In general for the sum in equation 3.7, the case $\hat{\mathbf{B}}_r = \hat{\mathbf{B}}_s$ describes auto-relaxation, $\hat{\mathbf{B}}_r \neq \hat{\mathbf{B}}_s$ for a given relaxation mechanism m describes cross-relaxation between operators, and terms with sums over relaxation mechanisms (m, n) describes cross-correlated relaxation between them.

Three notes on the underlying assumptions of equation 3.4; Firstly, the ensemble ("sample") average of the transformed time dependent Hamiltonian is zero. Secondly, it is assumed that we can only observe NMR phenomena in the range $\tau_c \leq t \leq 1/\Gamma_{rs}$, where in isotropic solutions such as water τ_c is

the overall rotational tumbling time. Third, the spectral density function described here is valid for isotropic solutions (water) in the classical limit.

3.2 Spectral Density

The spectral density function $J(\omega)$ describes the contribution of re-orientational molecular motions to NMR relaxation, and is defined as the real Fourier transform of the correlation function $C(\tau)$ (eq. 5.92, 5.93 in [57]):

$$J(\omega) = Re\left\{\int_{-\infty}^{\infty} C(\tau)e^{-i\omega\tau}d\tau\right\} \tag{3.8}$$

$$C(\tau) = \overline{c_0(t)c_0(t+\tau)Y_2^0[\Omega(t)]Y_2^0[\Omega(t+\tau)]} \tag{3.9}$$

where $C(\tau)$ describes stochastic diffusive processes, described by the previously mentioned stochastic spatial functions (F_2^0) decomposed into second order spherical harmonic functions, $Y_2^0[\Omega(t)]$ (eq. 5.91 in [57]):

$$F_2^0(t) = c_0(t)Y_2^0[\Omega(t)] \tag{3.10}$$

where $\Omega(t) = (\theta(t), \phi(t))$ are polar angles defining the principal direction of different spin interactions in relation to B_0. As discussed in section 3.3, these directions differ for the main interactions in spin 1/2 nuclei, namely dipole-dipole and CSA. Likewise, the $c_0(t)$ coupling terms also differ between interactions. These terms are generally constant on a NMR relaxation time-scale, which allows them to be factored out from the spectral density as (eq. 5.94 in [57]):

$$j(\omega) = c_{00}J(\omega) \tag{3.11}$$

where c_{00} is a time-independent coupling term. The spectral density function quantifies molecular motion at different frequencies. The measurable relaxation rates sample the spectral density function at discrete frequencies, which are related to the transition probabilities between the different spin quantum states defined in the last chapter [56]. The molecular motions at these frequencies modulate the spin transitions, inducing relaxation [55].

3.2.1 Model free Modelling

The spectral density function (SDF) contains information on molecular motion on a biologically important time scale, which makes it of high interest to model it from NMR relaxation data [24]. This is often done using the Model free (MF) modelling approach, which was independently created by Wennerström & Halle [65], and Lipari & Szabo [66, 67]. This section is in large parts based on an excellent MF-review by Halle [68]. MF models are based on a decoupling assumption and a symmetry assumption. These can take different forms, depending on the specific MF model used.

The decoupling assumption refers to the decoupling of global and internal molecular motions. In water based solvents, global motions generally refer to rotational diffusion. The decoupling assumption can be invoked in two different ways; First, motions can be assumed to be time-scale separated. That is, local motions are averaged out on the global motion timescale, and vice versa. Secondly, the motions can be assumed to be statistically independent. Time-scale separation implies statistical independence, but not the other way around. Here, local motions can take place on the same time-scale as global motions, assuming they are separate. An example is chemical exchange processes, which will be discussed below. These decoupling assumptions are generally assumed to hold for global, folded proteins, but not for intrinsically disordered proteins (IDP), for which there is a hydrodynamic coupling between local motion and global rotational diffusion. The effect of the respective decoupling approximations on the correlation function $C(\tau)$ are the following:

$$C(t) = \begin{cases} (1) & C_I(t)C_G(t) \\ (2) & C_I(t) + C_G(t) \end{cases} \tag{3.12}$$

where $C_I(t)$ is the internal correlation function, and $C_G(t)$ is the global correlation function. Different symmetry approximations can be applied to either of these correlation functions. An assumption applied in paper II of this thesis is that of isotropic global motion. Another asymmetry assumption that is often made on the internal motion is that it is isotropic (unidirectional), coupled with an assumption that the internal correlation function exhibit exponential decay. When applying the assumptions of statistical independence and isotropic global motion, the correlation function can be described as [68]:

$$C(t) = e^{-t/\tau_m}\left(S^2 + (1 - S^2) \cdot C_{int}(t)\right) \tag{3.13}$$

where S^2 is the order parameter, defined as $S^2 = C_{int}(\infty)$ and τ_m is the correlation time of the global process. The correlation function is normalised, which means that the order parameter is bounded on the interval [0 1]. If applying the additional assumptions made by Lipari & Szabo [66, 67] off an exponentially decaying internal correlation function and isotropic internal motion, the Fourier transform of $C(t)$ leads to the spectral density function [24]:

$$j(\omega) = \frac{2}{5}\left(\frac{S^2\tau_m}{1 + (\omega\tau_m)^2} + \frac{(1 - S^2)\tau}{1 + (\omega\tau)^2}\right) \tag{3.14}$$

where τ is the internal correlation time. From these expressions, the meaning of the term model-free modelling becomes more clear. In essence, the spectral density function is divided into the three components $C_G(\tau)$, $C_I(\tau)$ and S^2, with no assumption made on the molecular mechanisms underlying these parameters. Instead, such insight needs to be gained from the project specific context. In particular, the division of internal motion into $C_I(\tau)$ and S^2 is often used, as these parameters respectively provide kinetic and equilibrium information. The magnitude of the order parameter is sometimes interpreted as the restriction on the angular fluctuation of a given bond vector, such as NH. This can exemplified by the Gaussian axial fluctuation model [69], where the fluctuation is assumed to follow a Gaussian potential, yielding the approximate order parameter:

$$S^2 \approx 1 - 3\sin^2(\theta)\sigma_{phi}^2 \tag{3.15}$$

where the angle θ describes the 'cone' the bond vector diffuses around.

3.3 Spin Interactions

Following is a description of the two main relaxation mechanisms present in ^{15}N nuclei of ^{1}H-^{15}N spin pairs, namely dipole-dipole and CSA relaxation. These interactions are averaged to zero over different directions as the molecule tumbles in the solution [70]. However, the fluctuation of the

interaction is still present, and drives relaxation [56]. For both relaxation mechanisms the relevant relaxation coupling constant will be defined, which is the c_{00} term in the spectral density expression of equation 3.11. In addition to dipolar and CSA relaxation, certain spin systems can experience quadrupolar and paramagnetic relaxation (see for example chapter 5 in [57]).

3.3.1 Dipole-Dipole (DD) relaxation

Dipole relaxation is generated by the fluctuations in the magnetic field of nearby nuclei due to molecular motion (see section 5.4.1 in [57]). The dipolar coupling constant d, can be written as (eq 5.116 in [57]):

$$d = (\mu_0/4\pi)\hbar\gamma_H\gamma_N r_{NH}^{-3} \tag{3.16}$$

where μ_0 is the vacuum permeability constant, $\hbar$ is the reduced plank constant, (γ_H, γ_N) are the gyromagnetic ratios of hydrogen and nitrogen respectively, and r_{NH} is the effective NH chemical bond length. The effective bond length in the dipolar coupling is longer than what is measured in neutron scattering experiments, due to inherent *zero-point* fluctuations in the energy potential of chemical bond lengths. That is, fluctuations independent of thermal energy [71]. For the backbone NH-bond length (r_{NH}), neutron scattering measurements give $r_{NH} \approx 1.02$ Å, while NMR measurements and quantum mechanical calculations give $r_{NH} \approx 1.04$ Å [71,72]. Both values are used in analysis of NMR relaxation data. For model-free modelling this merely scales the estimated S^2-parameters. This makes the exact value used relatively unimportant for comparative studies such as paper I of this thesis [71].

3.3.2 Chemical shift Anisotropy (CSA) relaxation

The distribution of electrons surrounding nuclei in molecules is not symmetric, which gives rise to anisotropic shielding of the nuclei, i.e. the shielding depend on the molecular orientation relative to the static magnetic field. This induces NMR relaxation as molecular motion modify the CSA orientation in relation to the static magnetic field. For ^{15}N, the CSA interaction tensor is typically assumed to be axially symmetric. This allows the CSA coupling constant to be defined as (table 5.8 in [57]):

$$c = (\Delta\sigma\gamma_N B_0)/\sqrt{(}3) \tag{3.17}$$

where $\Delta\sigma$ is the CSA parameter, $\Delta\sigma = \sigma_{11} - (\sigma_{22} + \sigma_{33})/2$. The CSA can be measured using either solid state NMR or via solution NMR relaxation measurements (see for example [73, 74]). It can be noted that the CSA shielding exhibit equivalent zero-point fluctuations as the chemical bond mentioned in 3.3.1 above. However, for the nitrogen CSA in NH-spin pairs these fluctuations have a very small impact, and are often ignored [75].

3.4 Auto-relaxation rates

The Bloch equations treats NMR relaxation as mono-exponential decay with the phenomenological rate constants R_1, R_2. In addition to describing additional relaxation pathways, more rigorous treatments of NMR relaxation allow R_1, R_2 to be described in terms of spin couplings and spectral densities. Solving equation 3.7 for the dipolar coupling contributions to the auto-relaxation rates R_1, R_2 of the ^{15}N spin of an isolated ^{1}H-^{15}N spin pair gives the result [24]:

$$R_1^{dd} = \frac{1}{4}d^2\left[j(\omega_H - \omega_N) + 3j(\omega_N) + 6j(\omega_H + \omega_N)\right] \tag{3.18}$$

$$R_2^{dd} = \frac{1}{8}d^2\left[j(0) + j(\omega_H - \omega_N) + 3j(\omega_N) + 6j(\omega_H) + 6j(\omega_H + \omega_N)\right] \tag{3.19}$$

where d^2 is the squared dipolar coupling constant. The corresponding relaxation contributions from the CSA coupling is [24]:

$$R_1^{csa} = c^2 j(\omega_N) \tag{3.20}$$

$$R_2^{csa} = \frac{1}{6}c^2\left[4j(0) + 3j(\omega_N)\right] \tag{3.21}$$

where c^2 is the squared CSA coupling constant. The total relaxation rate constant is the sum of the contributions from the dipolar and CSA coupled relaxation. The frequencies of the spectral density that appear in these

relaxation rates are related to the spin transition probabilities. The higher frequency components of the spectral density appear for dipolar relaxation, as it involves an interaction between a *pair* of dipoles, while the CSA interacts with the static magnetic field [76] (see for example chapter 20 in [56]). Finally, we can note that equation 3.4 is a sum over different operators. For non-isolated spin pairs this manifests as as spin network, which will be further discussed in the context of homonuclear cross-relaxation in section 3.7.

3.5 Chemical Exchange

Proteins interchange between different inherent structural conformations and interaction complexes [25], such as the different protein - ligand interactions studied in papers I-III of this thesis. In NMR spectroscopy, such interchange is termed chemical exchange [25], which is measurable using NMR relaxation measurements, provided that the interchanging states have different chemical shifts [57], and the exchange occurs on an NMR accessible time-scale. For the simplest case, we have a two-state exchange process between states (M_1, M_2) with the exchange rates (k_1, k_{-1}):

$$M_1 \underset{k_{-1}}{\overset{k_1}{\rightleftarrows}} M_2 \tag{3.22}$$

In addition, we can define an average exchange rate $k_{ex} = 1/k_1 + k_{-1}$. Chemical exchange contributes a R_{ex}-term to the transverse auto-relaxation rate constant R_2, as the random jumps between states with different chemical shifts de-phase transverse magnetisation [77]. For backbone ^{1}H-^{15}N spin pairs, this exchange can be described using McConnell's extension to the Bloch equations to include rate constants [78]. For two site exchange, the time evolution of the transverse magnetisation for the two sites (M_1, M_2) in the absence of RF-pulses (free precession) is given by (eq. 5.180 in [57]):

$$\begin{bmatrix} M_1(t) \\ M_2(t) \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} M_1(0) \\ M_2(0) \end{bmatrix} \tag{3.23}$$

where the a_{jk}-coefficients depend on the frequency difference $\Delta\omega = \omega_2 - \omega_1$ between states, the relative populations p_1, p_2 of each state, and the

relaxation rate constants ($R^0_{2,1}$, $R^0_{2,2}$) of each state in the absence of chemical exchange. The effective R_2 rate (and the corresponding linewidth) and the effective signal offset depends on the ratio between k_{ex} and $\Delta\omega$. This ratio can be used to classify chemical exchange processes into different regimes [25]. The slow exchange regime is defined by $k_{ex} \ll \Delta\omega$, where $R_2 \approx R_{2,0}$, and the offset difference is large between the peaks of the exchanging states. The intermediate exchange regime is defined by $k_{ex} \approx \Delta\omega$, where the peaks coalesce into a single, highly broadened peak. For $k_{ex} \geq \Delta\omega$, this coalesced peak increasingly narrows, representing a population weighted average of the magnetisation of the interchanging states.

As will be discussed in chapter 4, the de-phased magnetisation can be refocused by CPMG and $R_{1\rho}$ based NMR relaxation experiments. These experiments introduce new controlled variables, namely the CPMG refocusing frequency ν_{cpmg} and the effective spin lock frequency ω_{SL} [25]. In addition, $\Delta\omega$ depends on the static magnetic field strength (B_0). By varying experiment dependent combinations of (ν_{cpmg}, ω_{SL}, B_0), $\Delta\omega$ and k_{ex} can be fitted. The frequency difference $\Delta\omega$ gives structural information about the exchanging states [25, 79], while k_{ex} provides kinetic information (as discussed in paper III).

3.6 Cross-correlated dipole/CSA relaxation

The dipolar and CSA spin interactions vary with B_0 in the same way, as they both depend on the relative orientation of the studied spin system. The correlation between these mechanisms generate cross-correlated relaxation. This relaxation mechanism adds to the overall relaxation of magnetisation with spin state +1/2, and subtracts from the relaxation of magnetisation with spin state -1/2 [76]. The relaxation rate constants of this cross-correlated relaxation mechanism can be calculated from the second term of equation 3.7 as [80]:

$$\eta_z = -\sqrt{3}cdP_2(\cos\beta)j(\omega_N) \tag{3.24}$$

$$\eta_{xy} = -\sqrt{3}cdP_2(\cos\beta)(4j(0) + 3j(\omega_N)) \tag{3.25}$$

where η_z and η_{xy} are the cross-correlated relaxation rate constants of longitudinal and transverse magnetisation respectively. $P_2(x) = (3x^2 - 1)/2$ is the second order Legendre polynomial, and β is the angle between the

principal axes of the dipolar and CSA tensors. The β-angle has been estimated to around 20° [73] for backbone amides, and should be included as a correction term in the auto relaxation rates R_1, R_2 at high static magnetic fields [81]. An interesting property of the transverse relaxation rate η_{xy} is that it is inherently chemical exchange (R_{ex}) free. This makes it useful for unambiguous estimates of the spectral density at zero frequency (J(0)) [80].

3.7 The Solomon equation and dipolar cross-relaxation

The longitudinal dipolar relaxation of interacting (non-decoupled) spins can be derived from the master equation (3.4 above). The result is equivalent to the commonly used Solomon equations [82], which extended for N interacting spins are (eq. 5.12, in [57]):

$$\frac{d\Delta I_{kz}(t)}{dt} = -\rho_k \Delta I_{kz}(t) - \sum_{j \neq k} \sigma_{kj} \Delta I_{jz}(t) \tag{3.26}$$

where $\Delta I_{kz}(t)$ refers to longitudinal magnetisation that has been disturbed from equilibria ($\Delta I_{kz}(t) = I_{kz}(t) - I^0_{kz}$ where I^0_{kz} is the equilibrium magnetisation state). The longitudinal auto-relaxation ρ is equivalent to R_1 in the Bloch formulation, where rate ρ_k of spin k is composed of contributions from all other interacting spins (eq. 5.13, in [57]):

$$\rho_k = \sum_{k \neq j} \rho_{kj} \tag{3.27}$$

In addition, there are cross-relaxation rate terms σ_{kj} between each interacting spin pair. For a given IS spin pair, σ can be expressed as a combination of the double and zero quantum transition probabilities (eq. 20.17 in [56]):

$$\sigma = \frac{1}{4} d^2 \left(6j(\omega_I + \omega_S) - j(\omega_I - \omega_S)\right) \tag{3.28}$$

For homonuclear spin systems, such as ^{1}H in protonated proteins, σ is dominated by $j(0)$-terms [83]. This affects ^{15}N spin state selective experiments in non-decoupled spectra, as those studied in paper V, as spin flips of the ^{1}H spin interchanges the spin state selective states [84]. This induces an

effect similar to chemical exchange, increasing the relaxation of each studied spin state by an amount equivalent to the longitudinal relaxation rate of ^{1}H [76]. In addition, it serves to average out the relaxation rates of the different spin states [84].

3.7.1 Dipolar cross-relaxation in isolated heteronuclear ^{1}H-^{15}N spin pairs

For nitrogen spin in an isolated ^{1}H-^{15}N spin pair, equation 3.26 can be written as (eq. 5.11 in [57]):

$$\frac{d\Delta N_z(t)}{dt} = -\rho\Delta N_z(t) - \sigma_{NH}\Delta H_z(t) \tag{3.29}$$

with the dipolar cross-relaxation rate σ_{NH} given by [24]:

$$\sigma_{NH} = \frac{1}{4}d^2\left(6j(\omega_H + \omega_N) - j(\omega_H - \omega_N)\right) \tag{3.30}$$

CHAPTER 4

Practical aspects of NMR relaxation

NMR experimentation is a broad and versatile area. This chapter is by necessity restricted to practical aspects of solution state ^{1}H-^{15}N NMR relaxation experiments, although many of the concepts are more general.

4.1 The NMR spectrometer

Following is a brief listing of the central components and functions of an NMR spectrometer. More details can be found for example in chapter 13 of *Understanding NMR spectroscopy* [55].

- The static field magnet. That is to say, the typically superconducting magnet generating the static magnetic field, B_0 (T).
- The probe-head ('probe'), a replaceable insert in the center of the magnet, containing the RF coils, a temperature regulation unit, and the shim coils ('shims'). The shims refer to multiple small, adjustable magnets used to calibrate the static magnetic field.
- Related to the shim coils is the deuterium signal lock. NMR samples typically contain a small amount of deuterated solvent (see section 4.2.5). By monitoring this signal, the NMR spectrometer can keep the magnetic field stable and adjust for drift and other disturbances during experiments from for example the RF pulses.

- The transmitter generates rf pulses and sends them to the probe. The frequency of these signals is called the carrier frequency, which is typically placed in the middle of the signal frequency range.

- The sample signal is amplified by amplifiers and acquired by the receiver, which includes an analogue to digital converter (ADC). As the NMR signal is stored digitally, it is acquired at discrete time-points and not continuously.

4.1.1 Quadrature Detection

NMR signal frequencies are typically defined by their offset ($\Omega = \omega_0 - \omega_{rf}$) to the carrier frequency (ω_{rf}). The sign of Ω is defined in a right handed coordinate system. Signals are recorded as the 1D projection on the receiver RF-coil (detector) in the probe. This generates a signal on the format $\cos(\Omega t)$, which does not suffice to distinguish the relative sign of the signal. The basis of quadrature detection is to split the signal to the receiver in two paths. Each split signal is mixed with a reference signal at the carrier frequency . The reference signals are 90° phase shifted relative to each other, which means that the two mixed signals will have cosine and sine modulation respectively.

The real Fourier transform of the cosine modulated signal gives two positive peaks at the frequencies $\pm\Omega$, while the sine modulated signal gives a positive peak with frequency $+\Omega$, and a negative peak at frequency -Ω. The difference of these data-sets gives a single peak at the correct frequency $+\Omega$. For more details, see chapter 3.2.3 of *Protein NMR spectroscopy* [57].

4.1.2 NMR pulse programming

NMR pulse sequences are implemented on the NMR spectrometer as scripts ('*pulse programs*') that are coded in manufacturer specific programming languages. Special pulses, for example adiabatic ramps discussed in section 4.6.3 are typically calculated and written as separate files which are read in by the pulse program during the experiment.

An important consideration when designing pulse programs is the sign of the pulse phases. In two articles, Malcolm Levitt and Ole Johannessen [85, 86] investigated how NMR spectrometers interpret pulse programs to generate the RF-pulses. They note that the generated RF pulses by design

can't have any meaningful sign. This contrasts to nuclear spins, as the gyro-magnetic ratio (γ) has a sign that defines the direction of nutation in the coordinate system. Specifically, for pulses with flip angles other than π, the magnetisation will be rotated in different directions depending on the sign of γ. They find additional related issues that have smaller practical implications, as they were found to cancel out.

4.2 NMR samples

This section describes five studies highlighting different aspects of NMR sample conditions and their importance in NMR experiments.

4.2.1 Buffer ion mobility

The overall sensitivity (Signal/Noise ratio, S/N) of an NMR spectrometer is proportional to [87]:

$$S/N \propto 1/\sqrt{T_c R_c + T_a(R_c + R_s) + T_s R_s} \tag{4.1}$$

Where T_c and R_c is the temperature and resistance of the radio frequency (rf) coils, T_a is the temperature of the pre-amplifier, and T_s and R_s is the temperature and resistance of the sample. The two factors that are controllable by the user are T_s and R_s. The choice of T_s is restricted by factors including protein stability, protein dynamics, and temperature of interest. The sample resistance is directly proportional to the concentration and mobility of any ionic buffer components. It is often possible to choose low mobility buffers to increase NMR sensitivity. For example, with everything else being equal, having HEPES buffer instead of triphosphate buffer can result in 2.3 times higher sensitivity [87].

4.2.2 Protein – Buffer interactions

Buffers can directly interact with and influence protein dynamics. Wong and co-workers have shown significant differences in millisecond (ms) protein dynamics for different buffers using ^{15}N NMR CPMG dispersion experiments [88]. For the three studied proteins, they saw relatively small chemical shift changes combined with significant site-specific changes to ms

dynamics. They note that while the studied proteins all contain charged sites that likely binds to buffer, in general protein-buffer behaviour is likely very protein specific and hard to predict [88].

4.2.3 Salt concentration and pH

Kukic and Co-workers [89] studied the effects of pH and salt concentration on Hen Egg White Lysozyme (HEWL). First, they studied the effect of ionic strength on the pKa of titratable groups, by tracking chemical shift changes. Two out of four tracked residues showed a significant effect. The difference in effect on the residues was suspected to be related to differences in solvent accessibility and hydrogen bonding patterns. Further, they looked at salt-induced chemical shift changes ($\Delta\delta_{solvent}$), decomposed as:

$$\Delta\delta_{solvent} = \Delta\delta_b + \Delta\delta_{el} + \Delta\delta_{vdw} + \Delta\delta_{hb} \tag{4.2}$$

where $\Delta\delta_b$ is the bulk magnetic suspectibility of the solvent, $\Delta\delta_{el}$ is screening of electrostatic forces, $\Delta\delta_{vdw}$ is changes in van der Waal forces, and $\Delta\delta_{hb}$ is changes in hydrogen bonding patterns. More specifically, $\Delta\delta_b$ refers to any protein wide average change in chemical shift with salt concentration, which in the case of HEWL was found for the $^1H^N$ nuclei. $\Delta\delta_{el}$ is a salt-screening effect, which was found for the titratable groups discussed above.

4.2.4 Viscosity

The impact of sample solution viscosity on protein dynamics was investigated by Wallerstein and Akke [90]. Many small molecular compounds (including many pharmaceuticals) are not water soluble, so in protein – ligand studies it is common to solubilise the ligand using a low volume fraction of dimethyl sulfoxide (DMSO) in the sample. This significantly increases the viscosity even at low fractions of DMSO (on the order of a few percent). They found that the impact of DMSO on the (^{1}H, ^{15}N) chemical shifts of Gal-3 was relatively small, which suggests a low degree of direct protein – DMSO interaction. The biggest impact was on the rotational correlation time, τ_c. As expected, τ_c scaled proportionally with viscosity. This has significance for comparison between different samples and studies, not the least for cases where different samples have been used to measure

different relaxation rates. The impact of DMSO on ms protein dynamics is the topic of paper III.

4.2.5 Deuterium exchange

A ubiquitous additive to protein solution NMR samples is deuterated H_2O (^2H_2O), as it is used for the lock signal mentioned in section 4.1. Just as normal water, these deuterons exchange with the protein amides. As ^1H and ^2H have different chemical shifts (the *'solvent induced isotope shift (SIIS)'*), this exchange generates an R_{ex} contribution to the transverse auto-relaxation rate constant, R_2. This phenomenon was investigated by Kumari and Co-workers [91], by measuring ^{15}N-CPMG dispersion experiments on the IDP α-synuclein and the globular PDZ2 domain of human phosphatase (hPTP1E).

They found that for both proteins, solvent exposed protein residues experience a significant increase in R_{ex} with ^2H_2O concentration. For the highly solvent exposed α-synuclein, this amounted to contributions up to around 15 s^{-1}. Since this effect is typically on an intermediate timescale, it can be refocused at relatively modest CPMG refocusing frequencies of around $\nu_{cpmg} = 500$ Hz. This makes it a very minor issue for standard measurements of R_2, but a potentially significant issue for dispersion experiments quantifying chemical exchange processes. They note that this issue can be minimised by either placing all D_2O in an isolated insert, or by minimising the added volume fraction of D_2O to ideally 1% or lower.

4.3 Measuring multi-dimensional NMR spectra

This section discusses basic aspects of two dimensional NMR experiments. The principles for NMR experiments involving additional dimensions are highly similar. See for example reference [57] for further details. The gyromagnetic ratio of hydrogen-1 (^1H) is $\gamma_h = 267.522$ (10^6 rad s^{-1} T^{-1}), nearly 10 times higher than that of nitrogen-15 (^{15}N, $\gamma_n = -27.106$ (10^6 rad s^{-1} T^{-1}) [56]. As the magnitude of the magnetisation vector is directly proportional to the gyromagnetic ratio, ^1H-^{15}N experiments typically start with ^1H magnetisation. This magnetisation is transferred via spin echoes (or other methods such as CRINEPT [92]) to the ^{15}N-nuclei. The FID (and possibly different relaxation rates) of this indirect dimension is sampled,

and the magnetisation is transferred back to ^{1}H for standard quadrature detection.

4.3.1 Sampling the FID of indirect dimensions

This section is a brief overview, for more details see for example chapter 8 of *Understanding NMR spectroscopy* [55]. The indirect dimension is not directly acquired. Instead, following the transfer of magnetisation to the ^{15}N-nuclei, there is a period of free precession. This is called the t_1-period, and its purpose is labelling the magnetisation with the larmor frequency of the heteronuclei. By re-running the experiment multiple times while incrementing the t_1-period, each acquired 1D-spectra will have experienced an extra frequency labelling and amplitude modulation. By combining these spectra we can then reconstruct the FID of this indirect dimension. As for any acquisition period, decoupling (section 2.1.2) of some form is typically required during the t_1-period. In addition, just as for 1D-experiments, both cosine and sine-modulated signals are needed to distinguish positive and negative offset-frequencies. Due to the nature of how the indirect dimension is acquired, quadrature detection cannot be employed directly. Instead, each experiment is run twice for each t_1-increment with different phases of the magnetisation at the start of the t_1-period. In other words, the cosine and sine components are acquired separately.

The lineshape of a Fourier transformed multidimensional complex signal is the product of the two component signals. As noted in chapter 1.2, the lineshape of a complex signal has a real absorptive and an imaginary dispersive lineshape. This means that both the real and imaginary components of the product of two signal lineshapes will contain an undesirable mixture of absorptive and dispersive lineshapes. This can be circumvented with the States (or States-Haberkorn) [93] method of phase discrimination in indirect dimensions. The basic principle is to only combine the real component of the cosine modulated signal with the real component of the sine-modulated signal. The product of these signals will then generate purely absorptive lineshapes. The States method of phase discrimination is arguably the most analogous to quadrature detection in the direct dimension. But, for modern NMR experiments it is more common to use either an extension of States-Haberkorn called States-TPPI [94] or the different sensitivity enhanced methods such as preservation of equivalent pathways (PEP) [95]. These methods differ in aspects including their intrinsic sensitivity, peak folding properties and axial peak positioning. Peak folding

refers to how peaks outside the studied frequency range are folded into the spectra. Axial peaks originate with magnetisation that have relaxed back to equilibrium along B_0 during the t_1-period, and therefore lack frequency labelling [96].

4.3.2 NMR relaxation experiments

The main practical difference when running a relaxation experiment is that in addition to the frequency labelling periods (t_1, t_2 etc.), we need to introduce a relaxation period. This relaxation period aims to either sample the typically mono-exponential relaxation decay (R_1, R_2) or the magnetisation build-up curve from cross-relaxation pathways (NOE, η_{xy}, η_z). In a similar fashion to an indirect dimension, these decay and build-up curves are sampled one point at a time for every acquisition period. However, this doesn't require quadrature detection, as we directly acquire both the sign and magnitude of each peak signal. So in addition to the typically two or more standard dimensions (^1H, ^2H, ^{15}N, ^{13}C), for relaxation experiments we need to add an additional so called pseudo-dimension [57].

4.4 Processing NMR data

In addition to the Fourier transform, there are several other processing methods that are commonly applied to NMR spectroscopy data. This includes linear prediction, zero filling, solvent filtering and baseline correction [57]. A method that puts special requirements on processing is non-uniform sampling (NUS), described in section 4.8.2 below. As an example, the next section will describe window functions, which will help in the later description of the Accordion method in section 4.8.1.

4.4.1 Window Functions

This description is based on chapter 5 of *Understanding NMR spectroscopy* [55]. The NMR signal over time (FID) decays exponentially with the transverse relaxation rate constant R_2, as described by equation 2.30. This means that for longer FID sampling times, we gain resolution but with diminishing protein signal. On the other hand, the average random noise level is constant with time. Taken together, this means that at longer

time-points the FID becomes increasingly dominated by noise. The FID can be optimized for sensitivity or resolution by multiplying it with a suitable mathematical function, denoted *window function*. For example, multiplying the FID by an exponential decay function will emphasize signal intensity at the cost of resolution and linewidth. Conversely, a rising exponential function will emphasise sharp, resolved peaks with the cost of increased noise.

4.5 Measuring longitudinal relaxation rates

For the case of an isolated ^{1}H-^{15}N spin pair, the dipolar longitudinal relaxation of the ^{15}N spin is described by equation 3.29. Starting from this equation, the next two subsections will describe the basis of measuring the ^{15}N R_1 auto-relaxation rate and σ cross-relaxation rate to the ^{1}H nuclei.

4.5.1 The Longitudinal auto-relaxation rate, R_1

To start with, ^{15}N magnetisation is inverted to phase -Z. By decoupling (de-phasing) the ^{1}H-magnetisation during the following relaxation period, dipolar cross-relaxation and cross-correlated dipol/CSA relaxation between the ^{15}N and ^{1}H spin becomes zero. The evolution of the ^{15}N-magnetisation can then be described by the first term of the Solomon equation:

$$\frac{d\Delta N_z(t)}{dt} = -R_1 \Delta N_z(t) \tag{4.3}$$

The solution of this differential equation for $N_z(t)$ (where $\Delta N_z(t) = N_z(t) - N_z^0$) is a mono-exponential decay on the form:

$$N_z(t) = C_1 e^{-R_1 t} + N_z^0 \tag{4.4}$$

We know that for $N_z(0)$, $e^{-R_1^N t} = 1$, so $C_1 = N_z(0) - N_z^0$. This leads to the expression:

$$N_z(t) = (N_z(0) - N_z^0) e^{-R_1 t} + N_z^0 \tag{4.5}$$

which is the basis of *inversion recovery experiments* measuring longitudinal auto-relaxation rates. These have been measured for papers I, II, and IV.

4.5.2 The steady-state Heteronuclear NOE

Applying a continuous rf-pulse matched to the ^{1}H larmor frequency for a long period ($\approx 5/R_1^N$) will saturate the ^{1}H-spin. This means that the average populations of the ^{1}H spin states will be equalised and $H_z(t)$ becomes 0. This will put the ^{15}N in a steady state (N_z^{ss}), meaning that $\frac{d\Delta N_z(t)}{dt} = 0$. Equation 3.29 then becomes:

$$0 = -R_1(N_z^{ss} - N_z^0) - \sigma_{NH}(0 - H_z^0) \tag{4.6}$$

Solving for $\langle N_z^{ss}\rangle/\langle N_z^0\rangle$, and setting $\langle H_z^0\rangle/\langle N_z^0\rangle = \gamma_H/\gamma_N$ gives (equation 5.149 in [57]):

$$\langle N_z^{ss}\rangle/\langle N_z^0\rangle = 1 + \frac{\sigma_{NH}\gamma_H}{R_1\gamma_N} \tag{4.7}$$

where the second term is called the NOE enhancement η_{NH}. This η_{NH}-factor can be measured as the difference in magnetisation between a reference spectrum and a spectrum with saturated ^{1}H spins [97, 98]. This dipolar cross-relaxation rate was measured for paper I.

4.6 The transverse relaxation rate and Chemical Exchange

This section focus on the measurement of the transverse auto-relaxation rate and chemical exchange phenomena in the intermediate to fast exchange regime using Carr-Purcell Meiboom-Gill (CPMG) and $R_{1\rho}$ based relaxation experiments. As noted in chapter 3, these experiments are closely related since the chemical exchange relaxation rate term (R_{ex}) manifests as a contribution to the transverse auto-relaxation rate, R_2. There are other experiments designed for measuring slower chemical exchange, including adapted versions of the $R_{1\rho}$ experiment, Chemical Exchange Saturation Transfer (CEST) and ZZ-exchange type experiments (for an overview, see reference [99]).

4.6.1 Overview of the CPMG and $R_{1\rho}$ experiments

Typical heteronuclear CPMG or $R_{1\rho}$ experiments starts with the transfer of magnetisation from ^{1}H to the nucleus of interest (such as ^{15}N). The transferred magnetisation is flipped into the transverse plane at the beginning of the relaxation period, where either a CPMG-pulse train or $R_{1\rho}$-spin lock is used to modulate the R_{ex} contribution to R_2. The effective R_2 relaxation rate is then estimated from the resulting exponential decay.

The CPMG-pulse train is in effect a repeated series of 180°-pulses that refocuses the R_{ex} contribution, similarly to a normal spin echo. For an example of how it works, start with a ^{15}N-nuclei N in chemical environment A, with Larmor frequency ω_A. At the start of the relaxation period, N has a pure x-phase. This is followed by a free precession delay of length τ_{cpmg}. During the time of this delay, the magnetization vector of N will rotate $\omega_A\tau_{cpmg}$ radians. Next is a π-pulse around the x-axis, rotating the magnetization vector of N to a phase of $2\pi - \omega_A\tau_{cpmg}$ radians. This is followed by another delay of equal length (τ_{cpmg}) to the first. During this second delay the magnetization vector of N will rotate an additional $\omega_A\tau_{cpmg}$ radians, ending at 2π radians. That means nuclei with different larmor frequencies will be refocused to the same phase at the end of the second delay.

The complexity increases if the nuclei randomly jumps to and from another chemical environment B (due to protein dynamics *etc.*). This also makes the Larmor frequency jump from ω_A to ω_B. If the average time spent in A and B is shorter than the total delay time $2\tau_{cpmg}$, the effective larmor frequency can differ between the first and second delay period. The refocusing effect will then be decreased. The degree of refocusing is proportional to the amplitude of the signal peak, allowing us to probe chemical exchange by studying the peak amplitude as a function of the refocusing frequency, $\nu_{cpmg} = 1/(4\tau_{cpmg})$. The chemical exchange contribution to R_2 for a given nuclei is suppressed for τ_{cpmg} values where $k_{ex}\tau_{cpmg} \ll 1$ [100] for a given given chemical exchange rate k_{ex}.

The $R_{1\rho}$ experiment works in a similar fashion. The difference is that instead of a series of pulses, this experiment uses a continuous pulse. Similar to a π-pulse this will continuously rotate the magnetisation around the axis of the pulse, effectively locking the magnetisation in place. A continuous pulse puts greater strain on the spectrometer, which has historically limited the RF power output (although that might be changing [101]). This leads to some special requirements for the $R_{1\rho}$ experiment that will be discussed in section 4.6.3. The lower power of the continuous pulse will increase

off-resonance effects. However, these are relatively well defined for this experiment, which enables the characterisation of chemical exchange by varying both the pulse frequency ω_{sl} and and power, as these both modify the effective field strength.

Both CPMG and $R_{1\rho}$ experiments involve the application of a significant amount of rf power. This can cause sample heating. If the experiment involves different amounts of input RF power for different spectra (such as spectra with different ν_{cpmg}), the sample heating must be compensated. This is done by adding a variable pulse to the beginning of the experiment, which is calculated to compensate the difference in RF output between spectra [102]. For more details about the CPMG and $R_{1\rho}$ relaxation experiments, see for example chapter 8 of [57].

4.6.2 Special considerations for CPMG dispersion experiments

A potential error source for CPMG experiments is the time-dependent interchange between in-phase and anti-phase magnetisation, as these have different relaxation rates. This will naturally occur in the presence of cross-correlated dipole/CSA relaxation. Special care has to be taken to suppress this pathway via pulsing on the amide protons during the CPMG-train [100]. Such interchange of magnetisation is also mediated by the J-coupling. This pathway is suppressed if the τ_{cpmg} delays are short enough to fulfill [100]:

$$\sin(2\pi J\tau_{cpmg})/2\pi J\tau_{cpmg} \approx 1 \tag{4.8}$$

This is typically fulfilled for ordinary measurements of the auto-relaxation rate constant R_2, where τ_{cpmg} is kept minimal to suppress chemical exchange contributions. However, CPMG *dispersion* experiments aim to quantify chemical exchange processes, which requires longer τ_{cpmg} delays. For this reason *relaxation compensated* CPMG experiments have been created that includes swapping blocks that interchange in-phase and anti-phase magnetisation at fixed time-points. These swapping blocks were originally introduced in the context of spin state selective experiments [103], discussed in section 4.7 below. For a discussion on relaxation compensated CPMG experiments, see section 8.3.4 of [57].

In addition, CPMG-pulse trains benefit from an even number of π-pulses, as this partially compensates potential flip angle errors [104]. A related issue is that hardware and heating limitations mean that the RF-power of

the π-pulses generally is insufficient to ensure proper rotations for all peaks in the sample. Peaks with frequencies far away from the carrier experience off-resonance effects, which partially rotates the magnetisation away from the transverse plane. These effects can be evaluated via simulation [105], and partially compensated for via pulse sequence design [106].

4.6.3 Special considerations for $R_{1\rho}$ dispersion experiments

As mentioned above, $R_{1\rho}$ experiments can be compared to CPMG experiments in the limit $\tau_{cpmg} \rightarrow 0$. This means that condition 4.8 is fulfilled, and anti-phase evolution via the J-coupling is not of any concern. However, care needs to be taken to avoid anti-phase evolution from dipolar-CSA cross-correlated relaxation [107], in direct analogue to the CPMG experiment.

For normal RF-pulses the flip angle can be calibrated by looking at the signal amplitude response to different pulse lengths and power settings. Given the lower power of the continuous spin lock pulses used for $R_{1\rho}$ experiments, these are typically calibrated using nutation experiments [108] in the form proposed by Guenneugues and co-workers [109]. The basis of nutation experiments is to place a continuous rf pulse orthogonally to the initial longitudinal magnetisation, and to increment the experiment over time. This will frequency-label the magnetisation with the frequency of the continuous pulse, analogously to how an indirect dimension is created. This will generate a spectrum with RF frequency in the axis of the indirect dimension, and the peaks visible at their experienced RF frequency.

As $R_{1\rho}$ experiments typically employ lower power pulses during the relaxation period, the off-resonance effects will be more significant. To ensure an effective spin lock, the magnetisation from each signal needs to be *aligned* with its effective field. Hansen and Kay developed alignment sequences for this purpose [110], based on the insight that off-resonance effects can be accounted for using suitable free precession periods. These constructs allow for very precise alignment of signals within $\pm\omega_{sl}$ rad/s of the carrier frequency. For alignment of signals with greater offsets ms long adiabatic sweep pulses are often used [111], that sweep over the entire frequency range of interest, and aligns signals as it passes. To allow proper spin locking, this requires that the adiabatic condition is fulfilled (equation 3.122 in [57]):

$$\left|\frac{d\theta}{dt}\right| \ll \omega_{eff} \tag{4.9}$$

That is to say, the sweep rate ($\frac{d\theta}{dt}$) needs to be much slower than the magnetisation precession rate around the effective field.

4.7 Spin state selective methods and TROSY

The special properties of coupled NMR signals from different spin states can be exploited in several ways. The distance between the peaks is a direct measure of the J-coupling between coupled nuclei. This measurement was the original purpose of two pulse sequence elements used for the selection of specific spin states, the S^3E [103] and S^3CT [112]. The different effective relaxation rates of spin state selective peaks is commonly exploited in TROSY (transverse relaxation optimised spectroscopy) [113] experiments, where the slowly relaxing peak is selected before the detection period. The η_{xy} mediated interchange between in-phase and anti-phase magnetization is sometimes used for magnetization transfer between ^{1}H and heteronuclei, called cross relaxation-enhanced polarization transfer (CRINEPT) [92]. Both TROSY and CRINEPT especially benefit the sensitivity of fast relaxing molecules such as large proteins. Finally, the difference in the effective relaxation rate can be used to measure the dipolar-CSA cross-correlated relaxation rates (η_z, η_{xy}) themselves, which have been done in several experiments ([114–116]), including paper V of this thesis.

4.8 Speeding up NMR measurements

The inherent insensitivity of NMR spectroscopy has necessitated the development of methods for increased sensitivity and reduced measurement time. NMR spectroscopy has seen a large overall increase in sensitivity from the introduction of ^{1}H-detected heteronuclear experiments (see chapter 7 of [57]), cryogenically cooled probe heads (see section 3.1 of [57]), and sensitivity enhanced detection schemes [95]. Other notable experimental methods that can increase sensitivity include the SOFAST [117, 118] approach and single-scan spectroscopy [119]. Nonetheless, it is still of high interest to find methods to decrease the acquisition time, especially for

higher dimensional experiments (ND>2). This is highly relevant for relaxation experiments, given the added inherent sensitivity loss resulting from the relaxation periods.

Relaxation dispersion experiments can be measured using constant time (CT) relaxation periods [120, 121].This refers to only measuring a single point of the relaxation decay for a given offset, spin lock or CPMG refocusing frequency. A reference experiment absent of the relaxation period is also measured. Reducing the number of sampled points in a decay curve can be very efficient in general [122], as it allows for a significant time saving. It also simplifies the analysis in cases where the sampled nucleus experiences other, refocusing (ν_{cpmg}, ω_{off}) independent relaxation processes such as homo-nuclear dipolar cross relaxation [120]. However, cases where the initial population is variable dependent (as for off-resonance $R_{1\rho}$), semi-empirical correction terms might be required [121]. The sampling of other indirect dimensions can also be reduced. A strategy that is suitable for chemical exchange, which often affects only a subset of residues, is to measure individual peaks using either 1D-experiments [121] or via the F_1F_2-selective approach [123] that selects single 2D peaks. The next section describes a reduced dimensionality method called Accordion, which was used for paper IV and V of this thesis in combination with non-uniform sampling (NUS).

4.8.1 Accordion Spectroscopy

The basic idea behind accordion experiments [124] is to synchronize the incrementation of the relaxation period (t_r) with the indirect dimension labelling period (t_1). Several accordion implementations exist; here I will demonstrate the constant-time version of Mandel and Palmer [125]. Assume an experiment with transverse ^{15}N magnetization at the beginning of the relaxation period, and a t_1-period kept at constant total length by the addition of a relaxation compensating period T for short values of t_1. Then if we increment t_r by t_1, we get the following time signal:

$$S_f(t_1) = S_0 e^{-i\omega} e^{-R_i k \Delta t_1} e^{-R_2 (k+1) \kappa \Delta t_1} \tag{4.10}$$

where R_i is the transverse relaxation rate during the t_1 period including contributions from inhomogeneities in the static magnetic field, k is the incrementing index, and κ is a scaling factor between the time step in the t_r

and t_1 periods ($= t_r/t_1$). It is apparent in equation 4.10 that we get an 2D-spectrum where the relaxation period contributes an additional exponential decay to the ^{15}N time signal. This is analogous to a window function (section 4.4.1), which broadens the peak linewidth to $R_f = R_i + \kappa R_2$. In order to deconvolute R_2 contribution to the effective relaxation rate constant R_f, at least one additional experiment is necessary. Mandel and Palmer [125] proposed an additional experiment where t_r is initially at its maximum ($n_{max}\kappa\Delta t_1$) and decremented with t_1:

$$S_r(t_1) = S_0 e^{-i\omega} e^{-R_i k \Delta t_1} e^{-R_2(n-k)\kappa\Delta t_1} \tag{4.11}$$

where the kR_2 factor has the reverse sign compared to equation 4.10. This means that the peak linewidth from this experiment will be $R_r = R_i - \kappa R_2$. This sign difference allows the estimation of R_2 from R_f, R_r as:

$$R_2 = \frac{R_f - R_r}{2\kappa} \tag{4.12}$$

4.8.2 Non-uniform sampling (NUS)

NMR samples does not generate signals at all frequencies in a measurement range. Also, as noted the time-points far out on the FID are dominated by noise. Altogether, this provides the rationale for the development of non-uniform sampling (NUS) schemes of indirect dimensions (for a brief overview, see section 9.3.1 of [57]). When using NUS in an experiment, the sampling scheme (the distribution of sampled points in the FID) is determined beforehand. One example of NUS sample scheme generation is to follow the sinusoidal weighted Poisson-Gap distribution [126], written as:

$$f(k|\lambda) = \frac{\lambda^k e^{-\lambda}}{k!} \tag{4.13}$$

Where k is the index of sampled points, and λ is the mean gap between them. λ is defined as a sinusoidal wave, $\lambda = \Lambda \sin(\theta)$. The θ parameter is varied linearly between $(0,\ \pi)$, but can be shifted by Λ. For example, shifting to $(0,\ 2/\pi)$ weighs up early points in the FID in a similar fashion to the exponential window function described in section 4.4.1. The Fourier transform should not be directly used for NUS datasets [127]. A method to

reconstruct the full FID from the NUS sampling is necessary. The method used for this thesis work is the DSURE (damped super-resolution estimator) algorithm [128]. The details of this statistical model is beyond the scope of this thesis, but the essence is that each FID is assumed to consist of a sum of exponentially decaying sinusoidal signals. DSURE reconstruction stands out in that it ensures a linear correspondence between signal amplitudes and peak volumes, and that it doesn't require the number of signals to be previously known [129]. This can be contrasted with for example maximum entropy (MaxEnt) reconstruction, which is inherently non-linear [130].

CHAPTER 5

Other important Methods

This chapter briefly describes other experimental methods used for this thesis work. It is important to note that different methods have different strengths and are more often than not complementary to each other. Integrating different methods with each other is central to the rapidly advancing field of Integrative Structural Biology [131].

5.1 Isothermal Titration Calorimetry (ITC)

The basis of calorimetry is the measure of heat exchange in chemical reactions. This heat is proportional to the change in enthalpy per mole of reactant. Given that heat exchange is nearly ubiquitous, calorimetric methods can be used for a wide range of studies. There are several different types of calorimeters available that provide different information [132]. The basic components of an isothermal titration calorimeter (ITC) instrument are a pair of measurement cells (sample+reference), a syringe to inject reactant into the sample cell, a stirrer, and a heating block to keep constant temperature in both cells [133].

The heat exchange is measured as the energy needed to keep constant temperature in the sample cell during titration of reactant. This energy typically decreases along the titration, as the sample is increasingly saturated with reactant. This will generate a plot of energy over time, consisting

of a series of peaks corresponding to each injection [134]. These peaks are integrated, and the resulting binding isotherm fitted to estimate the dissociation constant (K_d) and the stoichiometric ratio between sample and reactant. Once these parameters are found, it is straightforward to calculate the Gibbs free energy (ΔG), the enthalpy (ΔH) and entropy (ΔS) of binding assuming proper temperature calibration [135]. This calculation will give the free energy of everything that happens in the sample cell during the titration, including contributions from coupled processes such as protonation [135] and mixing of sample and reactant [134].

Mixing effects should be corrected for by running reference experiments measuring the effects of diluting the sample in solvent, diluting the reactant in solvent, and mixing the different solvents used for sample and reactant [134]. Sample preparation is very important for good ITC data. The solvent buffer of reactant and sample should be matched to minimize the free energy of mixing. The quantity of added reactant needs to be sufficient to saturate the sample [134]. The concentration of sample and reactant needs to be high enough to produce a detectable heat exchange. Also, for ideal data the isotherm needs to have a good curvature. This can be related to low values of the parameter $c = [S]/K_d$, where [S] is the sample concentration [132].

ITC experiments have been reported for measurement of a wide range of reactions, including protein-protein interactions, protein-ligand interactions, enzyme kinetics, and interactions of proteins with other molecules such as membranes, polysaccharides, and nucleic acids [135]. For this thesis work, ITC data were used in paper III to measure the free energy of binding between the CRD of Galectin-3 and a small ligand in the presence of DMSO.

5.2 Molecular Dynamics (MD) simulations

Simulations of molecular systems, such as proteins, can provide unparalleled control and atomistic level of detail. When used in conjunction, simulations and experimental methods can complement and guide each other [136]. Hypothetically, simulations can be performed in an exact way by solving the quantum mechanical Schrödinger equation for the studied system. However, neither this equation nor the comparatively simpler Newtonian equations of motions can be solved analytically for any but the simplest systems [137]. Instead, these equations are typically solved numerically in incremental time steps. For the study of proteins, this is usually

done using classical mechanics, that is to say Newtonian physics. Quantum mechanical simulations, while sometimes more exact, are much more computationally demanding and often unnecessary for the research questions at hand [137].

Central to classical simulation methods is the ergodic hypothesis, which says that for sufficiently long time periods, the average behaviour of a system does not depend on its initial conditions. This enables the computation of system properties averaged over either time or different ensembles [137]. Monte Carlo (MC) simulation is one example of ensemble averaging methods. The method calculates the energy of a given system configuration (for example the positions of all atoms), then randomly perturbs the configuration and calculates the change in energy, and then accept the new configuration with a probability that depend on the energy change. This is repeated to sample a full ensemble of system configurations [138].

For time averaging and time resolved studies, the main method is molecular dynamics (MD) simulations. The principle of MD is to solve the trajectories of all particles in a system as a function of time, using Newton's laws of motion [139] (as expressed by the more general Hamilton equations [138]). The resulting *trajectory* is a description of the time evolution of the system. Doing this requires a description of the forces acting on each particle, which is provided by (semi)-empirical force fields. These typically describe three types of forces, through-bond forces, local through-space (van der Waal) forces and electrostatic (Coulomb) forces. Force-fields can be defined using different methods, including via quantum mechanical calculations and comparison to experimental NMR data [138, 139].

MD simulations can be run in different conditions, for example using a constant number of particles, volume, and temperature (the canonical ensemble) or with a constant number of particles, pressure, and temperature (the isothermal-isobaric ensemble). The solvent can be modelled explicitly, or as an average energy potential acting on the system. If the solvent is explicit, the system is typically encased in a periodic boundary box [138]. To reduce the computational burden of simulations, different simplified models of the molecules or forces acting on the system can be used. These are called coarse-grained simulations and are highly useful for slower processes and larger systems, where full atomistic simulations might be unnecessary or too costly [140]. Finally, an increased range of conformations can be reached by enhanced sampling methods. These methods include disturbing the system by permutation of individual atoms (alchemical methods), simulation at several temperatures in parallel (parallel tempering), and

by imposing a energy penalty on the system against revisiting previously sampled conformations (metadynamics) [141].

The area of use for MD simulations is extremely large, and includes study of conformational changes, membrane transport, protein folding, and ligand binding [139]. In this thesis, MD simulations are used in paper II. A common use of MD is the calculation of order parameters for comparison with NMR data. These parameters can be calculated from MD trajectories, either directly by calculating internal time correlation functions (TCFs) or via the iRED protocol [142, 143].

5.3 Solid state NMR

So far, the principles of NMR described in chapters 2, 3 and 4 have been written with solution NMR spectroscopy in mind. As the name implies, this is distinguished from solid-state NMR (ssNMR) spectroscopy based on the state of the studied sample. Studying solid samples using ssNMR is highly useful in the pharmaceutical industry, as ssNMR can be used to analyze formulated pharmaceuticals in detail in a non-destructive way. It is capable of quantifying the packing states of ingredients used for a formulation in the final, solid form. In addition, it provides information on the stability and dynamics of the ingredients in the formulated product [144]. This also makes it an important tool for pharmaceutical research into understanding and predicting these properties. In addition, large biological systems such as membrane proteins and protein fibrils are well suited for ssNMR [145].

NMR data acquired from solid samples differ from solution measurements in that three additional interactions are active. These are heteronuclear dipolar couplings, chemical shift anisotropy and homonuclear dipolar couplings. These interactions where previously discussed in the context of NMR relaxation in chapter 3. The directionality of these interactions are defined by $(3\cos^2(\theta) - 1)$, where the angle θ relates the bond vector (or largest axis in the case of CSA) to the static magnetic field [70].

There are two main ways to modulate these interactions, either via spinning the sample, or by aligning the angle of the sample such that the directional term of the interactions become 0. This happens at the *magic angle* of 54.74°. When used together, this is known as Magic Angle Spinning (MAS). Spinning at speeds comparable to the strength of the interaction coupling averages it. Unfortunately, homonuclear ^{1}H interac-

tions often have magnitudes of 100kHz or more, making them impractical to significantly modulate this way. Thus, ssNMR experiments often use other less sensitive nuclei for detection, limiting the sensitivity. To improve sensitivity, cross polarization techniques based on Hartmann-Hahn matching (discussed briefly in chapter 4) can be used to transfer magnetization from ^{1}H nuclei to the nuclei of interest [70]. Another method used to increase sensitivity is dynamic nuclear polarization (DNP), which is based on transferring magnetization directly from electrons [145]. Given that these three interactions can be observed directly using ssNMR, this technique can provide unique information. ssNMR data from measurements of the ^{19}F CSA was used in paper II.

5.4 X-ray Crystallography

As mentioned in chapter 1, X-ray crystallography is a common [146] way to find the structure of molecules, such as proteins. This requires the molecule of study to be trapped in crystals. These can be grown from pure, concentrated sample solutions in combination with different salts and solvent. Inside the crystal, individual protein molecules are packed into repeated unit cells. The exact packing depends on crystallization conditions [16]. The crystals themselves are packed rather loosely, with holes and channels taking up to 60% of the total volume [147]. The resulting data quality depends on how well the protein is packed in the crystals [16].

X-rays used for crystallography are typically produced by synchrotron facilities, where electrons are accelerated to near the speed of light. The light emitted by these electrons has high intensity and can be filtered to become monochromatic. During measurement, the light passes through the crystal and is diffracted by the electrons in the studied protein, creating a structure specific diffraction pattern. The crystal itself is rotated to capture the diffraction pattern from different angles, and often cooled with nitrogen gas to minimize radiation damage from the beam [16]. Cooling the crystal to 100K generally improves data quality [147]. A simple illustration of the relationship between the diffraction pattern and the molecular structure is Bragg's law. This is formulated as $2d\sin(\theta) = n\lambda$, where d is the distance between two atoms, θ is the angle of diffraction, $2d\sin(\theta)$ is the length of the path difference between crystal and detector for two beams, n is an integer, and λ is the wavelength of the beam. This means that the diffracted light of two beams hitting different atoms constructively adds up when the

path difference between them is a multiple of the beam wavelength [147].

Each spot on the diffraction pattern is defined by three parameters, amplitude (measured), wavelength (controlled by the synchrotron source), and the phase which usually cannot be measured directly [16]. Instead, the phase can be inferred using one of many possible methods [148]. The collected diffraction pattern is Fourier transformed and used to generate an electron density map [16]. This map is a time-averaged description of the electron cloud, meaning that flexible sections of a molecule will have a less well-defined electron density. Hydrogen atoms are rarely visible as they have an inherently low electron density [149]. Assuming the amino acid sequence is known, this can be fitted to the electron density to generate a structure [16]. In addition to atomic coordinates, protein structure files typically contain occupancy factors, which describes the fraction of time an atom is found in the given position. Also included are atomic displacement parameters (B-factors), which estimates the atomic positional probability density [149].

An important application of X-ray crystallography is in the pharmaceutical industry, where it is often used to study protein – ligand interactions [149]. Improvements in robotics has allowed for large scale semi-automated crystallography trials [146, 149]. Protein structures determined via X-ray crystallography were used for papers I, II, and III.

CHAPTER 6

Guide to the Papers

6.1 Paper I

In this paper we use NMR relaxation experiments to study the dynamics of three BRD4 constructs both in apo and bound to peptides from the N-terminal sequence of histone 4 (residues 1–16) with lysine acetylation on K5, K8, K12 and K16 (tetra-acetylated; denoted H4Kac-4). The BRD4 constructs are the isolated bromodomain 1 (BD1), the isolated bromodomain 2 (BD2), and the two domains in tandem with a disordered linker region. The dissociation constants (K_d) of the bromodomains to H4Kac-4 were estimated by titrating the peptides to each BRD4 construct and tracking the chemical shift changes. For the isolated domains, K_d was estimated to 9 μM and 74 μM for BD1 and BD2 respectively, while the corresponding values for the tandem construct are 15 μM and 125 μM. These titrations also serve as an in-solution confirmation of the expected binding patterns to the bromodomains as determined by earlier crystal structures.

For all constructs ^{15}N NMR relaxation data was acquired, consisting of longitudinal (R_1) and transverse (R_2) auto-relaxation rates, heteronuclear NOEs, and transverse dipole-CSA cross-relaxation rates (η_{xy}). It could be directly observed from the differences in R_1 and R_2 that the rotational diffusion time (τ_m) was the shortest (faster diffusion) for isolated BD1, followed by isolated BD2 and the longest for the tandem construct. The difference between isolated BD1 and BD2 is somewhat surprising, given the high de-

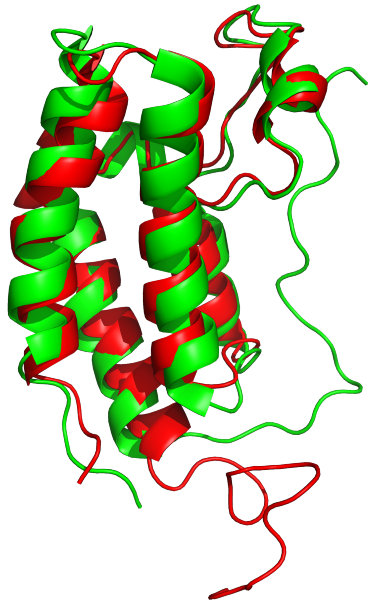


Figure 6.1: Aligned crystal structures of the BRD4 bromodomains. BD1 is shown in green (pdb: 4CLB) and BD2 is shown in red (2LSP). The structures were aligned and the image rendered using the PyMOL software package [33].

gree of similarity between the structures of the two domains (see figure 6.1). While this difference could in part be due to increased chemical exchange present in BD2, the difference persists when when calculating the rotational diffusion time using a combination of R_1, η_{xy}, and NOE data, which does not have any chemical exchange contribution. This led to the conclusion that the isolated BD2 domain is likely self-associating at the concentrations used for NMR experiments (in the hundreds of micromolar). This finding is also backed up by model free (MF) modelling of the relaxation data, and HPLC size exclusion experiments. The self-association of BD2 has been noted for other, much longer BRD4 constructs where it was linked to the phosphorylation of a site downstream of BD2.

When it comes to faster ps-ns internal dynamics of the different constructs, BD1 and BD2 showed similar behaviour. The NMR data of the linker indicated possible transient structure formation. The binding of H4Kac-4 was qualitatively shown to increase the overall flexibility of each domain.

6.2 Paper II

In this paper we investigate the dynamics of two fluorinated ligand diastereomers in complex with the Galectin-3 CRD, see figure 6.2. The difference in protein dynamics of the same two complexes have been previously studied [150]. Ensemble refinement calculations in that study also showed significant differences in ligand dynamics.

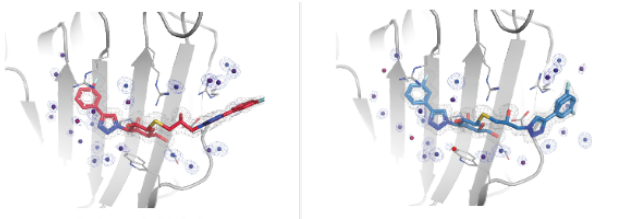


Figure 6.2: Binding pocket of the Galectin-3 CRD in complex with two diastereomer ligands. Both ligands have a fluorine in each end of the ligand, with the (from this viewpoint) leftmost fluorine position more buried inside the protein binding pocket. Left: R-ligand, Right: S-ligand.

We acquired ^{19}F NMR experiments measuring longitudinal, transverse auto-relaxation rates and constant time relaxation dispersions for both complexes. In addition, we had access to previously published MD trajectories [150] and unpublished solid state NMR data of the fluorine CSA for both complexes. Overall, the NMR data was qualitatively consistent with the ensemble refinement, with the biggest difference between the fluorine positions for a given ligand, rather than between the complexes. The fluorine position on the left side of each panel in figure 6.2 is buried deeper inside the protein binding pocket and showed more restricted motions relative to the fluorine position on the right.

Model free (MF) modelling was performed under three assumptions: (1), we can model inter-spatial ^{19}F - ^{1}H interactions as one effective interaction, (2) we can treat the fluorine flexibility as that of angular rotations around a single bond vector, (3) differences in CSA averaging between solid state and solution NMR are negligible. The MF models appear to give realistic results, in qualitative agreement with MD derived order parameters. A small but consistent difference with MD could be due to insufficient MD sampling and a poor force field for small ligands.

6.3 Paper III

In this article we investigated the effect of dimethyl sulfoxide (DMSO) on protein ligand binding dynamics. The CRD of Galectin-3 in complex with a small molecule ligand was studied in the presence of 0, 2, 6, and 10% DMSO (v/v). We acquired isothermal calorimetry (ITC) data and constant-time CPMG dispersion data at each DMSO concentration. The ITC data showed a small but appreciable increase in the ligand dissociation constant (K_d) with DMSO. Two main areas of the protein showed CPMG dispersions. One is on the opposite side of the ligand binding site (around residues 200-220), previously implicated in the dimerization of intact Galectin-3. The other area is the binding pocket itself, which can be identified from crystal structures and the chemical shift changes upon binding. The fitted exchange rates from the CPMG dispersion data showed a decrease with DMSO. Previous studies of Galectin-3 complexes have shown that ligand exchange dominates the dispersions of the residues in the binding pocket. This allowed us to resolve the specific *on* and *off* rates of the ligand exchange by combining the ITC and NMR data. By fitting the DMSO dependence of the on-rates we could then define a success rate of binding for the protein-ligand complexes of around 1%.

6.4 Paper IV

This paper describes a fast method for measuring the transverse auto-relaxation rate constant, R_2. This method combines two existing techniques, non-uniform sampling (NUS) and accordion. Designing accordion pulse sequences requires that the relaxation period can be incremented in small steps. This presents a challenge for R_2 measurements, as these typically contain a CPMG-block during the relaxation period, which have a relatively large minimum size increment.

Instead, the pulse sequence presented in the article uses a spin lock during the relaxation period, which has the significant advantage that it can be of practically arbitrary length. However, this also requires that the protein magnetisation is aligned with the spin lock at the start of the relaxation period. Therefore, we implemented both adiabatic tan/tanh ramps and hard pulse alignment sequences. Both versions of the pulse sequence were tested and found to agree well with a reference experiment run with CPMG-blocks but without accordion or NUS.

6.5 Paper V

This paper is similar to paper IV in that it describes an experimental NMR method that utilises accordion and NUS. However, the goal for this paper is the measurement of the ^{15}N transverse dipolar-CSA cross-relaxation rates η_{xy}. One of the simpler ways to measure this relaxation rate is to acquire a standard 2D HSQC but without ^{1}H decoupling during relaxation. This approach has two major issues. First, decoupled spectra have more peaks which can create spectral overlap. Secondly, the relaxation decay curves will generally be multi-exponential.

To resolve these issues, we utilised the spin state selective editing (S3E) module, as this allows us to individually select the two different proton spin states. This also has the advantage that the initial relaxation will be mono-exponential, as there is no initial second population to exchange with. For this accordion experiment, the auto-relaxation rate of each peak is much greater than the cross-relaxation rate. This means that the second population relaxes faster than it builds up, ensuring mono-exponentiality. However, any attempt to refocus chemical exchange contributions to the auto-relaxation rate needs to be made with caution, as off-resonance magnetisation can be expected to show multi-exponential behaviour. Thus, the experiment presented here lacks this functionality, and this issue is left for future studies. Generally, we find good agreement with cross-relaxation rates measured with reference experiments.

References

[1] J. M. Berg, J. L. Tymoczko, and L. Stryer, *Biochemistry (6th ed.)*. W.H. Freeman and Company, 2007.

[2] D. Goodsell, "Molecule of the month: Ubiquitin," 2004.

[3] D. Goodsell, "Molecule of the month: Capsaicin receptor trpv1," 2020.

[4] D. Dutta, Shuchismita ; Goodsell, "Molecule of the month: Hemoglobin," 2003.

[5] D. Goodsell, "Molecule of the month: Aquaporin," 2014.

[6] D. Goodsell, "Molecule of the month: Pepsin," 2000.

[7] D. Goodsell, "Molecule of the month: Titin," 2015.

[8] I. Gashaw, P. Ellinghaus, A. Sommer, and K. Asadullah, "What makes a good drug target?," *Drug Discovery Today*, vol. 16, pp. 1037–1043, dec 2011.

[9] P. Imming, C. Sinning, and A. Meyer, "Drugs, their targets and the nature and number of drug targets," *Nature Reviews Drug Discovery*, vol. 5, pp. 821–834, oct 2006.

[10] R. Santos, O. Ursu, A. Gaulton, A. P. Bento, R. S. Donadi, C. G. Bologa, A. Karlsson, B. Al-Lazikani, A. Hersey, T. I. Oprea, and J. P. Overington, "A comprehensive map of molecular drug targets," *Nature Reviews Drug Discovery*, vol. 16, pp. 19–34, jan 2017.

[11] L. H. Hurley and F. L. Boyd, "DNA as a target for drug action," *Trends in Pharmacological Sciences*, vol. 9, no. 11, pp. 402–407, 1988.

[12] I. Haq and J. Ladbury, "Drug-DNA recognition: energetics and implications for design," *Journal of Molecular Recognition*, vol. 13, pp. 188–197, jul 2000.

[13] J. B. Chaires and J.-L. Mergny, "Targeting DNA," *Biochimie*, vol. 90, pp. 973–975, jul 2008.

[14] M. S. Searle, "NMR Studies of Drug—DNA interactions," *Progress in Nuclear Magnetic Resonance Spectroscopy*, vol. 25, pp. 403–480, jan 1993.

[15] B. Liu, H. Shi, and H. M. Al-Hashimi, "Developments in solution-state NMR yield broader and deeper views of the dynamic ensembles of nucleic acids," *Current Opinion in Structural Biology*, vol. 70, pp. 16–25, oct 2021.

[16] J. Branden, Carl ; Tooze, *introduction to Protein Structure*. Garland Science, 1999.

[17] A. E. Ferentz and G. Wagner, "NMR spectroscopy: a multifaceted approach to macromolecular structure," *Quarterly Reviews of Biophysics*, vol. 33, pp. 29–65, feb 2000.

[18] K. M. Yip, N. Fischer, E. Paknia, A. Chari, and H. Stark, "Atomic-resolution protein structure determination by cryo-EM," *Nature*, vol. 587, pp. 157–161, nov 2020.

[19] A. Kryshtafovych, T. Schwede, M. Topf, K. Fidelis, and J. Moult, "Critical assessment of methods of protein structure prediction (CASP)—Round XIII," *Proteins: Structure, Function, and Bioinformatics*, vol. 87, pp. 1011–1020, dec 2019.

[20] V. N. Uversky, "Introduction to Intrinsically Disordered Proteins (IDPs)," *Chemical Reviews*, vol. 114, pp. 6557–6560, jul 2014.

[21] P. E. Wright and H. J. Dyson, "Intrinsically disordered proteins in cellular signalling and regulation," *Nature Reviews Molecular Cell Biology*, vol. 16, pp. 18–29, jan 2015.

[22] M. Orozco, "A theoretical view of protein dynamics," *Chem. Soc. Rev.*, vol. 43, no. 14, pp. 5051–5066, 2014.

[23] M. D. Miller and G. N. Phillips, "Moving beyond static snapshots: Protein dynamics and the Protein Data Bank," *Journal of Biological Chemistry*, vol. 296, p. 100749, jan 2021.

[24] V. A. Jarymowycz and M. J. Stone, "Fast Time Scale Dynamics of Protein Backbones: NMR Relaxation Methods, Applications, and Functional Consequences," *Chemical Reviews*, vol. 106, pp. 1624–1671, may 2006.

[25] A. G. Palmer and H. Koss, "Chemical Exchange," in *Methods in Enzymology*, vol. 615, pp. 177–236, Elsevier Inc., 1 ed., 2019.

[26] R. Ishima and D. A. Torchia, "Protein dynamics from NMR.," *Nature structural biology*, vol. 7, pp. 740–3, sep 2000.

[27] T. Misteli, "Protein Dynamics: Implications for Nuclear Architecture and Gene Expression," *Science*, vol. 291, pp. 843–847, feb 2001.

[28] T. Mittag, L. E. Kay, and J. D. Forman-Kay, "Protein dynamics and conformational disorder in molecular recognition," *Journal of Molecular Recognition*, vol. 23, no. 2, pp. n/a–n/a, 2009.

[29] S.-R. Tzeng and C. G. Kalodimos, "Protein dynamics and allostery: an NMR view," *Current Opinion in Structural Biology*, vol. 21, pp. 62–67, feb 2011.

[30] M. Kovermann, P. Rogne, and M. Wolf-Watz, "Protein dynamics and function from solution state NMR spectroscopy," *Quarterly Reviews of Biophysics*, vol. 49, p. e6, mar 2016.

[31] A. Sekhar and L. E. Kay, "An NMR View of Protein Dynamics in Health and Disease," *Annual Review of Biophysics*, vol. 48, pp. 297–319, may 2019.

[32] H. J. Dyson and P. E. Wright, "NMR illuminates intrinsic disorder," *Current Opinion in Structural Biology*, vol. 70, pp. 44–52, oct 2021.

[33] Schrödinger, LLC, "The PyMOL molecular graphics system, version 1.8." Available at: http://www.pymol.org/pymol, November 2015.

[34] L. Johannes, R. Jacob, and H. Leffler, "Galectins at a glance," *Journal of Cell Science*, vol. 131, no. 9, 2018.

[35] D. T. Logan, "Hunting down hydrogen: Applying neutron macromolecular crystallography to galectins," *Biochemist*, vol. 41, no. 2, pp. 20–23, 2019.

[36] S. Bänfer and R. Jacob, "Galectins in intra-and extracellular vesicles," *Biomolecules*, vol. 10, no. 9, pp. 1–12, 2020.

[37] A. Bartolazzi, "Galectins in cancer and translational medicine: From bench to bedside," *International Journal of Molecular Sciences*, vol. 19, no. 10, 2018.

[38] H. Leffler, "Galectin History, Some Stories, and Some Outstanding Questions," *Trends in Glycoscience and Glycotechnology*, vol. 30, no. 172, pp. SE129–SE135, 2018.

[39] R. Tazhitdinova and A. V. Timoshenko, "The Emerging Role of Galectins and O-GlcNAc Homeostasis in Processes of Cellular Differentiation," *Cells*, vol. 9, no. 8, 2020.

[40] F. Barake, A. Soza, and A. González, "Galectins in the brain: advances in neuroinflammation, neuroprotection and therapeutic opportunities," *Current opinion in neurology*, vol. 33, no. 3, pp. 381–390, 2020.

[41] A. Hajmirza, A. Emadali, A. Gauthier, O. Casasnovas, R. Gressin, and M. B. Callanan, "BET family protein BRD4: An emerging actor in NFκB signaling in inflammation and cance," *Biomedicines*, vol. 6, no. 1, pp. 1–9, 2018.

[42] P. Filippakopoulos and S. Knapp, "Targeting bromodomains: Epigenetic readers of lysine acetylation," *Nature Reviews Drug Discovery*, vol. 13, no. 5, pp. 337–356, 2014.

[43] M. E. White, J. M. Fenger, and W. E. Carson, "Emerging roles of and therapeutic strategies targeting BRD4 in cancer," *Cellular Immunology*, vol. 337, no. February, pp. 48–53, 2019.

[44] B. N. Devaiah, A. Gegonne, and D. S. Singer, "Bromodomain 4: a cellular Swiss army knife," *Journal of Leukocyte Biology*, vol. 100, no. 4, pp. 679–686, 2016.

[45] S. Lin and L. Du, "The therapeutic potential of BRD4 in cardiovascular disease," *Hypertension Research*, vol. 43, no. 10, pp. 1006–1014, 2020.

[46] A. Werner, A. G. Manford, and M. Rape, "Ubiquitin-Dependent Regulation of Stem Cell Biology," *Trends in Cell Biology*, vol. 27, pp. 568–579, aug 2017.

[47] R. B. Buxton, *Introduction to functional magnetic resonance imaging: principles and techniques*. Cambridge university press, 2009.

[48] M. Billeter, G. Wagner, and K. Wüthrich, "Solution NMR structure determination of proteins revisited," *Journal of Biomolecular NMR*, vol. 42, pp. 155–158, nov 2008.

[49] C. K. Larive, G. A. Barding, and M. M. Dinges, "NMR Spectroscopy for Metabolomics and Metabolic Profiling," *Analytical Chemistry*, vol. 87, pp. 133–146, jan 2015.

[50] J. L. Markley, R. Brüschweiler, A. S. Edison, H. R. Eghbalnia, R. Powers, D. Raftery, and D. S. Wishart, "The future of NMR-based metabolomics," *Current Opinion in Biotechnology*, vol. 43, pp. 34–40, feb 2017.

[51] P. Selenko and G. Wagner, "Looking into live cells with in-cell NMR spectroscopy," *Journal of Structural Biology*, vol. 158, pp. 244–253, may 2007.

[52] A. Y. Maldonado, D. S. Burz, and A. Shekhtman, "In-cell NMR spectroscopy," *Progress in Nuclear Magnetic Resonance Spectroscopy*, vol. 59, pp. 197–212, oct 2011.

[53] M. Edén, "Zeeman truncation in NMR. II. Time averaging in the rotating frame," *Concepts in Magnetic Resonance Part A: Bridging Education and Research*, vol. 43, no. 4, pp. 109–126, 2015.

[54] M. Edén, "Zeeman truncation in NMR. I. The role of operator commutation," *Concepts in Magnetic Resonance Part A: Bridging Education and Research*, vol. 43, no. 4, pp. 91–108, 2015.

[55] J. Keeler, *Understanding NMR spectroscopy.* Chichester, West Sussex, United Kingdom: John Wiley & sons, Ltd., 2 ed., 2010.

[56] M. H. Levitt, *Spin dynamics: basics of nuclear magnetic resonance.* John Wiley & Sons, 2001.

[57] J. Cavanagh, W. J. Fairbrother, A. G. Palmer, N. J. Skelton, and M. Rance, *Protein NMR Spectroscopy.* Academic Press, 2007.

[58] M. Mehring, *NMR Basic principles and progress.* Berlin: Springer Verlag, 1976.

[59] O. W. Sørensen, G. W. Eich, M. H. Levitt, G. Bodenhausen, and R. R. Ernst, "Product operator formalism for the description of NMR pulse experiments," *Progress in Nuclear Magnetic Resonance Spectroscopy*, vol. 16, pp. 163–192, jan 1984.

[60] T. D. Claridge, *High-resolution NMR techniques in organic chemistry*. Elsevier, 2016.

[61] F. Bloch, "Nuclear Induction," *Physical Review*, vol. 70, pp. 460–474, oct 1946.

[62] R. Ghose, "Average Liouvillian theory in nuclear magnetic resonance - Principles, properties, and applications," *Concepts in Magnetic Resonance*, vol. 12, no. 3, pp. 152–172, 2000.

[63] A. G. Redfield, *The Theory of Relaxation Processes*, vol. 1. Academic Press Inc., 1965.

[64] R. K. Wangsness and F. Bloch, "The dynamical theory of nuclear induction," *Physical Review*, vol. 89, pp. 728–739, feb 1953.

[65] B. Halle and H. Wennerström, "Interpretation of magnetic resonance data from water nuclei in heterogeneous systems," *The Journal of Chemical Physics*, vol. 75, no. 4, pp. 1928–1943, 1981.

[66] G. Lipari and A. Szabo, "Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 1. Theory and range of validity," *Journal of the American Chemical Society*, vol. 104, no. 17, pp. 4546–4559, 1982.

[67] G. Lipari and A. Szabo, "Model-Free Approach to the Interpretation of Nuclear Magnetic Resonance Relaxation in Macromolecules. 2. Analysis of Experimental Results," *Journal of the American Chemical Society*, vol. 104, no. 17, pp. 4559–4570, 1982.

[68] B. Halle, "The physical basis of model-free analysis of NMR relaxation data from proteins and complex fluids," *The Journal of Chemical Physics*, vol. 131, p. 224507, dec 2009.

[69] R. Brüschweiler and P. E. Wright, "NMR Order Parameters of Biomolecules: A New Analytical Representation and Application to the Gaussian Axial Fluctuation Model," *Journal of the American Chemical Society*, vol. 116, no. 18, pp. 8426–8427, 1994.

[70] D. D. Laws, H.-M. L. Bitter, and A. Jerschow, "Solid-state NMR spectroscopic methods in chemistry.," *Angewandte Chemie (International ed. in English)*, vol. 41, pp. 3096–129, sep 2002.

[71] D. A. Case, "Calculations of NMR dipolar coupling strengths in model peptides.," *Journal of biomolecular NMR*, vol. 15, pp. 95–102, oct 1999.

[72] L. Yao, B. Vögeli, J. Ying, and A. Bax, "NMR Determination of Amide N−H Equilibrium Bond Length from Concerted Dipolar Coupling Measurements," *Journal of the American Chemical Society*, vol. 130, pp. 16518–16520, dec 2008.

[73] L. Yao, A. Grishaev, G. Cornilescu, and A. Bax, "Site-specific backbone amide 15N chemical shift anisotropy tensors in a small protein from liquid crystal and cross-correlated relaxation measurements," *Journal of the American Chemical Society*, vol. 132, no. 12, pp. 4295–4309, 2010.

[74] K. Loth, P. Pelupessy, and G. Bodenhausen, "Chemical Shift Anisotropy Tensors of Carbonyl, Nitrogen, and Amide Proton Nuclei in Proteins through Cross-Correlated Relaxation in NMR Spectroscopy," *Journal of the American Chemical Society*, vol. 127, pp. 6062–6068, apr 2005.

[75] S. Tang and D. A. Case, "Calculation of chemical shift anisotropy in proteins," *Journal of Biomolecular NMR*, vol. 51, pp. 303–312, nov 2011.

[76] M. Goldman, "Interference effects in the relaxation of a pair of unlike spin- 1 2 nuclei," *Journal of Magnetic Resonance (1969)*, vol. 60, no. 3, pp. 437–452, 1984.

[77] D. E. Woessner, "Nuclear Transfer Effects in Nuclear Magnetic Resonance Pulse Experiments," *The Journal of Chemical Physics*, vol. 35, pp. 41–48, jul 1961.

[78] H. M. McConnell, "Reaction Rates by Nuclear Magnetic Resonance," *The Journal of Chemical Physics*, vol. 28, pp. 430–431, mar 1958.

[79] P. Neudecker, P. Robustelli, A. Cavalli, P. Walsh, P. Lundström, A. Zarrine-Afsar, S. Sharpe, M. Vendruscolo, and L. E. Kay, "Structure of an intermediate state in protein folding and aggregation.," *Science (New York, N.Y.)*, vol. 336, pp. 362–6, apr 2012.

[80] C. D. Kroenke, J. P. Loria, L. K. Lee, M. Rance, and A. G. Palmer, "Longitudinal and transverse 1H-15N dipolar/15N chemical shift anisotropy relaxation interference: Unambiguous determination of rotational diffusion tensors and chemical exchange effects in biological macromolecules," *Journal of the American Chemical Society*, vol. 120, no. 31, pp. 7905–7915, 1998.

[81] D. Fushman and D. Cowburn, "The effect of noncollinearity of 15N-1H dipolar and 15N CSA tensors and rotational anisotropy on 15N relaxation, CSA/dipolar cross correlation, and TROSY.," *Journal of biomolecular NMR*, vol. 13, pp. 139–47, feb 1999.

[82] I. Solomon, "Relaxation Processes in a System of Two Spins," *Physical Review*, vol. 99, pp. 559–565, jul 1955.

[83] D. S. Weaver and E. R. Zuiderweg, "Protein proton-proton dynamics from amide proton spin flip rates," *Journal of Biomolecular NMR*, vol. 45, no. 1-2, pp. 99–119, 2009.

[84] L. E. Kay, L. K. Nicholson, F. Delaglio, A. Bax, and D. Torchia, "Pulse sequences for removal of the effects of cross correlation between dipolar and chemical-shift anisotropy relaxation mechanisms on the measurement of heteronuclear T1 and T2 values in proteins," *Journal of Magnetic Resonance (1969)*, vol. 97, pp. 359–375, apr 1992.

[85] M. H. Levitt, "The Signs of Frequencies and Phases in NMR," *Journal of Magnetic Resonance*, vol. 126, no. 2, pp. 164–182, 1997.

[86] M. H. Levitt and O. G. Johannessen, "Signs of Frequencies and Phases in NMR: The Role of Radiofrequency Mixing," *Journal of Magnetic Resonance*, vol. 142, pp. 190–194, jan 2000.

[87] A. E. Kelly, H. D. Ou, R. Withers, and V. Dötsch, "Low-conductivity buffers for high-sensitivity NMR measurements," *Journal of the American Chemical Society*, vol. 124, no. 40, pp. 12013–12019, 2002.

[88] M. Wong, G. Khirich, and J. P. Loria, "What's in Your Buffer? Solute Altered Millisecond Motions Detected by Solution NMR," *Biochemistry*, vol. 52, pp. 6548–6558, sep 2013.

[89] P. Kukic, F. O'Meara, C. Hewage, and J. Erik Nielsen, "Coupled effect of salt and pH on proteins probed with NMR spectroscopy," *Chemical Physics Letters*, vol. 579, pp. 114–121, jul 2013.

[90] J. Wallerstein and M. Akke, "Minute Additions of DMSO Affect Protein Dynamics Measurements by NMR Relaxation Experiments through Significant Changes in Solvent Viscosity," *ChemPhysChem*, vol. 20, no. 2, pp. 326–332, 2019.

[91] P. Kumari, L. Frey, A. Sobol, N.-A. Lakomek, and R. Riek, "15N transverse relaxation measurements for the characterization of µs–ms

dynamics are deteriorated by the deuterium isotope effect on 15N resulting from solvent exchange," *Journal of Biomolecular NMR*, vol. 72, pp. 125–137, dec 2018.

[92] R. Riek, G. Wider, K. Pervushin, and K. Wüthrich, "Polarization transfer by cross-correlated relaxation in solution NMR with very large molecules," *Proceedings of the National Academy of Sciences*, vol. 96, pp. 4918–4923, apr 1999.

[93] D. States, R. Haberkorn, and D. Ruben, "A two-dimensional nuclear overhauser experiment with pure absorption phase in four quadrants," *Journal of Magnetic Resonance (1969)*, vol. 48, pp. 286–292, jun 1982.

[94] D. Marion, M. Ikura, R. Tschudin, and A. Bax, "Rapid recording of 2D NMR spectra without phase cycling. Application to the study of hydrogen exchange in proteins," *Journal of Magnetic Resonance (1969)*, vol. 85, pp. 393–399, nov 1989.

[95] M. Rance, "Improved techniques for homonuclear rotating-frame and isotropic mixing experiments," *Journal of Magnetic Resonance (1969)*, vol. 74, pp. 557–564, oct 1987.

[96] W. P. Aue, E. Bartholdi, and R. R. Ernst, "Two-dimensional spectroscopy. Application to nuclear magnetic resonance," *The Journal of Chemical Physics*, vol. 64, pp. 2229–2246, mar 1976.

[97] F. Ferrage, A. Piserchio, D. Cowburn, and R. Ghose, "On the measurement of 15N–1H nuclear Overhauser effects," *Journal of Magnetic Resonance*, vol. 192, pp. 302–313, jun 2008.

[98] F. Ferrage, A. Reichel, S. Battacharya, D. Cowburn, and R. Ghose, "On the measurement of 15N-1H nuclear Overhauser effects. 2. Effects of the saturation scheme and water signal suppression," *Journal of Magnetic Resonance*, vol. 207, no. 2, pp. 294–303, 2010.

[99] A. G. Palmer, "Chemical exchange in biomacromolecules: Past, present, and future," *Journal of Magnetic Resonance*, vol. 241, pp. 3–17, apr 2014.

[100] A. G. Palmer, N. J. Skelton, W. J. Chazin, P. E. Wright, and M. Rance, "Suppression of the effects of cross-correlation between dipolar and anisotropic chemical shift relaxation mechanisms in the measurement of spin-spin relaxation rates," *Molecular Physics*, vol. 75, pp. 699–711, feb 1992.

[101] J. G. Reddy, S. Pratihar, D. Ban, S. Frischkorn, S. Becker, C. Griesinger, and D. Lee, "Simultaneous determination of fast and slow dynamics in molecules using extreme CPMG relaxation dispersion experiments," *Journal of Biomolecular NMR*, vol. 70, pp. 1–9, jan 2018.

[102] A. C. Wang and A. Bax, "Minimizing the effects of radio-frequency heating in multidimensional NMR experiments," *Journal of Biomolecular NMR*, vol. 3, no. 6, pp. 715–720, 1993.

[103] A. Meissner, J. Duus, and O. W. Sørensen, "Spin-State-Selective Excitation. Application for E.COSY-Type Measurement of JHH Coupling Constants," *Journal of Magnetic Resonance*, vol. 128, pp. 92–97, sep 1997.

[104] S. Meiboom and D. Gill, "Modified Spin-Echo Method for Measuring Nuclear Relaxation Times," *Review of Scientific Instruments*, vol. 29, pp. 688–691, aug 1958.

[105] D. M. Korzhnev, E. V. Tischenko, and A. S. Arseniev, "Off-resonance effects in 15N T2 CPMG measurements.," *Journal of biomolecular NMR*, vol. 17, pp. 231–7, jul 2000.

[106] G. N. Yip and E. R. Zuiderweg, "A phase cycle scheme that significantly suppresses offset-dependent artifacts in the R2-CPMG 15N relaxation experiment," *Journal of Magnetic Resonance*, vol. 171, pp. 25–36, nov 2004.

[107] F. Massi, E. Johnson, C. Wang, M. Rance, and A. G. Palmer, "NMR R1ρ Rotating-Frame Relaxation with Weak Radio Frequency Fields," *Journal of the American Chemical Society*, vol. 126, no. 7, pp. 2247–2256, 2004.

[108] H. C. Torrey, "Transient nutations in nuclear magnetic resonance," *Physical Review*, vol. 76, no. 8, pp. 1059–1068, 1949.

[109] M. Guenneugues, P. Berthault, and H. Desvaux, "A method for determining B1 field inhomogeneity. Are the biases assumed in heteronuclear relaxation experiments usually underestimated?," *Journal of magnetic resonance (San Diego, Calif. : 1997)*, vol. 136, pp. 118–126, 1999.

[110] D. F. Hansen, D. Yang, H. Feng, Z. Zhou, S. Wiesner, Y. Bai, and L. E. Kay, "An Exchange-Free Measure of 15 N Transverse Relax-

ation: An NMR Spectroscopy Application to the Study of a Folding Intermediate with Pervasive Chemical Exchange," *Journal of the American Chemical Society*, vol. 129, pp. 11468–11479, sep 2007.

[111] F. A. Mulder, R. A. De Graaf, R. Kaptein, and R. Boelens, "An Off-resonance Rotating Frame Relaxation Experiment for the Investigation of Macromolecular Dynamics Using Adiabatic Rotations," 1998.

[112] M. D. Sørensen, A. Meissner, and O. W. Sørensen, "Spin-state-selective coherence transfer via intermediate states of two-spin coherence in IS spin systems: Application to E.COSY-type measurement of J coupling constants," *Journal of Biomolecular NMR*, vol. 10, no. 2, pp. 181–186, 1997.

[113] K. Pervushin, R. Riek, G. Wider, and K. Wuthrich, "Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution," *Proceedings of the National Academy of Sciences*, vol. 94, pp. 12366–12371, nov 1997.

[114] P. R. Vasos, J. B. Hall, and D. Fushman, "Spin-state selection for increased confidence in cross-correlation rates measurements," *Journal of Biomolecular NMR*, vol. 31, no. 2, pp. 149–154, 2005.

[115] Y. Liu and J. H. Prestegard, "Direct measurement of dipole-dipole/CSA cross-correlated relaxation by a constant-time experiment," *Journal of Magnetic Resonance*, vol. 193, no. 1, pp. 23–31, 2008.

[116] D. S. Weaver and E. R. Zuiderweg, "Z κ: AA transverse relaxation optimized spectroscopy NMR experiment measuring longitudinal relaxation interference," *Journal of Chemical Physics*, vol. 128, no. 15, 2008.

[117] P. Schanda and B. Brutscher, "Very fast two-dimensional NMR spectroscopy for real-time investigation of dynamic events in proteins on the time scale of seconds," *Journal of the American Chemical Society*, vol. 127, no. 22, pp. 8014–8015, 2005.

[118] T. Kern, P. Schanda, and B. Brutscher, "Sensitivity-enhanced IPAP-SOFAST-HMQC for fast-pulsing 2D NMR with reduced radiofrequency load," *Journal of Magnetic Resonance*, vol. 190, pp. 333–338, feb 2008.

[119] A. Tal and L. Frydman, "Single-scan multidimensional magnetic resonance," *Progress in Nuclear Magnetic Resonance Spectroscopy*, vol. 57, pp. 241–292, oct 2010.

[120] F. A. Mulder, N. R. Skrynnikov, B. Hon, F. W. Dahlquist, and L. E. Kay, "Measurement of slow (μs-ms) time scale dynamics in protein side chains by 15N relaxation dispersion NMR spectroscopy: Application to Asn and Gln residues in a cavity mutant of T4 lysozyme," *Journal of the American Chemical Society*, vol. 123, no. 5, pp. 967–975, 2001.

[121] D. M. Korzhnev, V. Y. Orekhov, and L. E. Kay, "Off-resonance R1ρ NMR studies of exchange dynamics in proteins with low spin-lock fields: An application to a fyn SH3 domain," *Journal of the American Chemical Society*, vol. 127, no. 2, pp. 713–721, 2005.

[122] J. Jones, P. Hodgkinson, A. Barker, and P. Hore, "Optimal Sampling Strategies for the Measurement of Spin–Spin Relaxation Times," *Journal of Magnetic Resonance, Series B*, vol. 113, pp. 25–34, oct 1996.

[123] E. Walinda, D. Morimoto, M. Shirakawa, and K. Sugase, "F 1 F 2-selective NMR spectroscopy," *Journal of Biomolecular NMR*, vol. 68, pp. 41–52, may 2017.

[124] G. Bodenhausen and Ernst. R. R., "Direct determination of rate constants of slow dynamic processes by two-dimensional "accordion" spectroscopy in nuclear magnetic resonance," *Journal of the American Chemical Society*, vol. 104, pp. 1304–1309, mar 1982.

[125] A. M. Mandel and A. G. Palmer, "Measurement of Relaxation-Rate Constants Using Constant-Time Accordion NMR Spectroscopy," *Journal of Magnetic Resonance, Series A*, 1994.

[126] S. G. Hyberts, D. P. Frueh, H. Arthanari, and G. Wagner, "FM reconstruction of non-uniformly sampled protein NMR data at higher dimensions and optimization by distillation," *Journal of Biomolecular NMR*, vol. 45, pp. 283–294, nov 2009.

[127] S. G. Hyberts, A. G. Milbradt, A. B. Wagner, H. Arthanari, and G. Wagner, "Application of iterative soft thresholding for fast reconstruction of NMR data non-uniformly sampled with multidimensional Poisson Gap scheduling," *Journal of Biomolecular NMR*, vol. 52, pp. 315–327, apr 2012.

[128] M. Juhlin, F. Elvander, J. Sward, and A. Jakobsson, "Fast Gridless Estimation of Damped Modes," in *ISPACS 2018 - 2018 International Symposium on Intelligent Signal Processing and Communication Systems*, 2018.

[129] G. Carlström, F. Elvander, J. Swärd, A. Jakobsson, and M. Akke, "Rapid NMR Relaxation Measurements Using Optimal Nonuniform Sampling of Multidimensional Accordion Data Analyzed by a Sparse Reconstruction Method," *The Journal of Physical Chemistry A*, vol. 123, pp. 5718–5723, jul 2019.

[130] P. Schmieder, A. S. Stern, G. Wagner, and J. C. Hoch, "Quantification of Maximum-Entropy Spectrum Reconstructions," *Journal of Magnetic Resonance*, vol. 125, pp. 332–339, apr 1997.

[131] H. van den Bedem and J. S. Fraser, "Integrative, dynamic structural biology at atomic resolution—it's about time," *Nature Methods*, vol. 12, pp. 307–318, apr 2015.

[132] M. W. Freyer and E. A. Lewis, "Isothermal Titration Calorimetry: Experimental Design, Data Analysis, and Probing Macromolecule/Ligand Binding and Kinetic Interactions," in *Methods in Cell Biology*, vol. 84, pp. 79–113, Elsevier, 2008.

[133] E. A. Lewis and K. P. Murphy, "Isothermal Titration Calorimetry," in *Protein-Ligand Interactions*, vol. 305, pp. 001–016, New Jersey: Humana Press, 2005.

[134] M. A. O'Neill and S. Gaisford, "Application and use of isothermal calorimetry in pharmaceutical development," *International Journal of Pharmaceutics*, vol. 417, pp. 83–93, sep 2011.

[135] R. Ghai, R. J. Falconer, and B. M. Collins, "Applications of isothermal titration calorimetry in pure and applied research-survey of the literature from 2010," *Journal of Molecular Recognition*, vol. 25, pp. 32–52, jan 2012.

[136] M. Karplus and J. A. McCammon, "Molecular dynamics simulations of biomolecules," *Nature Structural Biology*, vol. 9, pp. 646–652, sep 2002.

[137] D. Frenkel, M. Klein, M. Parrinello, and B. Smit, eds., *Understanding Molecular simulation: From Algorithms to Applications*. San Diego: Academic Press, 2nd ed., 2002.

[138] E. Paquet and H. L. Viktor, "Molecular Dynamics, Monte Carlo Simulations, and Langevin Dynamics: A Computational Review," *BioMed Research International*, vol. 2015, pp. 1–18, 2015.

[139] R. O. Dror, R. M. Dirks, J. Grossman, H. Xu, and D. E. Shaw, "Biomolecular Simulation: A Computational Microscope for Molecular Biology," *Annual Review of Biophysics*, vol. 41, pp. 429–452, jun 2012.

[140] M. G. Saunders and G. A. Voth, "Coarse-Graining Methods for Computational Biology," *Annual Review of Biophysics*, vol. 42, pp. 73–93, may 2013.

[141] V. Spiwok, Z. Sucur, and P. Hosek, "Enhanced sampling techniques in biomolecular simulations," *Biotechnology Advances*, vol. 33, pp. 1130–1140, nov 2015.

[142] J. J. Prompers and R. Brüschweiler, "General framework for studying the dynamics of folded and nonfolded proteins by NMR relaxation spectroscopy and MD simulation," *Journal of the American Chemical Society*, 2002.

[143] Y. Gu, D.-W. Li, and R. Brüschweiler, "NMR Order Parameter Determination from Long Molecular Dynamics Trajectories for Objective Comparison with Experiment," *Journal of Chemical Theory and Computation*, vol. 10, pp. 2599–2607, jun 2014.

[144] R. T. Berendt, D. M. Sperger, E. J. Munson, and P. K. Isbester, "Solid-state NMR spectroscopy in pharmaceutical research and analysis," *TrAC Trends in Analytical Chemistry*, vol. 25, pp. 977–984, nov 2006.

[145] M. Li, W. Xu, and Y. Su, "Solid-state NMR spectroscopy in pharmaceutical sciences," *TrAC Trends in Analytical Chemistry*, vol. 135, p. 116152, feb 2021.

[146] A. Wlodawer, W. Minor, Z. Dauter, and M. Jaskolski, "Protein crystallography for aspiring crystallographers or how to avoid pitfalls and traps in macromolecular structure determination," *FEBS Journal*, vol. 280, no. 22, pp. 5705–5736, 2013.

[147] A. Ilari and C. Savino, "A Practical Approach to Protein Crystallography," in *Methods in Molecular Biology*, vol. 1525, pp. 47–78, Springer Verlag, 2017.

[148] G. Taylor, "The phase problem," *Acta Crystallographica Section D Biological Crystallography*, vol. 59, pp. 1881–1890, nov 2003.

[149] L. Maveyraud and L. Mourey, "Protein X-ray crystallography and drug discovery," *Molecules*, vol. 25, no. 5, 2020.

[150] M. L. Verteramo, O. Stenström, M. M. Ignjatović, O. Caldararu, M. A. Olsson, F. Manzoni, H. Leffler, E. Oksanen, D. T. Logan, U. J. Nilsson, U. Ryde, and M. Akke, "Interplay between Conformational Entropy and Solvation Entropy in Protein-Ligand Binding," *Journal of the American Chemical Society*, 2019.

Scientific publications

Paper I

Bromodomain Interactions with Acetylated Histone 4 Peptides in the BRD4 Tandem Domain: Effects on Domain Dynamics and Internal Flexibility

Sven Wernersson,[1,‡] Romel Bobby [2,4,‡] Liz Flavell,[3,5] Alexander G. Milbradt,[2] Geoffrey A. Holdgate,[2,6] Kevin J. Embrey,[2,7*] and Mikael Akke[1*]

[1] Biophysical Chemistry, Center for Molecular Protein Science, Department of Chemistry, Lund University, SE-221 00 Lund, Sweden; [2] Mechanistic and Structural Biology, Discovery Sciences, BioPharmaceuticals R&D, AstraZeneca, Cambridge, UK; [3] Discovery Biology, Discovery Sciences, BioPharmaceuticals R&D, AstraZeneca, Cambridge Science Park, Cambridge, UK

ABSTRACT: The bromodomain and extra-terminal (BET) protein BRD4 regulates gene expression via recruitment of transcriptional regulatory complexes to acetylated chromatin. Like other BET proteins, BRD4 contains two bromodomains, BD1 and BD2, that can interact cooperatively with target proteins and designed ligands, with important implications for drug discovery. Here, we used NMR spectroscopy to study the dynamics and interactions of the isolated bromodomains, as well as the tandem construct including both domains and the intervening linker, and investigated the effects of binding a tetra-acetylated peptide corresponding to the tail of histone 4. Using ^{15}N spin relaxation, we determined the global rotational diffusion correlation times and residue-specific order parameters for BD1 and BD2. Isolated BD1 is monomeric in the apo state, but apparently dimerizes upon binding the tetra-acetylated peptide. Isolated BD2 partially dimerizes in both the apo and peptide-bound states. By contrast, in the tandem construct only BD2 shows significant dimerization. In agreement with this result, the peptide affinity is lower for both domains in the tandem construct than for the isolated domains. The backbone order parameters reveal marked differences between BD1 and BD2, primarily in the acetyl-lysine binding site where the ZA loop is more flexible in BD2. Peptide binding reduces the order parameters of the ZA loop in BD1, and the ZA and BC loops in BD2. The AB loop, located distally from the binding site, shows variable dynamics that reflect the different dimerization propensities of the domains. These results provide a basis for understanding target recognition by BRD4.

Epigenetic regulation of gene expression involves switching between chromatin conformations that are either compact, in which gene expression is silenced, or open, in which the transcriptional machinery can access DNA. Post-translational modification of histones constitutes an important determinant of such regulation that responds to physiological and environmental signals. Epigenetic 'writer' and 'eraser' enzymes introduce and remove, respectively, post-translational modifications of histones, while 'reader' domains recognize the modifications and aid in initiating transcription through various modes of action.[1] Acetylation of lysine side chains on histone tails is a central example of post-translational modification that is recognized by the 'reader' bromodomain (BD).[2] Among the many BDs present in the human genome, the BDs of the bromodomain and extra-terminal (BET) family of proteins have emerged as important class of transcriptional coactivators involved in cell cycle progression, transcriptional activation and elongation. In particular, the BET protein BRD4 can bind, not only to acetylated histones, but also directly to various transcription factors in an acetylation-dependent manner.[3,4] BRD4 is involved in transcription of oncogenes and pro-inflammatory cytokines and chemokines, making it an important target for the treatment of several diseases, including inflammation and cancer,[5–7] using small-molecule inhibitors of BDs.[4]

BET proteins contain two 110-residue long BDs, denoted BD1 and BD2, in addition to the extra-terminal domain, which is located C-terminally of the BDs. The domains are separated by long unstructured segments; in BRD4, the segment between BD1 and BD2 is roughly 180 residues long. The BET BDs show preference for binding di-acetylated peptides with the acetylated lysine (Kac) residues close in sequence.[4] The BD structure is an antiparallel bundle of four α-helices (α_Z, α_A, α_B, and α_C), where the two inter-helical loops (ZA and BC) at one end of the molecule form a hydrophobic pocket that binds the Kac-containing peptide motifs (Figure 1). Crystal structures have revealed the detailed interactions between BD residues and the Kac peptides.[4] The ZA-loop, comprising 11 residues (85–95) in BD1 and 16 residues (373–388) in BD2, is considerably longer and known to be much more flexible than the BC-loop, comprising 5 residues in both BD1 (140–144) and BD2 (433–437). Molecular dynamics simulations have indicated that the dynamics of residues in the BC- and ZA-loops lead to switching between occluded and open binding sites that are important for binding.[6,8–11] These and other observations have led to the concept that dynamics, rather than structure, is key to achieving inhibitor selectivity between BD1 and BD2.[8,12,13]

The mechanistic significance of the tandem arrangement of BDs has not been resolved fully, but it appears that BD1

alone is sufficient to bind BET proteins to chromatin and maintain steady-state gene expression, whereas both BD1 and BD2 are required to achieve rapid increase in gene expression in response to inflammatory signals.[7] While chromatin binding is known to involve each BD individually,[13,17] tandem BDs have been implicated in binding multiple acetyl-lysine containing targets at different points during transcription in a coordinated way,[18,19] and BD-mediated dimerization of BRD4 on chromatin has been detected in vivo.[20] Bivalent inhibitors of BET bromodomains bind simultaneously to BD1 and BD2,[19,21,22] demonstrating conformational flexibility of the intermediate linker region connecting the two domains. These observations indicate important functional roles of dimerization and flexibility of the tandem BD arrangement, but little is currently known about the extent of inter-domain flexibility, inter-domain interactions, or the dynamic consequences of binding Kac peptides corresponding to histone tails.[23] In order to reach a complete understanding of the role of dynamics in BRD4 function, it is thus critical to address not only the internal dynamics of individual BDs, but also the dynamics of the intact tandem BDs, including the relative orientational dynamics of the BDs and the dynamics of inter-domain segment.

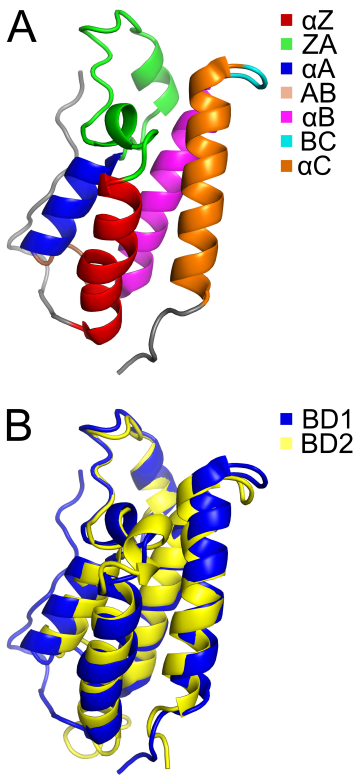


Figure 1. BRD4 bromodomain structures. (A) Overview of bromodomain structure, exemplified here by BD1 (PDB-id: 4CLB[14]), with the different secondary structure elements highlighted by color. The binding site is located at the top of the structure in this view. (B) Ribbon representation of superimposed structures of BD1 (blue, PDB-id: 4CLB[14]) and BD2 (yellow, PDB-id: 2LSP[15]). Bound ligands are not shown. The figure was prepared using PyMOL.[16]

Here, we report a comparative analysis of the ligand-binding properties and conformational dynamics of the tandem BDs from BRD4, as well as the isolated domains BD1 and BD2. Using NMR spectroscopy, we measure domain-specific affinities for a tetra-acetylated H4 histone peptide in the context of the tandem arrangement, and characterize the conformational dynamics of the BDs, as well as the linker region between them. Rotational diffusion correlation times reveal differences between BD1 and BD2 in their propensities to form dimers. Our results show that tetra-acetylated H4 histone peptides bind to the individual domains, rather than forming bivalent complexes involving both BD1 and BD2. The results pinpoint differences between the two BDs in their conformational dynamics on both fast (picosecond to nanosecond) and slow (microsecond to millisecond) timescales, most notably involving the ZA- and BC-loops. Furthermore, the two domains respond differently to binding Kac peptides.

MATERIALS AND METHODS

Protein expression and purification. BRD4 genes were cloned into pET28 expression vectors containing an N-terminal His_6-tag followed by a tobacco etch virus (TEV) protease site. Sequences for the individual constructs covered residues N44 to E168 for the N-terminal bromodomain, BD1; H341 to E460 for the C-terminal bromodomain, BD2; and N44 to E460 for the tandem BRD4(1,2). The TEV protease-digested BD1 and BRD4(1,2) constructs retained four non-native residues (G40, S41, H42, M43) prior to the native N44, whereas the BD2 construct retained a non-native Gly-Gly sequence prior to its native H341. Expression and purification closely followed the protocol described previously.[21] Uniform labelling with ^{15}N and ^{13}C isotopes was achieved by expression in minimal M9 medium with $^{15}NH_4Cl$ (Sigma Aldrich) and ^{13}C glucose (Cambridge Isotope Laboratories) as the sole sources of nitrogen and carbon, respectively. In addition, ^{15}N and ^{13}C labelled Celtone medium (Cambridge Isotope Laboratories) was supplemented to the growth medium at 5 g/L. For perdeuteration, either glucose-d_7 for U-[^2H,^{15}N] labelling or ^{13}C-glucose-d7 for [^2H,^{13}C,^{15}N] labelling was used in a M9/D_2O medium with supplements of ^2H variants of the Celtone Base Powder.

Size Exclusion Chromatography. Size exclusion chromatography was performed using Superdex 75 resin in a 3.2/300 column (Cytiva). The buffer contained 20 mM 4-(2-hydroxyethyl)piperazine-1-ethanesulfonic acid (HEPES) pH 7.4, 100 mM NaCl, and 1 mM tris(2-carboxyethyl)phosphine (TCEP). Samples were loaded as either 50 µl of 135 µM or 20 µl of 350 µM protein solution. Retention times were calibrated using 10 µl of Bio-Rad gel filtration standards.

NMR sample preparation. The NMR samples contained 0.135 mM protein, i.e., BD1, BD2 or BD4(1,2), dissolved in a buffer comprising 20 mM Na_2HPO_4, 1 mM TCEP, and

7%/93% D_2O/H_2O at pH 6.8. Peptide-bound samples contained in addition 0.90 mM H4Kac4 in the case of BD1 and BD2, or 1.5 mM H4Kac4 in the case of BRD4(1,2).

NMR spectroscopy. All NMR experiments were performed at 30 °C on Bruker AV 600 and AVIII 800 spectrometers equipped with 5-mm z-gradient $^1H/^{13}C/^{15}N$ TCI cryoprobes. Temperature calibration was performed with a 99.8% methanol-d_4 sample.[24] Proton chemical shifts were referenced to 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS), whereas ^{15}N and ^{13}C chemical shifts were indirectly referenced as described.[25] The assignment strategy for the backbone $^1H^N$, ^{15}N, $^{13}C'$, $^{13}C^\alpha$ and side-chain $^{13}C^\beta$ chemical shifts followed standard triple-resonance experiments.[26] Backbone assignments for isolated BD1 were obtained using ^1H-^{15}N SOFAST HMQC,[27,28] 3D CBCANH,[29] CBCA(CO)NH,[30] HN(CO)CA,[31] HNCA,[32] and HNCO[32] experiments. All subsequent triple-resonance experiments contained the TROSY-based detection scheme.[33] Backbone assignments for isolated BD2 are available from BMRB accession numbers 15057, 18439 and 19738. The BD2 assignments were manually confirmed using a 3D HNCACB experiment. The backbone assignments for the tandem bromodomain construct BRD4(1,2) were obtained using three-dimensional HNCACB, HN(CO)CACB, HNCA, HN(CO)CA, HNCO, HNCACO, (H)N(COCO)NH and (HN)CO(CO)NH experiments.[34] All multi-dimensional experiments were acquired using a non-uniform sampling scheme with Poisson gap distribution as described.[35] Spectra were processed with NMRPipe[36] and analyzed with CcpNmr analysis.[37] Non-uniformly sampled spectra were reconstructed using the istHMS method.[35] Figures containing NMR spectra were plotted using the Python based program nmrglue.[38]

NMR ^{15}N relaxation experiments. TROSY-based ^{15}N relaxation experiments[39] were performed at static magnetic field strengths of 14.1 T and 18.8 T. ^{15}N R_1 relaxation experiments were acquired with delays of: 100, 200, 300, 400, 500, 600, 800, 1000, 1300, 1600, 1900 and 2200 ms at 18.8 T; and 100, 200, 300, 400, 500, 600, 800, 1200, 1500, 1600 and 2000 ms at 14.1 T. ^{15}N R_2 relaxation experiments were acquired with delays of: 15.7, 31.4, 47.0, 62.7, 78.4, 94.1, 125.4, 141.1, 156.8, 188.2 and 203.8 ms at both 14.1 T and 18.8 T. Steady-state $\{^1H\}$-^{15}N-heteronuclear *NOE* data were measured from pairs of interleaved spectra recorded with or without ^1H saturation during the 7.0 s recycle delay, denoted *NOE* and control, respectively. ^1H saturation was applied as a train of high-power 120° pulses.[40] Transverse cross-correlation relaxation rate constants (η_{xy}) were measured as the difference in the relaxation rates of the TROSY and anti-TROSY components of the NH doublet,[41,42] using relaxation delays of (2, 4, 6, 15, 25, 40, 50) ms with duplicate data acquired at 15 ms and 25 ms. ^{15}N R_1, R_2 and η_{xy} relaxation rate constants were obtained by non-linear least squares fitting of peak intensities at measured relaxation delays, as implemented in the program relax,[43] whereas the steady-state $\{^1H\}$-^{15}N *NOE* values were calculated from peak intensity ratios obtained from spectra acquired in the presence and absence of proton saturation, with uncertainties in peak intensities estimated from the baseplane noise. Uncertainties of the relaxation rates were obtained using duplicate delays and standard errors were estimated from a sample of 500 Monte Carlo simulations of the uncertainties for each dataset.[44] Trimmed averages were calculated in MATLAB by first calculating the mean and standard deviation of the full data set. Data points outside of one standard deviation from the mean where then removed, and a new mean calculated from the remaining data points.

Model-free relaxation data analysis. The extended model-free formalism[45–48] was used to analyze the ^{15}N relaxation data for BRD4 with the method for the combined optimization of the global diffusion tensor and local model-free parameters implemented in the relax program (version 3.3.6).[49] The analysis assumed an N-H bond length (r_{NH}) of 1.02 Å and a CSA ($\Delta\sigma_{CSA}$) of –172 ppm. The N-H bond vector orientations were extracted from the BD1 X-ray and BD2 NMR derived structures, PDB entries 4CLB[14] and 2LSP[15], respectively. The ligands in the PDB-structures were removed prior to analysis. All residues included in the analysis were represented by relaxation data recorded at both field strengths.

Diffusion tensor analysis. The MATLAB based version of the rotdif program[50] was modified for use with η_{xy} transverse cross-correlated relaxation rates to estimate the overall rotational diffusion tensor. This approach avoids potential problems caused by exchange contributions to the transverse relaxation rate, which can occur in R_2 but not in η_{xy}. The ratio of the spectral densities $J(\omega_N)$ and $J(0)$ is calculated as:

$$\left(\frac{\eta_{xy}}{R_1'\,\kappa} - 1\right)^{-1} = \frac{3}{4}\frac{J(\omega_N)}{J(0)}$$

where

$$\kappa = -\frac{cdP_2(\theta)}{(c^2 + d^2)}$$

and d is the dipolar coupling constant $(\mu_0 h\gamma_H\gamma_N)/(8\pi^2 r^3_{NH})$, μ_0 is the permeability of free space, h is Planck's constant, γ_H and γ_N are the gyromagnetic ratios of hydrogen and nitrogen, respectively, r_{NH} is the N–H bond length (1.02 Å), c is the CSA coupling constant $(\omega_N\Delta\sigma_{CSA})/3$ with $\Delta\sigma_{CSA}$ = –172 ± 20 ppm, $P_2(\theta)$ is the second order Legendre polynomial, and θ = 15 ± 10° is the angle between the N-H internuclear vector and the unique axis of the chemical shielding tensor. The values of $\Delta\sigma_{CSA}$ and θ correspond to conservative averages taken from the literature.[51–53] R_1' is the longitudinal relaxation corrected for high-frequency spectral density:[54]

$$R_1' = R_1 - 7\left(\frac{0.921}{0.87}\right)^2 C_{HF} = 3(c^2 + d^2)J(\omega_N)$$

where the correction factor (C_{HF}) for the high frequency components is calculated under the assumption that $J(\omega) \propto \omega^{-2}$ at $\omega \approx \omega_H$, yielding

$$C_{HF} = -(\gamma_N/5\gamma_H)(1 - NOE)R_1 = d^2 J(0.87\omega_H)$$

The HYDRONMR software[55,56] was used to calculate diffusion tensors and correlation times for the isolated BD1 and BD2 domains, based on PDB structures 4CLB[14] and 2LSP,[15] respectively. The ligands in the PDB structures were removed prior to analysis. The temperature was set to 303 K and the solvent viscosity to $798\cdot10^{-6}$ Pa s. An effective atomic radius of 3 Å was used in accordance with Halle &

Davidovic.[57] The three principal values of the diffusion tensor are defined as D_{xx}, D_{yy}, and D_{zz}, with $D_{zz} \geq D_{yy} \geq D_{xx}$. Further, the rotational diffusion correlation time (τ_c) and diffusion anisotropy ($D_{\parallel}/D_{\perp}$) are obtained from the relationships:

$$D_{iso} = \frac{1}{3}\left(D_{xx} + D_{yy} + D_{zz}\right) = \frac{1}{6\tau_c}$$

$$\frac{D_{\parallel}}{D_{\perp}} = \frac{2D_{zz}}{(D_{xx} + D_{yy})}$$

Spectral density mapping. Spectral density mapping[58] was performed using data obtained at 14.1 T and 18.8 T with approximate expressions for $J(0.921\omega_H)$ and $J(0.955\omega_H)$ obtained by extrapolation from the static magnetic field dependence of the relaxation data using a first order Taylor series expansion of $J(\omega)$ at the $0.870\omega_H$ frequency[58]

$$J(\varepsilon\omega_H) = J(0.870\omega_H) + (\varepsilon - 0.870)\omega_H J'(0.870\omega_H)$$

where ε = 0.921 or 0.955, and $J'(0.870\omega_H)$ is the first derivative of $J(\omega)$ at the $0.870\omega_H$ frequency, estimated from the difference in $J(0.921\omega_H)$ between the different static magnetic field strengths. The spectral density mapping calculations used $\Delta\sigma$ = –173 ± 7 ppm and r_{NH} = 1.04 Å.[51]

Exchange contributions (R_{ex}) to R_2 were estimated by comparing auto- and cross-correlated relaxation rates, following published protocols.[59] In this approach R_{ex} can be estimated at a specified static magnetic field strength, (B_0) as:

$$R_{ex} = \Gamma_{auto} - \Gamma_{cross}$$

Γ_{auto} and Γ_{cross} were calculated from R_1, R_2, *NOE*, and η_{xy} data acquired at 18.8 T:

$$\Gamma_{auto} = R_2 - \frac{1}{2}R_1 - \left(\frac{3d^2}{4}\right)J(\omega_H) =$$

$$\left(\frac{d^2}{2}\right)J(0) + \left(\frac{2\gamma_N^2\Delta\sigma_{CSA}^2 J(0)}{9} + \Theta_{ex}\right)B_0^2$$

$$\Gamma_{cross} = \left(\frac{3d^2 + 4c^2}{4\sqrt{3}cdP_2(\cos\theta)}\right)\eta_{xy} - \frac{1}{2}R_1 + \frac{7}{8}d^2J(0.92\omega_H)$$
$$= \left(\frac{d^2}{2}\right)J(0) + \left(\frac{2\gamma_N^2\Delta\sigma_{CSA}^2 J(0)}{9}\right)B_0^2$$

where the spectral densities $J(\omega_H)$ and $J(0.92\omega_H)$ were obtained by spectral density mapping as described above, using θ = 19.6° ± 2.5°,[51] and $\Theta_{ex} = R_{ex}/B_0^2$. This analysis is sensitive to errors arising from site-specific variations in CSA.

Errors were propagated using Monte Carlo simulations.[44] For each residue 10,000 log-normal distributed points where randomly generated for each variable (R_1, R_2, *NOE*, η_{xy}, $\Delta\sigma_{CSA}$, etc.) using a standard deviation equal to the error estimated from the fitted relaxation data.

***K*d determination.** A peptide mimicking a tetra-acetylated histone 4 (H4) tail, comprising residues 1–16 of H4 with Nε-acetylation at K5, K8, K12, and K16 (denoted H4-Kac5Kac8Kac12Kac16) was purchased from Cambridge Research Biochemicals (Cambridge, UK). The peptide was dissolved at a concentration of 10 mM in 0.2 M Na_2HPO_4 pH 6.8. For binding titrations, all NMR samples contained 0.065 mM of U-[^{2}H,^{15}N]-labeled BRD4(1,2) or U-[^{13}C,^{15}N]-labeled BD1/BD2. The ^{1}H and ^{15}N chemical shift changes were followed by collecting ^{1}H-^{15}N TROSY experiments at six ligand concentrations: (30, 60, 119, 235, 461, and 889) μM. The combined chemical shift perturbation was calculated using the equation

$$\Delta\delta = \sqrt{(\Delta\delta_{HN})^2 + (\frac{1}{5}\Delta\delta_N)^2}$$

where $\Delta\delta_{HN}$ and $\Delta\delta_N$ denote the chemical shift differences between the peptide-bound and apo states for $^1H^N$ and ^{15}N, respectively.

We measured the dissociation constants for the interaction of BRD4 BDs with different ligands by monitoring the chemical shift changes of BRD4 BDs from the apo to the peptide-bound form during titration. When the exchange rate is greater than the chemical shift difference between the free and bound states, the observed chemical shift perturbation at each titration point ($\Delta\delta_{obs}$) is the population weighted average between the chemical shifts of the free and bound states obtained by the following mass action binding isotherm equation for binding to a single site (valid for BD1 and BD2):

$$\Delta\delta_{obs} = (\delta_b - \delta_f)\frac{(K_d + P_t + L_t) - \sqrt{(K_d + P_t + L_t)^2 - 4P_tL_t}}{2P_t}$$

where δ_b and δ_f are the chemical shifts of the bound and free states, respectively; P_t and L_t are the total concentrations of protein and ligand, respectively, at each titration point; and K_d the dissociation constant. In the case of BRD4(1,2), the chemical shift perturbations were analyzed using the corresponding coupled equations valid for simultaneous binding to two sites (1 and 2):

$$\Delta\delta_{obs,1} = (\delta_{b,1} - \delta_{f,1})\frac{L}{K_{d,1} + L}$$

$$\Delta\delta_{obs,2} = (\delta_{b,2} - \delta_{f,2})\frac{L}{K_{d,2} + L}$$

where L is the concentration of free ligand, $L = L_t - P_1L - P_2L$, and P_iL indicates the concentration of protein–ligand complex with the ligand bound to site i, which is obtained as

$$P_iL = \frac{P_t\, e}{(3\, K_{d,i} + e)}$$

where

$$e = 2(a^2 - 3b)^{1/2}\cos(d/3) - a$$

$$a = K_{d,1} + K_{d,2} + 2P_t - L_t$$

$$b = K_{d,1}K_{d,2}$$

$$d = \text{acos}\left(\frac{-2a^2 + 9ab - 27c}{2(a^2 - 3b)^{3/2}}\right)$$

$$c = b\,L_t$$

We performed global, non-linear fits of the above equations to the experimental titration data using the GraFit package version 6.0.5 (Erithacus Software Ltd., Staines).

RESULTS AND DISCUSSION

We used NMR spectroscopy to investigate ligand interactions, intramolecular dynamics, and rotational diffusion of the BRD4 bromodomains. NMR makes it possible to determine the domain-specific binding affinities in the context of the tandem construct. We compared the dynamics of the isolated domains and the tandem construct, in both the free and ligand-bound states. Furthermore, we characterized the dynamics of the disordered inter-domain linker segment and its effects on the structure, dynamics and interactions of the BDs. NMR provides an opportunity to study all of these properties under identical sample conditions; this contrasts with many previous studies of BET bromodomains, which have involved multiple methods involving different conditions to investigate the interactions of BD1 and BD2. By studying ligand binding and the rotational diffusion properties under identical conditions, we arrive at a consistent model for the coupled changes in structure, dynamics and interactions upon ligand binding to BRD4 that resolves partly conflicting interpretations of previously reported results.

Chemical shift differences between the isolated and tandem domains reveal differences in domain interactions. To enable residue-specific studies of BD interactions and dynamics, we previously reported the use of segmental labeling of BRD4(1,2) to assign the backbone amide ^{1}H and ^{15}N resonances of BD1 and BD2 in the tandem domain construct, BRD4(1,2), together with the inter-domain linker.[60] Here we compare the chemical shifts of the isolated domains with those recorded for the tandem construct. Resonance assignments were aided by previously reported chemical shift data sets of the BRD4 bromodomains, e.g. Biological Magnetic Resonance Bank entries 50145 and 50146.[61] Standard triple-resonance experiments (see Materials & Methods) were used to obtain backbone assignments of the isolated domains BD1 and BD2 at a level of 98% and 93%, respectively, for non-proline residues. For BRD4(1,2) standard triple-resonance experiments performed on both uniformly and segmentally labelled protein[21,60] enabled nearly complete assignments of BD1 (88%) and BD2 (87%). However, resonance assignment of the inter-domain linker region in BRD4(1,2) was non-trivial, due to severe resonance overlap in this region (T169–S348), which has properties characteristic of an intrinsically disordered protein region, including an abundance of repeat sequences and a high proportion of proline residues. Using special triple-resonance experiments designed to establish sequential connections across proline residues and stretches with high chemical shift degeneracy,[34] we succeeded to assign the inter-domain linker region to a completeness of 69%. In summary, despite a lower overall completeness of assignments in BRD4(1,2), a high percentage of backbone assignments were obtained for the two bromodomains. As might be expected, the 2D ^{1}H-^{15}N TROSY spectra of the tandem domain BRD4(1,2) show a high degree of resonance overlap with most of the resonances from the linker region having a narrow chemical shift dispersion in the ^{1}H dimension reflecting its largely unstructured nature (Fig. 2a). In addition, there is significant variation in the peak intensities, indicating non-uniform dynamics in BRD4(1,2).

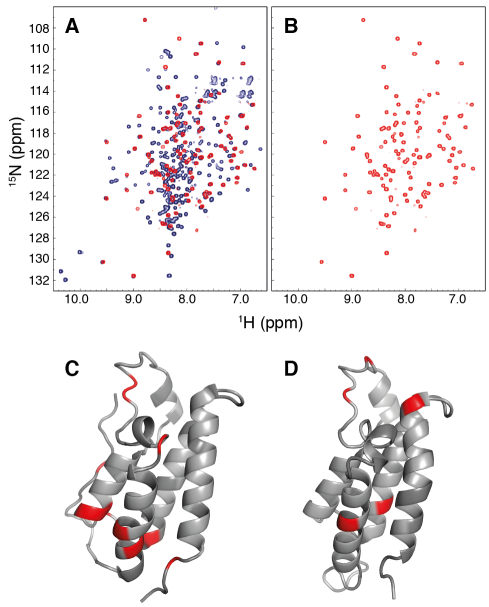


Figure 2. ^{15}N TROSY spectrum of the tandem BRD4-construct and chemical shift differences between isolated and tandem constructs. (A) Superimposed TROSY spectra for uniformly ^{15}N-labelled BRD4(1,2) (blue) and BD2 ^{15}N-labelled BRD4 (red), (B) BD2 ^{15}N-labelled BRD4 spectrum (red; identical to panel A). (C, D) Absolute backbone amide chemical shift differences ($\Delta\delta$) between tandem and isolated constructs mapped onto the structures of (C) BD1, PDB-id 4CLB,[14] and (D) BD2, PDB-id 2LSP.[15] Residues with $\Delta\delta > 0.1$ are colored red. Panels C and D were prepared using PyMOL.[16] FIGURES (Word Style "VA_Figure_Caption"). Each figure must have a caption that includes the figure number and a brief description, preferably one or two sentences. The caption should follow the format "Figure 1. Figure caption." All figures must be mentioned in the text consecutively and numbered with Arabic numerals. The caption should be understandable without reference to the text. Whenever possible, place the key to symbols in the artwork, not in the caption. To insert the figure into the template, be sure it is already sized appropriately and paste before the figure caption. For formatting double-column figures, see the instructions at the end of the template. Do NOT modify the amount of space before and after the caption as this allows for the rules, space above and below the rules, and space above and below the figure to be inserted upon editing.

The NMR spectra reveal chemical shift differences between BD1 in the isolated and tandem constructs. The overall chemical shift difference is 0.10 ppm with greater differences ($\Delta\delta$ = 0.37–0.45 ppm) observed for residues N54,

K55, Y118, W120, N121, and A122. These residues are all located adjacent to the C-terminus of BD1 and residues 118–122 are situated in the AB-loop, showing that the linker affects this region of BD1 in the tandem construct (Fig. 2c). In the case of BD2, the chemical shift differences are much smaller with an average of 0.03 ppm, but still indicate perturbations ($\Delta\delta$ = 0.11–0.17 ppm) in primarily three regions around residues I394, N428, K445, Q447, and E451. These residues are all located on the same side of the helix bundle (Fig. 2d). Interestingly, this region has previously been implicated in homo-dimerization based on homology with the BD1 domain of BRD2, which is known to form homo-dimers. [62,63] Thus, the differences in chemical shifts between the isolated domain and tandem construct might indicate that the presence of the linker affects the tendency of BD2 to form homo-dimers; however, the chemical shift data alone cannot resolve which species is more prone to dimerize. We address this issue further below using relaxation measurements.

Binding of tetra-acetylated H4 peptide to BD1 and BD2. We investigated ligand binding to the isolated and tandem bromodomains of BRD4 by monitoring chemical shift changes during titration with a tetra-acetylated peptide comprising the N-terminal sequence of histone 4 (residues 1–16) with lysine acetylation on K5, K8, K12 and K16 (denoted H4Kac4), as shown in Fig. 3. The chemical shift perturbations allowed us to identify the residues engaged in association with H4Kac4. Residues in BD1 and BD2 with chemical shift changes greater than one standard deviation above the average are primarily located in the ZA and BC loops and also in the αA, αB and αC helices (Fig. 3c, d), thereby verifying that the binding modes observed in solution generally agree with those expected from the database of available structures.[4] The spectra further showed that the exchange between free and bound forms range from slow-to-intermediate to intermediate-to-fast on the chemical shift time scale, reflecting the variation in chemical shift perturbation upon peptide binding. For instance, the highly conserved residues N140 in BD1 and N433 in BD2, which are known to be key in mediating stable interactions with the ligand, appeared in the slow-exchange regime, as a consequence of their greater change in chemical shift upon binding. The majority of the BRD4 resonances affected by binding appeared in the fast-exchange regime and thus allowed straightforward tracking of the backbone amide $^{1}H^{N}$ and ^{15}N chemical shifts from the free to the bound form during the titration, as described next. We used the titration-dependent chemical shift changes for six backbone amides of BD1 (K55, W75, W81, V87, N121 and E151) and five backbone amides of BD2 (W374, D381, C391, F426 and E438), as well as the indole NH group of W374, to calculate the dissociation constant (K_d) of the bromodomain-peptide interactions (Fig. 3; Table 1) using equations describing ligand binding to a single site, in the case of the isolated domains, or two sites simultaneously, in the case of the tandem domain BRD4(1,2). The NMR-derived affinities for H4Kac4 show significant differences between the two isolated domains, with K_d values of 9 μM for BD1 and 74 μM for BD2 (Table 1), in good agreement with previous results.[64] These dissociation constants can be compared with results obtained for H4 octapeptides mono-acetylated on either K5 or K16, which have mutually similar K_d values of roughly 300 μM and 120 μM for isolated BD1 and BD2, respectively.[65]

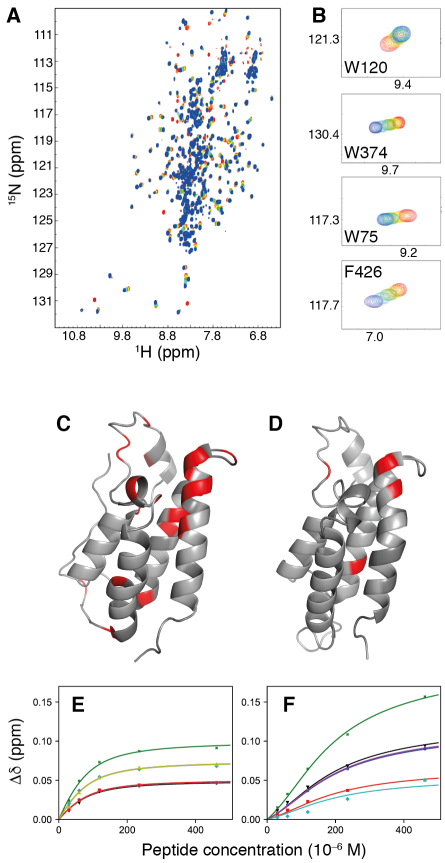


Figure 3. Binding of tetra-acetylated H4 peptide to the BRD4 bromodomains. (A) Superimposed ^{15}N-TROSY spectra of BRD4(1,2) tracking chemical shift changes during titration with the H4Kac4 peptide. (B) Close-up views of chemical shift changes observed for 4 representative residues. (C, D) Backbone amide chemical shift differences $\Delta\delta$ > 0.1 between apo and H4Kac4 bound states, colored red on the (C) BD1 structure, PDB-id: 4CLB,[14] and (D) on the BD2 structure, PDB-id: 2LSP.[15] (E, F) Binding isotherms from chemical shift perturbations ($\Delta\delta_{obs}$) of H4Kac4 binding to the BD1 (E) and BD2 (F) domains of BRD4(1,2). Panel E includes data for residues L55 (blue), W75 (purple), W81 (light green), V87 (dark green), W120 (red), N121 (teal), and G151 (black). Panel F includes data for W374 backbone (red) and side chain (black), D381 (dark green), C391 (purple), F426 (blue), and E438 (teal). Panels C and D were prepared using PyMOL.[16] C391 (purple), F426 (blue), and E438 (teal). Panels C and D were prepared using PyMOL.[16] C391 (purple), F426 (blue), and E438 (teal). Panels C and D were prepared using PyMOL.[16]

Similar results have been obtained for 16-mer H4 peptides mono-acetylated on K5 or K12, with K_d of 600 μM and 1 mM.[23] Thus, the comparison confirms that multiple

acetylation of the histone peptide leads to higher affinity for BRD4.[64,66] Furthermore, higher affinity of BD1, compared to BD2, has also been observed for di-acetylated transcription factor motifs.[61] In the tandem construct BRD4(1,2), the K_d values of the individual domains binding to H4Kac4 are 23 μM for BD1 and 125 μM for BD2 (Fig. 3e, f; Table 1). Thus, the affinity is higher for BD1 than BD2 in both the case of the individual domains and the tandem construct. The reduction in H4Kac4 affinity of each domain in BRD4(1,2) compared to the isolated domains, amounts to a modest reduction in the free energy of binding: 2.2 kJ/mol for BD1 and 1.3 kJ/mol for BD2, which can be interpreted as unfavorable coupling free energy between the domains in BRD4(1,2). BRD4 shows higher affinity for tetra-acetylated H4 motifs than for peptides with lower levels of acetylation, which has been ascribed to avidity effects.[64] This concept is supported by bivalent binding of a single peptide or synthetic ligand to two BDs in several crystal structures.[19,21,22,64,67,68] However, the unfavorable coupling free energy measured for both BD1 and BD2 in the tandem construct argues against this interpretation in the case of H4Kac4 binding to BRD4(1,2). Nonetheless, bivalent binding might play a role in the case of the isolated domains, because the reduced peptide affinity observed for the individual domains of BRD4(1,2) might reflect a reduced tendency to form bivalent complexes in the tandem construct, compared to the isolated domains. Furthermore, several residues exhibit differences between the isolated and tandem constructs in their chemical shift perturbations upon peptide binding, which might be explained by differential dimer formation, or possibly by interactions with the linker that might be altered by peptide binding. We return to this issue of dimerization below.

Table 1. Dissociation constants for H4Kac4-bromodomain complexes measured by NMR chemical shift perturbation.

Construct	K_d (10^{-6} M) BD1	K_d (10^{-6} M) BD2
Isolated[a]	9.1 ± 1.3	74.0 ± 1.6
tandem[b]	22.9 ± 2.1	125.1 ± 10.0

[a] determined using equations for one-site ligand binding. [b] determined using equations for two-site ligand binding.

Bromodomain dynamics: overview of ^{15}N relaxation measurements. To explore in more detail the potential domain interactions and the dynamics of the domains, we characterized the molecular dynamics of BRD4 bromodomains using ^{15}N nuclear spin relaxation measurements. We performed ^{15}N R_1, R_2 and steady-state heteronuclear {^{1}H}-^{15}N *NOE* experiments at static magnetic field strengths of 14.1 T and 18.8 T to quantify picoseconds to nanosecond (ps–ns) dynamics. In addition, we acquired TROSY-based cross-correlated relaxation (η_{xy}) experiments[41] at 18.8 T to aid in estimating the overall rotational correlation times (τ_c) and exchange contributions (R_{ex}) to the transverse relaxation rates. η_{xy}, caused by interference between ^{1}H–^{15}N dipole-dipole and ^{15}N chemical shift anisotropy (CSA) interactions, is not affected by chemical exchange and hence provides an improved estimate of τ_c, while comparison of R_2 and η_{xy} provides an assessment of exchange, which is essential in the present case where fast exchange between free and peptide-bound states or monomeric and dimeric species might otherwise complicate the analysis of τ_c, as well as the internal dynamics of the individual domains. Figure 4 provides an overview of the results for apo BRD4(1,2), which clearly identifies the two domains and demonstrates that the linker region between them is highly flexible as indicated by its higher R_1, lower R_2, lower η_{xy}, and lower *NOE* values compared to the BDs. However, we note that the linker region shows significant non-monotonous variation in R_1 and *NOE* among residues, indicating that it does not behave as a simple random coil-like chain, but most likely has propensity to form more ordered structure in certain regions. Supplementary Figure S1 shows the corresponding results for peptide-bound BRD4(1,2), as well as for apo and peptide-bound BD1 and BD2.

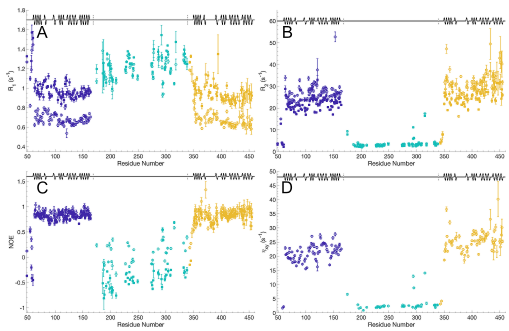


Figure 4. ^{15}N relaxation data for the tandem construct BRD4(1,2). (A) R_1 relaxation rate constants, (B) R_2 relaxation rate constants, (C) {^{1}H}-^{15}N *NOE*s, (D) η_{xy} relaxation rate constants. The secondary structure is indicated at the top of each panel, with α-helixes represented by waves and loops and the linker region as straight lines. The individual domains and the linker are indicated by color: blue, BD1; teal (green), linker; and yellow, BD2. Data acquired at 14.1 T and 18.8 T are indicated by filled squares and open circles, respectively.

Bromodomain dynamics: rotational diffusion of the BD1 and BD2 domains. We determined the rotational diffusion tensors of the ligand-free (apo) states of BD1 and BD2 domains in both the isolated and tandem constructs. The trimmed and weighted averages of the measured relaxation rates (Table S1) give a first indication of the relative difference in overall tumbling time (τ_c) of the different constructs. The R_1 relaxation rate is proportional to $1/\tau_c$, and the η_{xy} relaxation rate is proportional to τ_c. The R_2 relaxation rate can also be used to evaluate the global diffusion time, but it includes contributions from conformational exchange on the microsecond to millisecond timescales, R_{ex}, which complicate the analysis. For both the apo and peptide-bound forms of isolated BD2, the average values of R_1 and η_{xy} are significantly lower and higher, respectively, than the corresponding values for the isolated BD1, indicating slower global tumbling of BD2 compared to BD1. This difference persists in apo and peptide-bound BRD4(1,2), although it is slightly attenuated. The observed difference in τ_c is unexpected given that the isolated BD1 and BD2 domains

have the same molecular weight (17.5 kDa) and similar tertiary structures consisting of α-helical bundles (Fig. 1). Indeed, hydrodynamics calculations performed using HydroNMR[55,56] predict diffusion tensors with effective τ_c values of 7.9 ns for BD1 and 8.3 ns for BD2 (Table 2). We estimated the diffusion tensors of the different constructs using a modified version of the rotdif program,[50] which takes as input the measured R_1, *NOE* and η_{xy} relaxation rates, thereby avoiding exchange contributions to the transverse relaxation rate. It should be noted that diffusion analysis of the bromodomains is challenging, because most residues are located in α-helices with their ^{15}N-^{1}H bond vectors pointing along the largest principal axis of the anisotropic diffusion tensor (Fig. S2), and the scarcity of bond vectors oriented perpendicular to the unique diffusion axis leads to uncertainty in the estimated values. First, we analyzed the relaxation rate constants for the isolated apo states. The results for isolated apo BD1 indicate an anisotropic diffusion tensor with $D_{||}/D_{\perp} = 1.56 \pm 0.02$ and $\tau_c = 7.6 \pm 0.9$ ns, in good agreement with the value expected from the HydroNMR calculations (Table 2). By contrast, the results for isolated apo BD2 yield $D_{||}/D_{\perp} = 1.39 \pm 0.01$ and $\tau_c = 10.1 \pm 0.9$ ns. Notably, τ_c is considerably greater than the expected value, suggesting partial dimerization of this domain. Furthermore, the lower value of $D_{||}/D_{\perp}$ also suggests partial formation of a side-by-side dimer (which is expected to have a more spherical shape than the monomer) in line with previous hypotheses based on crystal structures[62,63] and the chemical shift differences between the isolated and tandem constructs described above. Using the dimeric structure reported for BD1 of BRD2, PDB id 2DVQ,[63] as a model for the tentative BD2 dimer, HydroNMR calculations predict a slightly greater value, $\tau_c = 13.8$ ns, than the experimentally determined one, indicating that BD2 of BRD4 is exchanging between monomeric and dimeric states.

Table 2. Diffusion tensor parameters of BRD4 bromodomains.[a]

Rotdif calculations				
State (construct)	D_{iso} (10^7 s^{-1})	$D_{||}/D_{\perp}$	τ_c (ns)	χ^2_{red}
apo BD1 (isolated)	2.2 ± 0.3	1.56 ± 0.02	7.6 ± 0.9	1.5
apo BD2 (isolated)	1.7 ± 0.1	1.39 ± 0.01	10.1 ± 0.9	2.5
apo BD1 (tandem)	1.3 ± 0.3	1.56 ± 0.01	13 ± 3	3.0
apo BD2 (tandem)	1.2 ± 0.3	1.59 ± 0.02	14 ± 3	4.8
H4Kac4 BD2 (isolated)	1.4 ± 0.1	1.59 ± 0.02	11.9 ± 0.9	2.8
H4Kac4 BD1 (tandem)	1.0 ± 0.3	1.69 ± 0.01	16 ± 4	3.9
H4Kac4 BD2 (tandem)	0.1 ± 0.4	1.92 ± 0.01	16 ± 7	5.4
HydroNMR calculations				
State (construct)	D_{iso} (10^7 s^{-1})	$D_{||}/D_{\perp}$	τ_c (ns)	χ^2_{red}
apo BD1 (monomer)[a]	2.1	1.4	7.9	n. a.[f]
apo BD2 (monomer)[b]	2.0	1.8	8.3	n. a.[f]
H4K12ac BRD2-BD1 (dimer)[c]	12	0.83	13	n. a.[f]
H4K5acK8ac BD1 (monomer)[d]	2.0	1.4	8.4	n. a.[f]
H4K8acK12ac BD1 (dimer)[e]	0.73	1.12	33	n. a.[f]

[a]PDB-ID: 4CLB,[14] [b]PDB-ID: 2LSP,[15] [c]PDB-ID: 2DVQ,[63] [d]PDB-ID: 3UVW,[64] [e]PDB-ID: 3UW9,[64] [f]not applicable

We validated the results for the apo forms of isolated BD1 and BD2 by size-exclusion chromatography (SEC), which showed that isolated apo BD1 is monomeric, whereas isolated apo BD2 elutes as a larger protein than expected and this effect is further pronounced at higher concentration, consistent with partial dimerization (Table 3). The result for BD2 contrasts with previous interpretations of SEC data for the wild-type and mutant forms, designed to disrupt the dimer interface, which suggested that BD2 is monomeric despite eluting as a larger species;[65] this study also did not detect any signs of heterodimer formation between BD1 and BD2. Previous ^{15}N NMR relaxation results have also suggested that BD2 has a greater hydrodynamic radius than does BD1 (in agreement with our results), but analytical ultracentrifugation experiments performed in the same study indicated that the domain is monomeric.[23] However, we note that analytical ultracentrifugation was conducted with significantly lower protein concentrations (by factors of 2–10) than those used in the NMR study. In addition, differences in sequence length (construct size) among the studied systems could play a role. For example, Liu et al. used a shorter construct of BD2, comprising residues 352–457 (compared to our version comprising 341–460), which lacks residues next to the proposed dimer interface.[62]

Next, we characterized the rotational diffusion properties of the domains in the context of the tandem construct, BRD4(1,2). The best-fit diffusion tensors of the individual domains in BRD4(1,2) are $D_{||}/D_{\perp} = 1.56 \pm 0.01$ and $\tau_c = 13 \pm 3$ ns for BD1, and $D_{||}/D_{\perp} = 1.59 \pm 0.02$ and $\tau_c = 14 \pm 3$ ns for BD2. The anisotropy of each domain in BRD4(1,2) is indistinguishable from that of the isolated BD1 domain, indicating that partial dimer formation of BD2 is significantly

reduced in the tandem construct, compared to the isolated domain, or altogether abolished. The higher value of τ_c for BD1 in BRD4(1,2) compared to the isolated domain is explained by the motional restriction imparted by the connection of the two domains via the linker.[69,70] The slightly greater τ_c value for the BD2 domain, compared to BD1, in BRD4(1,2) is in line with the difference in τ_c values predicted by HydroNMR for the isolated domains.

Table 3. Molecular size of different apo BRD4 constructs measured by size exclusion chromatography.

State (construct)	Concentration (μM)	Measured M_w (kDa)	Predicted M_w (kDa)
apo BD1 (isolated)	135	17.8	17.5
apo BD2 (isolated)	135	25.1	17.5
apo BRD4(1,2)	135	141.2	49.3
apo BD1 (isolated)	350	17.8	17.5
apo BD2 (isolated)	350	28.1	17.5
apo BRD4(1,2)	350	112.2	49.3

We studied the effect of peptide binding on the rotational diffusion of the bromodomains. In the H4Kac4-bound state, the isolated BD2 domain has $D_{||}/D_{\perp} = 1.59 \pm 0.02$ and $\tau_c = 11.9 \pm 0.9$ ns, reflecting a modest increase in τ_c compared to the partially dimeric apo state. We did not record η_{xy} for isolated peptide-bound BD1, precluding the rotational diffusion analysis described above, but model-free analysis (based on R_1, R_2, and *NOE*; see below) yields $D_{||}/D_{\perp} = 1.8$ and $\tau_c = 12.3$ ns, similar to the results for BD2 (SI Table S2). These results might suggest that the tetra-acetylated peptide binds to the isolated domains in a bivalent mode, thereby inducing dimerization. As noted above, a number of crystal structures of bromodomains have revealed bivalent binding of peptides[64,67,68] and synthetic inhibitors.[19,21,22] These structures show a great deal of variation in the relative orientation of the two domains, including side-by-side and fully extended head-to-head orientations. In all cases, the τ_c values predicted by HydroNMR (Table 2) for these dimers are considerably greater than the experimentally determined one, indicating that the peptide-bound form of isolated BD1 dimerizes transiently, similar to isolated BD2 (which apparently does so in both the apo and bound forms).

In the tandem construct, the peptide-bound BD1 and BD2 domains are characterized by $D_{||}/D_{\perp} = 1.69 \pm 0.01$ and $\tau_c = 16 \pm 4$ ns and $D_{||}/D_{\perp} = 1.92 \pm 0.01$ and $\tau_c = 17 \pm 7$ ns, respectively. The relatively large uncertainties in these results unfortunately precludes any firm assessment of potential dimer formation in this case.

Internal dynamics of BD1 and BD2 and effects of peptide binding. We analyzed the internal dynamics on the ps–ns timescale using the model-free (MF) formalism[45,47,71] and the dynamic effects of binding H4Kac4 to each domain. The MF analysis included the order parameters (S^2, S_f^2) and effective correlation times (τ_e, τ_s) of the sub-ns internal motion, together with the overall rotational diffusion (τ_c and $D_{||}/D_{\perp}$), while slower motions were treated simply as exchange contributions (R_{ex}) to R_2. MF analysis of the bromodomains is hampered to some extent by the limited range of ^{15}N-^{1}H bond vector orientations sampling the diffusion tensor, as described above. Furthermore, exchange between monomeric and dimeric states results in population-weighted averages of relaxation rates associated with the two different diffusion tensors,[72] which presents a potential caveat for the MF analysis of BD2 since the detailed structure of the hypothetical dimer is unknown. However, three reasonable assumptions make the analysis tractable: first, the N-H bond vector orientations in the molecular frame do not change upon dimer formation; second, the order parameter is identical in the monomer and dimer; and third, the diffusion tensor of the dimer is nearly isotropic. With these assumptions, MF analysis can be performed on BD2, while recognizing that the determined diffusion tensor principal values represent an effective apparent tensor. The MF analysis is further dependent on the quality of the structural models, because errors in the N–H bond vector orientations in the principal axis frame of the diffusion tensor affect the fitted MF parameters and often translate into artificial R_{ex} values in the range of 1–3 s^{-1}. To assess the impact of these effects on the fitted order parameters, we also performed MF fits using an effective correlation time (local τ_m) for each residue, without reference to the overall structure of the protein. The resulting two sets of MF parameters determined by these alternative approaches generally agree well with a mean deviation in S^2 of less than 0.03 ± 0.02 for all states and constructs.

The MF optimization generally resulted in back-calculated relaxation rates that are in good agreement with the experimental data (full set of fitted MF parameters and back-calculated relaxation rates are available via Mendeley Data). Overall, the resulting estimates of the global diffusion tensors (SI Table S2) appear to be fully consistent with the results from rotdif (Table 2) and SEC (Table 3), indicating a successful separation of global and local motions in the MF analysis. The final order parameters resulting from the MF analysis are shown in Figures 5 and 6.

We investigated how the internal dynamics differ between BD1 and BD2 in their isolated and tandem forms, and how each domain responds to ligand binding. We initially focus on the results for the isolated domains, because the underlying data are generally of higher quality compared to those for the tandem construct. The average order parameter of residues in α-helices differs between the two domains in the apo state (Figure 6a, b), with values of 0.90 for BD1 and 0.82 for BD2 (the standard error of the mean is less than 0.01 in each case), which can be compared with the average value of 0.88 ± 0.07 (one standard deviation) for residues in α-helices, obtained from a larger database of S^2 values in globular proteins.[73] The observed difference in S^2 indicates that BD1 is more rigid than BD2, a result that is in general agreement with recent molecular dynamics (MD) simulations[8,74] and amide-exchange mass-spectrometry.[75] Furthermore,

the overall stability towards unfolding in urea is also markedly different, with BD1 being more stable than BD2 toward loss of tertiary structure.[76]

The two domains show different profiles of S^2 values along the protein sequence, where in particular the ZA and AB loops have higher mobility in BD2, whereas the BC loop has similar mobility in the two domains (Fig. 5a, b). Two recent MD simulations both indicate a higher ZA mobility in BD2, as well as similar fluctuations of the BC loop in the two domains,[8,74] but only the study by Cheng et al. [8] shows an effect on the AB loop similar to our results.

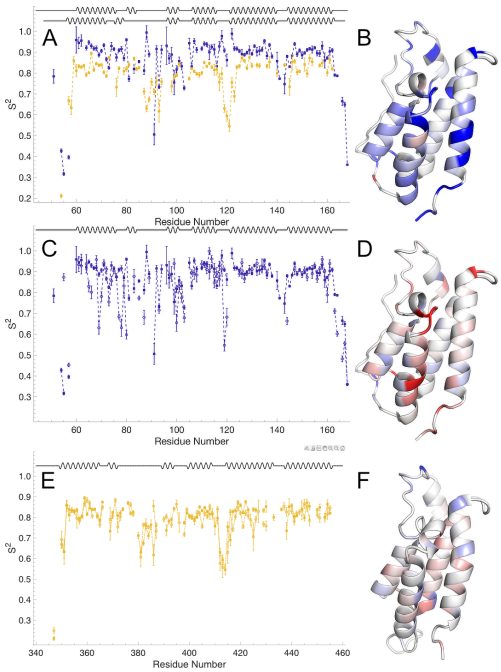


Figure 5. NMR order parameters (S^2) of isolated BRD4 bromodomains in the apo and H4 tetra-acetylated peptide (H4Kac4) bound states. (A) S^2 versus residue number for the apo states of BD1 (blue) and BD2 (yellow). The sequences are aligned via the BC-loop segment (residues 140–144 in BD1, 433–437 in BD2). (B) Difference in S^2 between apo BD1 and BD2 (data from panel A), color coded onto the BD1 structure, PDB: 4CLB.[14] (C) S^2 versus residue number for BD1, apo (filled squares), H4Kac4-bound (open circles). (D) Difference in S^2 between H4Kac4-bound and apo BD1, color coded onto the BD1 structure. (E) S^2 versus residue number for BD2, apo (filled squares), H4Kac4-bound (open circles). (F) Difference in S^2 between H4Kac4-bound and apo BD2, color coded onto the BD2 structure, PDB: 2LSP.[15] The black line at the top of panels A, C, and E indicates the location of loops (lines) and α-helices (wave); in panel A, the top and bottom lines refer to BD1 and BD2, respectively. The color coding in panels B, D, and F depicts differences in S^2, $\Delta S^2 = S^2$(HKac-4 bound) – S^2(apo), in the range [–0.3; 0.3] from red (negative), via white to blue (positive). Panels B, D and F were prepared using PyMOL.[16]

Figures 5c–f show comparisons of the order parameters for the apo and peptide-bound states of the two isolated domains. Upon peptide binding BD1 gains flexibility relative to the apo form (Fig. 5c, d), specifically in the ZA and AB loops, and apparently also in the BC loop, although there are few data points in this loop for the peptide-bound state. In BD2, peptide binding leads to lower order parameters in the ZA and BC loops, but not in the AB loop (Fig. 5e, f). As mentioned above, MD simulations have suggested that binding of various synthetic ligands can lead to increased conformational fluctuations of the bromodomains, where the relative changes in BD1 and BD2 depend sensitively on the ligand structure. Our present results now detail the response of the BDs to binding a natural H4Kac4 peptide. Increased flexibility of the loop segments in the peptide-bound state suggests that structure becomes slightly more expanded with increased hydrodynamic radius, which is in agreement with the results on rotational diffusion described above. The difference in flexibility of the AB loop is highly unexpected, since it is located at the opposite end of the four-helix bundle from the binding site, but its internal dynamics might reflect dimerization. The S^2 values are low for the AB loop in apo BD2, which is dimeric. Peptide binding to BD2 does not change the S^2 values of the AB loop and it does not appear to change the population of dimers. Apo BD1 has high S^2 values in the AB loop and it is monomeric. Peptide binding to BD1 leads to partial dimerization and reduction in the S^2 values of the AB loop. Thus, increased flexibility of the AB loop is likely the result of dimer formation.

Taken together, the order parameters show that the dynamic response to binding a natural acetylated peptide varies significantly between BD1 and BD2, demonstrating that the detailed amino acid sequence has dramatic consequences on the internal dynamics as well as the propensity to form dimers, despite the high degree of structural homology between the domains (cf. Fig. 1).

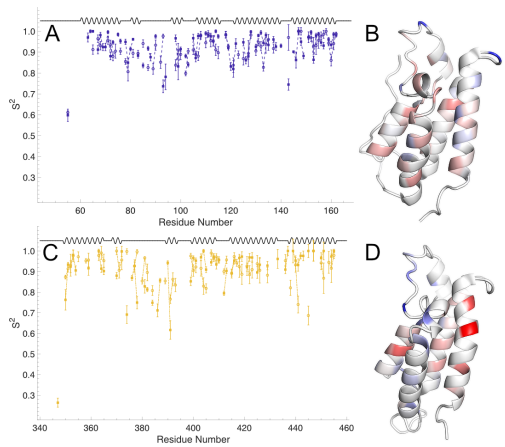


Figure 6. NMR order parameters (S^2) of tandem BRD4 bromodomains in the apo and H4Kac4-bound states. (A) S^2 versus residue number for BD1, apo (filled squares), H4Kac4 bound (open circles). (B) Difference in S^2 between HKac-4 bound and apo BD1, color coded onto the BD1 structure, PDB: 4CLB.[14] (C)

S^2 versus residue number for BD2, apo (filled squares) H4Kac4 bound (open circles). (D) Difference in S^2 between H4Kac4 bound and apo BD2, color coded onto the BD2 structure, PDB: 2LSP.[15].The black line at the top of panels A and C indicates the location of loops (lines) and α-helices (wave). The color coding in panels C and D depicts differences in S^2 , $\Delta S^2 = S^2$(HKac-4 bound) – S^2(apo), in the range [–0.3; 0.3] from red (negative), via white to blue (positive). Panels B, D and F were prepared using PyMOL.[16]

Compared to their isolated forms, both domains of BRD4(1,2) appear to be more rigid regardless of whether they are in the apo or peptide-bound states (Fig. 6a–d). In the tandem construct the BD1 domain shows relatively small changes in order parameter between the apo and peptide-bound states, indicating limited structural-dynamical changes upon binding. This result is in agreement with the small changes in the diffusion tensor reported above, and indicates that bivalent peptide binding occurs to a much lower extent in the tandem construct than in the isolated domain. The differences in order parameters between the apo and peptide-bound states of BD2 in BRD4(1,2) seem to indicate that peptide binding leads to slightly decreased flexibility of the ZA loop, but increased flexibility of the BC loop, whereas the AB loop is less affected.

Interpreting slower timescale exchange dynamics. The MF analysis results in conformational exchange contributions, R_{ex}, to the transverse relaxation rates for a relatively large number of residues, especially in BD2. In order to validate these results, we performed spectral density mapping,[58] based on the R_1, R_2, and *NOE* relaxation parameters. In the absence of exchange contributions to R_2, the spectral density component $J(0)$ should not depend on the static magnetic field strength (B_0). In the presence of exchange on the intermediate to fast time scale, spectral density mapping instead results in increased $J(0)$ values with increasing B_0. Thus, by plotting the $J(0)$ values extracted from the relaxation data sets obtained at 14.1 T and 18.8 T against one another, we identified those residues that deviate from the straight line with unit slope and zero intercept as likely to experience exchange (Fig. 7). This analysis clearly indicates that the isolated BD2 shows exchange in both the apo and H4Kac4-bound states, in agreement with the results presented above on τ_c that indicate exchange between monomeric and dimeric states. In contrast, exchange is less prominent in isolated BD1 and essentially absent in BRD4(1,2) for both domains. To further validate the MF-derived R_{ex} terms for isolated BD2, we compared these with exchange contributions estimated by comparing Γ_{auto} and Γ_{cross}, determined from linear combinations of relaxation rate constants involving either R_2 or η_{xy}, respectively, using the approach presented by Palmer and coworkers;[53,59] see Materials and Methods. This analysis confirms the larger R_{ex} contributions estimated by the MF approach for isolated BD2 in the apo and peptide-bound states (data not shown). There is no correlation between $\Delta\delta$ and R_{ex} determined for the peptide-bound state of BD2, indicating that the exchange is not due to exchange kinetics between free and bound states, but rather reflects intrinsic conformational dynamics on the micro- to millisecond timescale.

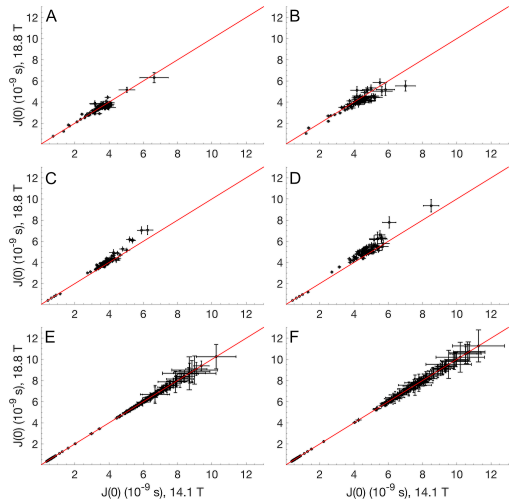


Figure 7. Spectral density values $J(0)$ determined by spectral density mapping of ^{15}N relaxation data measured at static magnetic field strengths of 14.1 T and 18.8 T. (A) apo BD1, (B) H4Kac4-bound BD1, (C) apo BD2, (D) H4Kac4-bound BD2, (E) apo BRD4(1,2), (F) H4Kac4-bound BRD4(1,2). Black symbols show pairs of $J(0)$ determined at B_0 = 14.1 T and 18.8 T with error bars indicating one standard deviation. The red line, with a slope of 1 and intercept of 0, is drawn to guide the eye.

Concluding remarks. We have investigated the dynamics of the individual bromodomains of BRD4 and their interactions with a tetra-acetylated peptide from histone 4, both in the context of tandem BRD4(1,2) and as isolated domains. We have identified notable differences between BD1 and BD2 in their propensities to form dimers, in their internal dynamics, and in the response of these characteristics to peptide binding. The results reported herein establish a basis for understanding the role of intrinsic bromodomain dynamics in governing interactions with acetylated histones and transcription factors, which in many cases seem to involve cooperative bromodomain binding.[4] Furthermore, the present paper describes differential dynamics of the two bromodomains that should provide valuable insights relevant to drug design initiatives to achieve inhibitor selectivity.

ASSOCIATED CONTENT

Supporting information: Figures showing all relaxation data, histograms of N-H bond vector distributions, trimmed means of relaxation datasets (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

* Kevin J. Embrey, kevin.embrey@astrazeneca.com; Mikael Akke, mikael.akke@bpc.lu.se

Present Addresses

[4] Roche Pharma Research and Early Development, pRED Informatics, Roche Innovation Center Basel, F. Hoffmann-La Roche Ltd, Grenzacherstrasse 124, 4070 Basel, Switzerland
[5] Peak Proteins Ltd., Birchwood House, Larkwood Way, Tytherington Business Park, Macclesfield, SK10 2XR, UK
[6] High-throughput Screening, Discovery Sciences, BioPharmaceuticals R&D, AstraZeneca, Alderley Park, UK
[7] New Modalities Product Development, Pharmaceutical Technology & Development, Operations, AstraZeneca, Macclesfield, UK

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript. ‡These authors contributed equally.

ACKNOWLEDGMENT

This work was supported by the Swedish Research Council (2018-4995 awarded to MA). We thank Göran Carlström and David Fushman for helpful discussions.

ABBREVIATIONS

BET, bromodomain and extra-terminal protein; BRD4, Bromodomain-containing Protein 4; BD1, Bromodomain 1; BD2, Bromodomain 2; CSA, chemical shift anisotropy; DSS, 4,4-dimethyl-4-silapentane-1-sulfonic acid; H4Kac4, tetra-acetylated peptide comprising the N-terminal sequence of histone 4 (residues 1–16) with lysine acetylation on K5, K8, K12 and K16; HEPES, 4-(2-hydroxyethyl)piperazine-1-ethanesulfonic acid; HMQC, heteronuclear multiple-quantum coherence; Kac, acetylated lysine residue; MF, model-free; MD, molecular dynamics; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser enhancement; SEC, size-exclusion chromatography; TCEP, tris(2-carboxyethyl)phosphine; TEV, tobacco etch virus; TROSY, transverse relaxation optimized spectroscopy.

REFERENCES

(1) Taverna, S. D., Li, H., Ruthenburg, A. J., Allis, C. D., and Patel, D. J. (2007) How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers. *Nat. Struct. Mol. Biol. 14*, 1025–1040.

(2) Dhalluin, C., Carlson, J. E., Zeng, L., He, C., Aggarwal, A. K., and Zhou, M. M. (1999) Structure and ligand of a histone acetyltransferase bromodomain. *Nature 399*, 491–496.

(3) Shi, J., and Vakoc, C. R. (2014) The Mechanisms behind the Therapeutic Activity of BET Bromodomain Inhibition. *Mol. Cell 54*, 728–736.

(4) Zaware, N., and Zhou, M. M. (2019) Bromodomain biology and drug discovery. *Nat. Struct. Mol. Biol. 26*, 870–879.

(5) Nicodeme, E., Jeffrey, K. L., Schaefer, U., Beinke, S., Dewell, S., Chung, C. W., Chandwani, R., Marazzi, I., Wilson, P., Coste, H., White, J., Kirilovsky, J., Rice, C. M., Lora, J. M., Prinjha, R. K., Lee, K., and Tarakhovsky, A. (2010) Suppression of inflammation by a synthetic histone mimic. *Nature 468*, 1119–1123.

(6) Filippakopoulos, P., Qi, J., Picaud, S., Shen, Y., Smith, W. B., Fedorov, O., Morse, E. M., Keates, T., Hickman, T. T., Felletar, I., Philpott, M., Munro, S., McKeown, M. R., Wang, Y., Christie, A. L., West, N., Cameron, M. J., Schwartz, B., Heightman, T. D., La Thangue, N., French, C. A., Wiest, O., Kung, A. L., Knapp, S., and Bradner, J. E. (2010) Selective inhibition of BET bromodomains. *Nature 468*, 1067–1073.

(7) Gilan, O., Rioja, I., Knezevic, K., Bell, M. J., Yeung, M. M., Harker, N. R., Lam, E. Y. N., Chung, C., Bamborough, P., Petretich, M., Urh, M., Atkinson, S. J., Bassil, A. K., Roberts, E. J., Vassiliadis, D., Burr, M. L., Preston, A. G. S., Wellaway, C., Werner, T., Gray, J. R., Michon, A. M., Gobbetti, T., Kumar, V., Soden, P. E., Haynes, A., Vappiani, J., Tough, D. F., Taylor, S., Dawson, S. J., Bantscheff, M., Lindon, M., Drewes, G., Demont, E. H., Daniels, D. L., Grandi, P., Prinjha, R. K., and Dawson, M. A. (2020) Selective targeting of BD1 and BD2 of the BET proteins in cancer and immunoinflammation. *Science (80-.). 368*, 387–394.

(8) Cheng, C., Diao, H., Zhang, F., Wang, Y., Wang, K., and Wu, R. (2017) Deciphering the mechanisms of selective inhibition for the tandem BD1/BD2 in the BET-bromodomain family. *Phys. Chem. Chem. Phys. 19*, 23934–23941.

(9) Raich, L., Meier, K., Günther, J., Christ, C. D., Noé, F., and Olsson, S. (2021) Discovery of a hidden transient state in all bromodomain families. *Proc. Natl. Acad. Sci. U. S. A. 118*, e2017427118.

(10) Steiner, S., Magno, A., Huang, D., and Caflisch, A. (2013) Does bromodomain flexibility influence histone recognition? *FEBS Lett. 587*, 2158–2163.

(11) Zhang, X., Chen, K., Wu, Y. D., and Wiest, O. (2017) Protein dynamics and structural waters in bromodomains. *PLoS One 12*, e0186570.

(12) Raux, B., Voitovich, Y., Derviaux, C., Lugari, A., Rebuffet, E., Milhas, S., Priet, S., Roux, T., Trinquet, E., Guillemot, J. C., Knapp, S., Brunel, J. M., Fedorov, A. Y., Collette, Y., Roche, P., Betzi, S., Combes, S., and Morelli, X. (2016) Exploring Selective Inhibition of the First Bromodomain of the Human Bromodomain and Extra-terminal Domain (BET) Proteins. *J. Med. Chem. 59*, 1634–1641.

(13) Picaud, S., Wells, C., Felletar, I., Brotherton, D., Martin, S., Savitsky, P., Diez-Dacal, B., Philpott, M., Bountra, C., Lingard, H., Fedorov, O., Müller, S., Brennan, P. E., Knapp, S., and Filippakopoulos, P. (2013) RVX-208, an inhibitor of BET transcriptional regulators with selectivity for the second bromodomain. *Proc. Natl. Acad. Sci. U. S. A. 110*, 19754–19759.

(14) Atkinson, S. J., Soden, P. E., Angell, D. C., Bantscheff, M., Chung, C. W., Giblin, K. A., Smithers, N., Furze, R. C., Gordon, L., Drewes, G., Rioja, I., Witherington, J., Parr, N. J., and Prinjha, R. K. (2014) The structure based design of dual HDAC/BET inhibitors as novel epigenetic probes. *Medchemcomm 5*, 342–351.

(15) Zhang, G., Liu, R., Zhong, Y., Plotnikov, A. N., Zhang, W., Zeng, L., Rusinova, E., Gerona-Nevarro, G., Moshkina, N., Joshua, J., Chuang, P. Y., Ohlmeyer, M., He, J. C., and Zhou, M. M. (2012) Down-regulation of NF-κB transcriptional activity in HIV-associated kidney disease by BRD4 inhibition. *J. Biol. Chem. 287*, 28840–28851.

(16) The PyMOL Molecular Graphics System. Schrödinger, LLC.

(17) Baud, M. G. J., Lin-Shiao, E., Cardote, T., Tallant, C., Pschibul, A., Chan, K.-H., Zengerle, M., Garcia, J. R., Kwan, T. T.-L., Ferguson, F. M., and Ciulli, A. (2014) A bump-and-hole approach to engineer controlled selectivity of BET bromodomain chemical probes. *Science (80-.).*

(18) Schröder, S., Cho, S., Zeng, L., Zhang, Q., Kaehlcke, K., Mak, L., Lau, J., Bisgrove, D., Schnölzer, M., Verdin, E., Zhou, M. M., and Ott, M. (2012) Two-pronged binding with bromodomain-containing protein 4 liberates positive transcription elongation factor b from inactive ribonucleoprotein complexes. *J. Biol. Chem.*

(19) Ren, C., Zhang, G., Han, F., Fu, S., Cao, Y., Zhang, F., Zhang, Q., Meslamani, J., Xu, Y., Ji, D., Cao, L., Zhou, Q., Cheung, K. lung, Sharma, R., Babault, N., Yi, Z., Zhang, W., Walsh, M. J., Zeng, L., and Zhou, M. M. (2018) Spatially constrained tandem bromodomain inhibition bolsters sustained repression of BRD4 transcriptional activity for TNBC cell growth. *Proc. Natl. Acad. Sci. U. S. A. 115*, 7949–6954.

(20) Wang, R., Li, Q., Helfer, C. M., Jiao, J., and You, J. (2012) Bromodomain protein Brd4 associated with acetylated chromatin is important for maintenance of higher-order chromatin structure. *J. Biol. Chem. 287*, 10738–10752.

(21) Waring, M. J., Chen, H., Rabow, A. A., Walker, G., Bobby, R., Boiko, S., Bradbury, R. H., Callis, R., Clark, E., Dale, I., Daniels, D. L., Dulak, A., Flavell, L., Holdgate, G., Jowitt, T. A., Kikhney, A., McAlister, M., Méndez, J., Ogg, D., Patel, J., Petteruti, P., Robb, G. R., Robers, M. B., Saif, S., Stratton, N., Svergun, D. I., Wang, W.,

Whittaker, D., Wilson, D. M., and Yao, Y. (2016) Potent and selective bivalent inhibitors of BET bromodomains. *Nat. Chem. Biol.*

(22) Tanaka, M., Roberts, J. M., Seo, H. S., Souza, A., Paulk, J., Scott, T. G., Deangelo, S. L., Dhe-Paganon, S., and Bradner, J. E. (2016) Design and characterization of bivalent BET inhibitors. *Nat. Chem. Biol.*

(23) Liu, Y., Wang, X., Zhang, J., Huang, H., Ding, B., Wu, J., and Shi, Y. (2008) Structural basis and binding properties of the second bromodomain of Brd4 with acetylated histone tails. *Biochemistry*.

(24) Findeisen, M., Brand, T., and Berger, S. (2007) A 1H-NMR thermometer suitable for cryoprobes. *Magn. Reson. Chem. 45*, 175–178.

(25) Markley, J. L., Bax, A., Arata, Y., Hilbers, C. W., Kaptein, R., Sykes, B. D., Wright, P. E., and Wüthrich, K. (1998) Recommendations for the presentation of NMR structures of proteins and nucleic acids. *J. Mol. Biol. 280*, 933–952.

(26) Cavanagh, J., Fairbrother, W. J., Palmer, A. G., III, and Skelton, N. J. (1995) Protein NMR Spectroscopy: Principles and Practice 2nd ed. Academic Press.

(27) Schanda, P., and Brutscher, B. (2005) Very Fast Two-Dimensional NMR Spectroscopy for Real-Time Investigation of Dynamic Events in Proteins on the Time Scale of Seconds. *J. Am. Chem. Soc. 127*, 8014–8015.

(28) Kern, T., Schanda, P., and Brutscher, B. (2008) Sensitivity-enhanced IPAP-SOFAST-HMQC for fast-pulsing 2D NMR with reduced radiofrequency load. *J. Magn. Reson. 190*, 333–338.

(29) Grzesiek, S., and Bax, A. (1992) An efficient experiment for sequential backbone assignment of medium-sized isotopically enriched proteins. *J. Magn. Reson. 99*, 201–207.

(30) Grzesiek, S., and Bax, A. (1992) Correlating backbone amide and side chain resonances in larger proteins by multiple relayed triple resonance NMR. *J. Am. Chem. Soc. 114*, 6291–6293.

(31) Bax, A., and Ikura, M. (1991) An efficient 3D NMR technique for correlating the proton and15N backbone amide resonances with the α-carbon of the preceding residue in uniformly15N/13C enriched proteins. *J. Biomol. NMR 1*, 99–104.

(32) Kay, L. E., Ikura, M., Tschudin, R., and Bax, A. (1990) Three-dimensional triple-resonance NMR spectroscopy of isotopically enriched proteins. *J. Magn. Reson. 89*, 496–514.

(33) Salzmann, M., Wider, G., Pervushin, K., and Wüthrich, K. (1999) Improved sensitivity and coherence selection for [15N,1H]-TROSY elements in triple resonance experiments. *J. Biomol. NMR 15*, 181–184.

(34) Yoshimura, Y., Kulminskaya, N. V, and Mulder, F. A. A. (2015) Easy and unambiguous sequential assignments of intrinsically disordered proteins by correlating the backbone 15N or 13C' chemical shifts of multiple contiguous residues in highly resolved 3D spectra. *J. Biomol. NMR 61*, 109–21.

(35) Hyberts, S. G., Milbradt, A. G., Wagner, A. B., Arthanari, H., and Wagner, G. (2012) Application of iterative soft thresholding for fast reconstruction of NMR data non-uniformly sampled with multidimensional Poisson Gap scheduling. *J. Biomol. NMR 52*, 315–327.

(36) Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR 6*, 277–293.

(37) Vranken, W. F., Boucher, W., Stevens, T. J., Fogh, R. H., Pajon, A., Llinas, M., Ulrich, E. L., Markley, J. L., Ionides, J., and Laue, E. D. (2005) The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins 59*, 687–96.

(38) Helmus, J. J., and Jaroniec, C. P. (2013) Nmrglue: an open source Python package for the analysis of multidimensional NMR data. *J. Biomol. NMR 55*, 355–367.

(39) Zhu, G., Xia, Y., Nicholson, L. K., and Sze, K. H. (2000) Protein dynamics measurements by TROSY-based NMR experiments. *J. Magn. Reson. 143*, 423–426.

(40) Kay, L. E., Torchia, D. A., and 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.

(41) Lakomek, N. A., Kaufman, J. D., Stahl, S. J., Louis, J. M., Grishaev, A., Wingfield, P. T., and Bax, A. (2013) Internal dynamics of the homotrimeric HIV-1 viral coat protein gp41 on multiple time scales. *Angew. Chemie - Int. Ed. 52*, 3911–3915.

(42) Lee, D., Hilty, C., Wider, G., and Wüthrich, K. (2006) Effective rotational correlation times of proteins from NMR relaxation interference. *J Magn Reson 178*, 72–76.

(43) d'Auvergne, E. J., and Gooley, P. R. (2008) Optimisation of NMR dynamic models I. Minimisation algorithms and their performance within the model-free and Brownian rotational diffusion spaces. *J. Biomol. NMR 40*, 107–119.

(44) Press, W. H., Flannery, B. P., Teukolsky, S. A., and Vetterling, W. T. (1986) Numerical Recipes. The Art of Scientific Computing. Cambridge University Press, Cambridge.

(45) Halle, B., and Wennerström, H. (1981) Interpretation of magnetic resonance data from water nuclei in heterogeneous systems. *J. Chem. Phys. 75*, 1928–1943.

(46) Lipari, G., and Szabo, A. (1982) 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.

(47) Clore, G. M., Szabo, A., Bax, A., Kay, L. E., Driscoll, P. C., and Gronenborn, A. M. (1990) 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.

(48) Halle, B. (2009) The physical basis of model-free analysis of NMR relaxation data from proteins and complex fluids. *J. Chem. Phys. 131*, 224507.

(49) d'Auvergne, E. J., and Gooley, P. R. (2007) Set theory formulation of the model-free problem and the diffusion seeded model-free paradigm. *Mol. Biosyst. 3*, 483–494.

(50) Walker, O., Varadan, R., and Fushman, D. (2004) Efficient and accurate determination of the overall rotational diffusion tensor of a molecule from 15N relaxation data using computer program ROTDIF. *J. Magn. Reson. 168*, 336–345.

(51) Yao, L., Grishaev, A., Cornilescu, G., and Bax, A. (2010) Site-specific backbone amide 15N chemical shift anisotropy tensors in a small protein from liquid crystal and cross-correlated relaxation measurements. *J. Am. Chem. Soc. 132*, 4295–4309.

(52) Fushman, D., Tjandra, N., and Cowburn, D. (1998) Direct measurement of 15N chemical shift anisotropy in solution. *J. Am. Chem. Soc. 120*, 10947–10952.

(53) Kroenke, C. D., Rance, M., and Palmer, A. G. (1999) Variability of the 15N chemical shift anisotropy in Escherichia coli ribonuclease H in solution. *J. Am. Chem. Soc. 121*, 10119–10125.

(54) Ghose, R., Fushman, D., and Cowburn, D. (2001) Determination of the rotational diffusion tensor of macromolecules in solution from NMR relaxation data with a combination of exact and approximate methods - Application to the determination of interdomain orientation in multidomain proteins. *J. Magn. Reson. 149*, 204–217.

(55) Bernado, P., Garcia de la Torre, J., and Pons, M. (2002) Interpretation of 15N NMR relaxation data of globular proteins using hydrodynamic calculations with HYDRONMR. *J. Biomol. NMR 23*, 139–150.

(56) Garcia de la Torre, J., Huertas, M. L., and Carrasco, B. (2000) HYDRONMR: prediction of NMR relaxation of globular proteins from atomic-level structures and hydrodynamic calculations. *J. Magn. Reson. 147*, 719–730.

(57) Halle, B., and Davidovic, M. (2003) Biomolecular hydration: from water dynamics to hydrodynamics. *Proc. Natl. Acad. Sci. U. S. A. 100*, 12135–12140.

(58) Farrow, N. A., Zhang, O., Szabo, A., Torchia, D. A., and Kay, L. E. (1995) Spectral density function mapping using 15N relaxation data exclusively. *J. Biomol. NMR 6*, 153–162.

(59) Palmer, A. G., Kroenke, C. D., and Loria, J. P. (2001) Nuclear magnetic resonance methods for quantifying microsecond-to-millisecond motions in biological macromolecules. *Methods Enzymol. 339*, 204–238.

(60) Williams, F. P., Milbradt, A. G., Embrey, K. J., and Bobby, R. (2016) Segmental isotope labelling of an individual bromodomain of a tandem domain BRD4 using Sortase A. *PLoS One 11*, e0154607.

(61) Patel, K., Solomon, P. D., Walshe, J. L., Ford, D. J., Payne, R. J., Wilkinson-White, L., Low, J. K. K., and Mackay, J. P. (2021) BET-Family Bromodomains Can Recognize Diacetylated Sequences from Transcription Factors Using a Conserved Mechanism. *Biochemistry 60*, 648–662.

(62) Wu, S. Y., and Chiang, C. M. (2007) The double bromodomain-containing chromatin adaptor Brd4 and transcriptional regulation. *J. Biol. Chem. 282*, 13141–12145.

(63) Nakamura, Y., Umehara, T., Nakano, K., Moon, K. J., Shirouzu, M., Morita, S., Uda-Tochio, H., Hamana, H., Terada, T., Adachi, N., Matsumoto, T., Tanaka, A., Horikoshi, M., Ozato, K., Padmanabhan, B., and Yokoyama, S. (2007) Crystal structure of the human BRD2 bromodomain: Insights into dimerization and recognition of acetylated histone H4. *J. Biol. Chem. 282*, 4193–4201.

(64) Filippakopoulos, P., Picaud, S., Mangos, M., Keates, T., Lambert, J. P., Barsyte-Lovejoy, D., Felletar, I., Volkmer, R., Müller, S., Pawson, T., Gingras, A. C., Arrowsmith, C. H., and Knapp, S. (2012) Histone recognition and large-scale structural analysis of the human bromodomain family. *Cell 149*, 214–231.

(65) Vollmuth, F., Blankenfeldt, W., and Geyer, M. (2009) Structures of the dual bromodomains of the P-TEFb-activating protein Brd4 at atomic resolution. *J. Biol. Chem. 284*, 36547–36556.

(66) Morinière, J., Rousseaux, S., Steuerwald, U., Soler-López, M., Curtet, S., Vitte, A. L., Govin, J., Gaucher, J., Sadoul, K., Hart, D. J., Krijgsveld, J., Khochbin, S., Müller, C. W., and Petosa, C. (2009) Cooperative binding of two acetylation marks on a histone tail by a single bromodomain. *Nature 461*, 664–668.

(67) Umehara, T., Nakamura, Y., Jang, M. K., Nakano, K., Tanaka, A., Ozato, K., Padmanabhan, B., and Yokoyama, S. (2010) Structural basis for acetylated histone H4 recognition by the human BRD2 bromodomain. *J. Biol. Chem. 285*, 7610–7618.

(68) Patel, K., Walport, L. J., Walshe, J. L., Solomon, P. D., Low, J. K. K., Tran, D. H., Mouradian, K. S., Silva, A. P. G., Wilkinson-White, L., Norman, A., Franck, C., Matthews, J. M., Mitchell Guss, J., Payne, R. J., Passioura, T., Suga, H., and Mackay, J. P. (2020) Cyclic peptides can engage a single binding pocket through highly divergent modes. *Proc. Nat. Acad. Sci. U. S. A. 117*, 26728–26738.

(69) Mulder, F. A. A., Bouakaz, L., Lundell, A., Venkataramana, M., Liljas, A., Akke, M., and Sanyal, S. (2004) Conformation and dynamics of ribosomal stalk protein L12 in solution and on the ribosome. *Biochemistry 43*, 5930–5936.

(70) Walsh, J. D., Meier, K., Ishima, R., and Gronenborn, A. M. (2010) NMR Studies on Domain Diffusion and Alignment in Modular GB1 Repeats. *Biophys. J. 99*, 2636–2646.

(71) Lipari, G., and Szabo, A. (1982) Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. I. Theory and range of validity. *J. Am. Chem. Soc. 104*, 4546–4559.

(72) Gill, M. L., and Palmer, A. G. (2014) Local isotropic diffusion approximation for coupled internal and overall molecular motions in NMR spin relaxation. *J. Phys. Chem. B 118*, 11120–11128.

(73) Goodman, J. L., Pagel, M. D., and Stone, M. J. (2000) Relationships between protein structure and dynamics from a database of NMR-derived backbone order parameters. *J. Mol. Biol. 295*, 963–978.

(74) Rodríguez, Y., Gerona-Navarro, G., Osman, R., and Zhou, M. M. (2020) In silico design and molecular basis for the selectivity of Olinone toward the first over the second bromodomain of BRD4. *Proteins Struct. Funct. Bioinforma. 88*, 414–430.

(75) Malvezzi, F., Stubbs, C. J., Jowitt, T. A., Dale, I. L., Guo, X., DeGnore, J. P., Degliesposti, G., Skehel, J. M., Bannister, A. J., and McAlister, M. (2021) Phosphorylation-dependent BRD4 dimerization and implications for therapeutic inhibition of BET family proteins.

(76) Lori, L., Pasquo, A., Lori, C., Petrosino, M., Chiaraluce, R., Tallant, C., Knapp, S., and Consalvi, V. (2016) Effect of BET missense mutations on bromodomain function, inhibitor binding and stability. *PLoS One 11*, e0159180-20.

Supplementary Information

Bromodomain Interactions with Acetylated Histone 4 Peptides in the BRD4 Tandem Domain: Effects on Domain Dynamics and Internal Flexibility

Sven Wernersson, Romel Bobby, Liz Flavell, Alexander G. Milbradt, Geoff A. Holdgate, Kevin J. Embrey, and Mikael Akke**

Corresponding Authors
* Kevin J. Embrey, kevin.embrey@astrazeneca.com; Mikael Akke, mikael.akke@bpc.lu.se

Supplementary Table S1. Trimmed averages of all NMR relaxation data.

14.1 T data

apo constructs	R_1 (s^{-1})[a]	R_2 (s^{-1})[b]	NOE[c]
BD1 (isolated)	1.33	13	0.79
BD2 (isolated)	1.09	15	0.80
BD1 (tandem)	0.97	22	0.81
BD2 (tandem)	0.94	26	0.79
H4Kac-4 bound constructs			
BD1 (isolated)	1.13	16	0.81
BD2 (isolated)	0.94	17	0.79
BD1 (tandem)	0.77	25	0.81
BD2 (tandem)	0.79	30	0.81

18.8 T data

apo constructs	R_1 (s^{-1})[a]	R_2 (s^{-1})[b]	NOE[c]	η_{xy} (s^{-1})[d]
BD1 (isolated)	0.97	15	0.83	11
BD2 (isolated)	0.80	18	0.84	14
BD1 (tandem)	0.60	26	0.87	22
BD2 (tandem)	0.67	32	0.86	25
H4Kac-4 bound constructs				
BD1 (isolated)	0.85	18	0.83	N.D.[e]
BD2 (isolated)	0.68	22	0.85	17
BD1 (tandem)	0.55	30	0.86	26
BD2 (tandem)	0.55	37	0.85	30

[a]The standard error of the mean (SEM) is ≤ 0.005 in all cases. [b]SEM ≤ 0.1 in all cases. [c]SEM $\leq$ 0.01 in all cases. [d]SEM ≤ 0.3 in all cases. [e]N.D. = No data.

Supplementary Table S2: Diffusion tensor parameters determined with the relax software suite.[a]

apo constructs			
Construct	D_{iso} (10^7 s^{-1})	$D_{\parallel}/D_{\perp}$	τ_c (ns)
BD1 (isolated)	1.90	1.41	8.8
BD2 (isolated)	1.62	1.30	10.3
BD1 (tandem)	1.23	1.80	13.5
BD2 (tandem)	1.06	1.12	15.7
H4Kac-4 bound constructs			
Construct	D_{iso} (10^7 s^{-1})	$D_{\parallel}/D_{\perp}$	τ_c (ns)
BD1 (isolated)	1.36	1.79	12.3
BD2 (isolated)	1.44	0.56	11.6
BD1 (tandem)	1.04	1.67	16.1
BD2 (tandem)	1.13	1.77	14.7

[a] D_{iso} is the trace of the diffusion tensor, ($D_{\parallel}/D_{\perp}$) is the anisotropy of the diffusion tensor, τ_c is the rotational diffusion correlation time. See the main text for definitions.

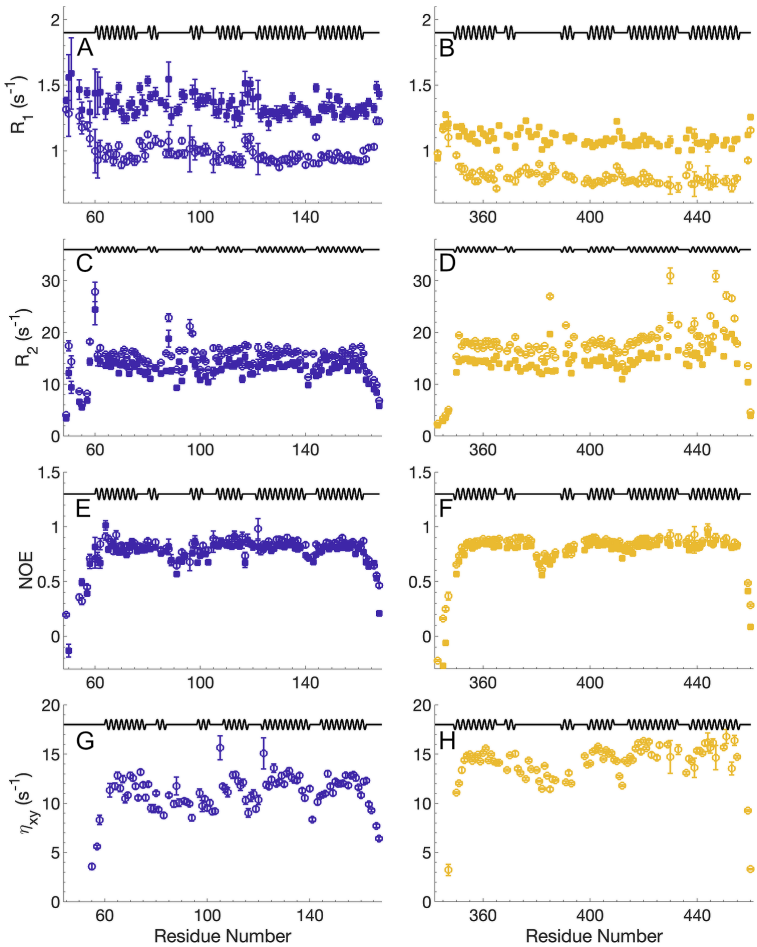


Figure S1. Relaxation data of apo, isolated BRD4 bromodomains at 14.1 T (filled squares) and 18.8 T (open circles). (A, B) R_1, (C, D) R_2, (E, F) $\{^1H\}$-^{15}N *NOE*, (G, H) η_{xy}. Left-hand column (A, C, E, G) apo isolated BD1. Right-hand column (B, D, F, H) apo isolated BD2.

The black line at the top of each panel shows the secondary structure, with α-helices represented as waves and intervening segments (loops and linker) as straight lines.

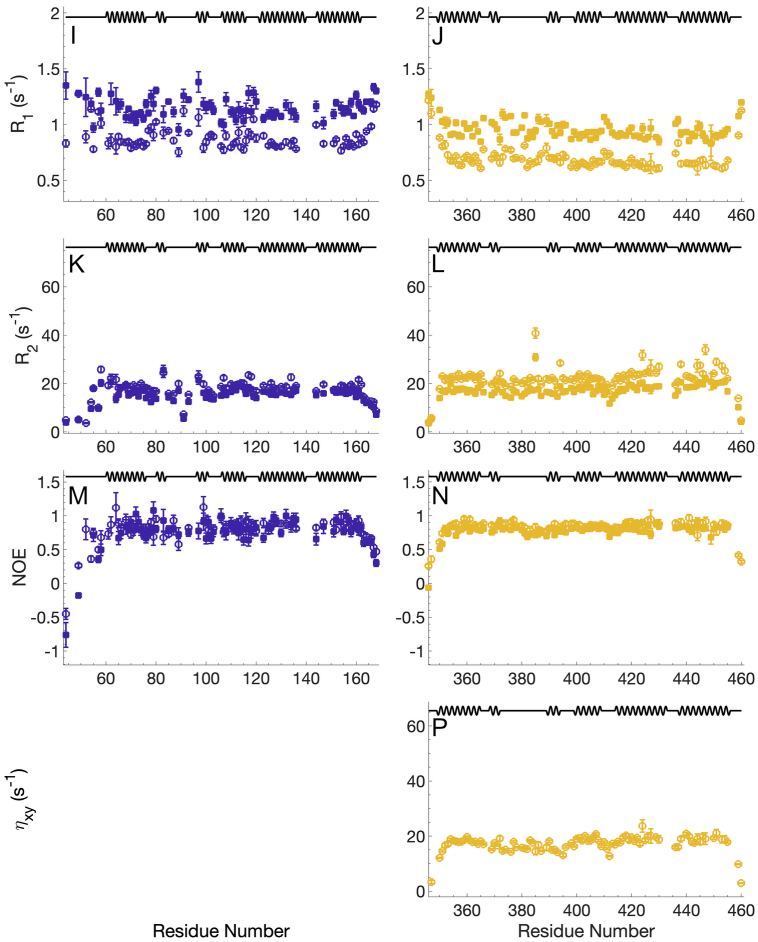


Figure S1, continued. Relaxation data of H4Kac-4 bound, isolated BRD4 bromodomains at 14.1 T (filled squares) and 18.8 T (open circles). (I, J) R_1, (K, L) R_2, (M, N) $\{^1H\}$-^{15}N *NOE*, (P) η_{xy}. Left-hand column (I, K, M) H4Kac-4 bound isolated BD1. Right-hand column (J, L, N, P) H4Kac-4 bound isolated BD2. The black line at the top of each panel shows the secondary structure, with α-helices represented as waves and intervening segments (loops and linker) as straight lines. Panel O is intentionally omitted because the corresponding data are not available.

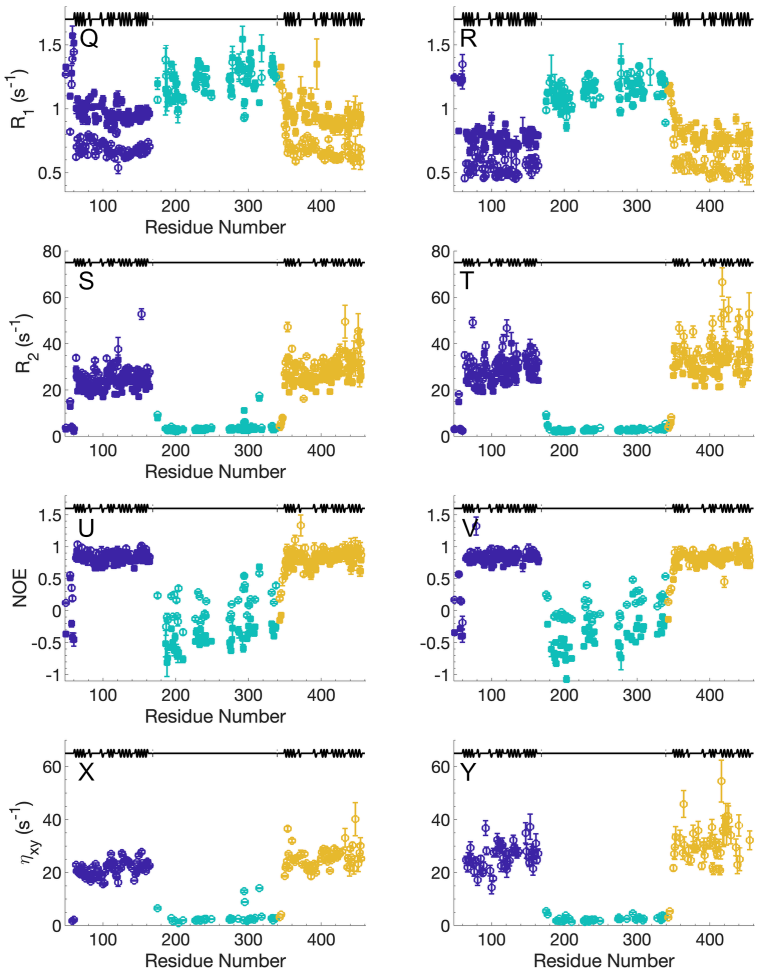


Figure S1, continued. Relaxation data of tandem BRD4(1,2) at 14.1 T (filled squares) and 18.8 T (open circles). (Q, R) R_1, (S, T) R_2, (U, V) $\{^1H\}$-^{15}N *NOE*, (X, Y) η_{xy}. Left-hand column (Q, S, U, X) Apo tandem BRD4(1,2). Right-hand column (R, T, V, Y) H4Kac-4 bound tandem BRD4(1,2). The black line at the top of each panel shows the secondary structure, with α-helices represented as waves and intervening segments (loops and linker) as straight lines.

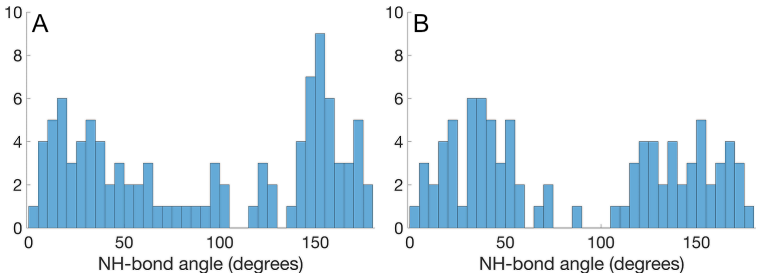


Figure S2. Histograms of N-H bond vector angles relative to the unique axis of the diffusion tensors calculated using rotdif. Each histogram contains 36 bins with a width of 5°. (A) Apo, isolated BD1, (B) Apo, isolated BD2.

Paper II

Conformational Dynamics of Designed Drug-like Ligands Bound to Proteins Determined by ^{19}F NMR Relaxation and Molecular Dynamics Simulations

Sven Wernersson[1], Olof Stenström[1], Simon Birgersson[1], Magdalena Riad [1,†], Ulrich Weininger[1,‡], Göran Carlström[2], Ulf J. Nilsson[2], Mikael Akke[1,*]

[1] Division of Biophysical Chemistry, Center for Molecular Protein Science, Department of Chemistry, Lund University, Sweden; [2] Centre for Analysis and Synthesis, Department of Chemistry, Lund University, Sweden.

Supporting Information Placeholder

ABSTRACT: Conformational fluctuations can contribute significantly to the entropy of various molecular processes, including the binding of drug-like ligands to proteins. It is of great interest to understand how residual ligand entropy in the protein-bound state contributes to affinity and specificity of binding. Pharmaceutical compounds often include fluorine atoms, which potentially can be used to probe molecular dynamics via the nuclear spin relaxation of the 100% naturally abundant ^{19}F isotope. We used ^{19}F NMR relaxation experiments and molecular dynamics simulations to determine the extent of conformational fluctuations of ligands bound to proteins. We measured ^{19}F relaxation rates at three different static magnetic field strengths and analyzed the results using the model-free formalism to determine order parameters for phenyl rings of protein-bound ligands. Our detailed analysis indicates that although ^{19}F relaxation depends on interactions with multiple surrounding protons the data can be represented reasonably well by a single order parameter for the aromatic ring plane.

Molecular recognition between proteins and low-molecular-weight ligands are central to biological processes. Understanding these interactions is a key aspect of drug design aiming to interfere with such processes for medicinal purposes. Rational structure-based drug design has advanced very significantly in recent years, due in large part to advances in computational chemistry, but still remains extremely challenging. The main challenge is that the free energy of complex formation involves a small difference between large terms originating from numerous different interactions between the protein, ligand and solvent molecules. Interaction energies depend sensitively on interatomic distances and orientations, which vary with the conformational dynamics. Furthermore, changes upon complex formation in the conformational fluctuations of the component parts of the system can result in significant entropic contributions to the binding free energy.[1–5]

Pharmaceutical compounds are often fluorinated because this leads to prolonged drug life-times in the patient and also enable highly specific interactions with molecular moieties in the target proteins.[6] Fluorine is highly suitable for NMR spectroscopic investigations due to its naturally abundant (100%) isotope ^{19}F, which has a nuclear spin quantum number of ½. Consequently, ^{19}F NMR is routinely used to monitor ligand binding to protein targets[7] and is also used to study proteins that have been selectively labeled with ^{19}F.[8]

Despite the common use of ^{19}F in studies aimed at screening ligand binding or other protein interactions, very few ^{19}F relaxation studies to probe dynamics have been reported due to inherent difficulties.[9–13] Interpreting ^{19}F relaxation in terms of conformational fluctuations of ligands bound to proteins is challenging because (i) the ^{19}F chemical shift anisotropy (CSA) might deviate from model values as a consequence of specific interactions with the protein atoms, and (ii) the ^{19}F spin typically experiences dipolar interactions with several ^{1}H spins in the ligand and protein. Here, we report order parameters for the F-C bond axis of fluorine atoms covalently attached to aromatic rings of designed ligands bound to galectin-3C. These order parameters can be interpreted in terms of fluctuations of the aromatic ring plane around the imaginary C^{δ}-C^{ζ} axis.

Here we address the conformational dynamics of two designed ligands bound to the carbohydrate recognition domain of galectin-3 (galectin-3C). The same protein-ligand complexes have been studied previously using a combination of protein NMR relaxation, isothermal titration calorimetry, MD simulations and X-ray crystallography,[14] which provides important background information to the present study.

We have previously studied the dynamics of galectin-3C in the same complexes as studied here using a combination of ^{15}N and ^{2}H spin relaxation, ensemble-refined X-ray diffraction data, and MD simulations.[14] The results from these different methods are in agreement and show that galectin-3C in complex with the S diastereomer has greater flexibility than it has in complex with the R diastereomer. The ensemble-refined X-ray data further indicate that the S ligand undergoes greater fluctuations than does R in the protein-bound state. However, the MD simulations did not reveal any significant difference in the dynamics of the two ligands. Here, we complement our previous NMR relaxation study of the protein dynamics and conformational entropy by directly probing the ligand dynamics using ^{19}F spin relaxation.

MATERIALS AND METHODS

^{19}F relaxation theory. The theoretical treatment initially assumes an isolated pair of ^{19}F and ^{1}H spins. In reality the ^{19}F spin is surrounded by two neighboring ^{1}H spins on the phenyl ring of the ligand and also by additional, more remote protons in both the protein and ligand. In the two-spin approximation, the single ^{1}H spin represents the total dipolar relaxation caused by all protons surrounding the ^{19}F spin (see below). The longitudinal relaxation of ^{19}F is given by:[15]

$$\begin{aligned}\frac{dF_z(t)}{dt} = &-R_1[F_z(t) - F_z^0] \\ &- \sigma_{FH}[H_z(t) - H_z^0] \\ &- \eta_z 2F_z H_z(t)\end{aligned} \tag{1}$$

and the transverse relaxation is given by:

$$\frac{dF_{xy}(t)}{dt} = -R_2 F_{xy}(t) - \eta_{xy} 2F_{xy} H_z(t) \tag{2}$$

where F_z and F_{xy} denote ^{19}F longitudinal and transverse magnetizations, respectively, R_1 and R_2 are the longitudinal and transverse auto-relaxation rate constants of ^{19}F, respectively, σ_{FH} is the rate constant for dipolar cross-relaxation between ^{19}F and ^{1}H, η_z and η_{xy} are the rate constants for longitudinal and transverse cross-correlated cross-relaxation involving the ^{19}F CSA and ^{19}F–^{1}H dipole-dipole interactions. If ^{1}H decoupling is applied the first expression simplifies to:

$$\frac{dF_z(t)}{dt} = -R_1[F_z(t) - F_z^0] + \sigma_{FH} H_z^0 \tag{3}$$

which has the solution

$$\frac{F_z(t)}{F_z^0} = 1 + \frac{\sigma_{FH}}{R_1}\frac{\gamma_H}{\gamma_F} - \left(2 + \frac{\sigma_{FH}}{R_1}\frac{\gamma_H}{\gamma_F}\right) e^{-R_1 t} \tag{4}$$

where γ_F and γ_H are the gyromagnetic ratios of ^{19}F and ^{1}H, respectively. The constant 2 can be replaced by a fitting parameter ϕ, which takes into account imperfect inversion of the longitudinal magnetization.

The transverse relaxation of ^{19}F with ^{1}H decoupling is described by a mono-exponential decay with relaxation rate constant R_2:

$$F_{xy}(t) = F_{xy}(0) e^{-R_2 t} \tag{5}$$

The relaxation rate constants depend on linear combinations of the spectral density, $J(\omega)$, sampled at specific frequencies.[15] In principle, each ^{19}F and ^{1}H$_i$ dipolar interaction should be described by its own spectral density function, $J_i(\omega)$, and the effective relaxation rates include the contributions from all of these interactions.

$$\begin{aligned}R_1 = \sum_i \frac{d_i^2}{4}[&J_i(\omega_H - \omega_F) + 3J_i(\omega_F) \\ &+ 6J_i(\omega_H + \omega_F)] \\ &+ c^2 J(\omega_F)\end{aligned} \tag{6}$$

$$\begin{aligned}R_2 = \sum_i \frac{d_i^2}{8}[&J_i(0) + J_i(\omega_H - \omega_F) \\ &+ 3J_i(\omega_F) + 6J_i(\omega_H) \\ &+ 6J_i(\omega_H + \omega_F)] \\ &+ \frac{c^2}{6}[4J(0) + 3J(\omega_F)] \\ &+ R_{ex}\end{aligned} \tag{7}$$

$$\begin{aligned}\sigma_{FH} = \sum_i \frac{d_i^2}{4}[&6J_i(\omega_H + \omega_F) \\ &- J_i(\omega_H - \omega_F)]\end{aligned} \tag{8}$$

where the sum runs over all ^{19}F and ^{1}H$_i$ dipolar interactions, $d_i = (\mu_0 h \gamma_H \gamma_F)/(8\pi^2 r^3_{FHi})$, μ_0 is the permeability of free space, h is Planck's constant, r_{FHi} is the distance between the ^{19}F nucleus and a specific ^{1}H nuclei (denoted with the subscript i), $c = \Delta\delta_{eff}\gamma_F B_0/3^{1/2}$ is the CSA interaction, $\Delta\delta_{eff} = \Delta\delta(1+\eta^2/3)^{1/2}$ is the generalized CSA, $\Delta\delta$ is the chemical shielding anisotropy, η is the chemical shielding asymmetry, B_0 is the static magnetic field strength, and R_{ex} is the exchange contribution to R_2.

The spectral density function can be expressed using the model-free formalism:[16–19]

$$\begin{aligned}J(\omega) = \frac{2}{5}\Bigg\{O_f^2\left[\frac{O_s^2\tau_c}{1+(\omega\tau_c)^2} + \frac{(1-O_s^2)\tau_{e,s}}{1+(\omega\tau_{e,s})^2}\right] \\ + \frac{(1-O_f^2)\tau_e}{1+(\omega\tau_e)^2}\Bigg\}\end{aligned} \tag{9}$$

where $O^2 = O_f^2 O_s^2$ with $O_f^2$and O_s^2 being the order parameters for internal motions on a fast ($\tau_e \leq 250$ ps) and slow ($\tau_{e,s} > 250$ ps) timescale, respectively, and τ_c

is the overall rotational correlation time. In the absence of motions on slower timescales, $O_s^2 = 1$. In case $O_s^2 < 1$, it is assumed that τ_e is sufficiently small to make the third term of eqn (9) negligible. Thus, two alternative representations are used involving the parameter sets $\{O_f^2, \tau_e\}$ or $\{O_f^2, O_s^2, \tau_{e,s}\}$.

Here, we perform model-free fits using an effective dipolar interaction constant $1/(r_{FHeff})^3$ and a single set of model-free parameters for each ^{19}F site. We also test the agreement between the experimental relaxation data and back-calculated relaxation rates derived from MD simulations wherein each specific ^{19}F and ^1H_i dipolar interaction is described with an individual model-free order parameter.

Protein expression and purification. Galectin-3C was expressed and purified following published protocols,[20,21] yielding a protein stock solution of 16 mg/ml in buffer consisting of 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, 140 mM NaCl, 2.7 mM KCl, pH 7.3, 2 mM ethylenediaminetetraacetic acid (EDTA), 4 mM tris(2-carboxyethyl)phosphine hydrochloride (TCEP), and 150 mM lactose. The protein stock solution was stored at 278 K.

Ligand synthesis and purification. The two diastereomeric compounds (2R)- and (2S)-2-hydroxy-3-(4-(3-fluorophenyl)-1H-1,2,3-triazol-1-yl)-propyl) 2,4,6-tri-O-acetyl-3-deoxy-3-(4-(3-fluorophenyl)-1H-1,2,3-triazol-1-yl)-1-thio-β-D-galactopyranoside (denoted R and S, respectively) have been described before.[14]

NMR sample preparation. ^{15}N-labeled galectin-3C was expressed, purified, and prepared as described previously.[14] Ligand stock solutions were prepared by dissolving the ligand in neat dimethyl sulfoxide (DMSO) to a concentration of 8.2 mM and 35 mM for S and R, respectively. The NMR samples were prepared with 0.2 mM ligand and 0.33 mM protein for S-galectin-3C and 0.26 mM protein for R-galectin-3C in 5 mM 4-(2-hydroxyethyl)-1-piperazinethanesulfonic acid (HEPES) buffer pH 7.4; thus, the ligand:protein ratio was kept below 1 so that the population of bound ligand was greater than 99% in both cases. Additional DMSO was added to yield a total concentration of 5% DMSO and 8% D_2O was added for the field-frequency lock.

^{19}F NMR relaxation experiments and data processing. ^{19}F relaxation experiments were acquired on NMR spectrometers operating at static magnetic field strengths (B_0) of 11.7 T, 14.1 T, and 16.4 T. In experiments acquired at 11.7 T and 14.1 T ^1H-decoupling was achieved with a single ^1H 180° pulse in the middle of the relaxation period, whereas experiments acquired at 16.4 T utilized WALTZ-65 decoupling.[22] Longitudinal ^{19}F inversion recovery experiments were acquired using relaxation delays of (2×0, 0.1, 0.2, 2×0.5, 0.8, 1.3, 1.9, 4) s at 11.7 T, (0, 0.1, 0.2, 0.4, 0.7, 1, 1.5, 2, 3) s at 14.1 T, and (0.01, 0.02, 0.04, 0.05, 0.08, 2×0.1, 0.2, 0.3, 0.5, 0.8, 2×1.0, 2.0, 3.0, 5.0) s at 16.4 T. Carr-Purcell-Meiboom-Gill (CPMG) transverse relaxation experiments were acquired using a CPMG refocusing frequency of ν_{CPMG} = 500 Hz, and relaxation delays of (0, 0.004, 2×0.008, 0.012, 0.016, 0.02, 0.028, 0.04) s at 11.7 T, (0, 0.004, 0.008, 0.012, 0.016, 0.02, 0.028, 0.036, 0.044) s at 14.1 T, and (0, 0.002, 2×0.004, 0.006, 2×0.008, 0.012, 0.016, 2×0.02, 0.028, 0.032, 0.04) s at 16.4 T. ^1H decoupling was applied during the CPMG relaxation period in order to suppress cross-correlated dipole-dipole/CSA relaxation.[23,24] Relaxation compensated CPMG relaxation dispersion experiments[25] were acquired at 16.4 T using a constant time period of 10 ms and CPMG refocusing frequencies of ν_{CPMG} = (0, 100, 3×200, 400, 600, 2×800, 1000, 1200, 2×1400, 1600, 2×1800, 2000) Hz.

All NMR relaxation datasets were processed using NMRPipe.[26] Each spectrum was processed twice, using an exponential window function matched to the peak width at half height ('matched filter') for each individual peak, followed by zero filling to twice the data size. All spectra were subsequently imported into MATLAB for further processing using in-house scripts. The volume of each peak was integrated using a Lorentzian lineshape, and the volume uncertainty estimated via 10,000 Monte-Carlo (MC) simulations for each peak, using the standard deviation of the baseline noise as an estimate of the uncertainty of the peak height. R_1 and σ_{FH} rate constants were determined by fitting the decoupled inversion recovery curves to eqn (4). R_2 rate constants were estimated by fitting the decoupled CPMG decay curves to mono-exponential decays, eqn (5). ^{19}F CPMG relaxation dispersions were fitted to the Carver-Richards equation[27] for two-state exchange in the form presented by Davis et al.[28] Parameter uncertainties were estimated from 10,000 MC simulations for each site. Due to the application of matched filters, the baseline noise was artificially low. This in turn propagates to the estimated parameter uncertainties. For this reason, these uncertainty estimates were scaled to achieve a reduced χ^2 value of 1.0 for each experimental fit.

Model-free analysis of ^{19}F relaxation data. Model-free (MF) parameters were optimized using an in-house MATLAB script. The dipole–dipole coupling constants (d_i) were calculated for all protons within different cut-off distances (defined for each dataset in the Results section) from each fluorine; distances were either measured directly in the X-ray crystal structures of the complexes with PDB-IDs 6QGF (R) and 6QGE (S), or calculated as averages over the MD trajectories. In addition, the coupling constants estimated from MD simulations were multiplied with the order parameters of the corresponding F–H vectors. The τ_c values reported previously[14] were adjusted as

described[29] to account for differences in sample viscosity due to the different DMSO concentrations of the samples used in this study compared to the previous one (see SI for details). The adjusted τ_c values were 8.14 ns and 8.17 ns for R-galectin-3C and S-galectin-3C, respectively, and were used as fixed parameters during MF optimization. We fitted 13 different models that include subsets of the following parameters: generalized order parameter squared, $O^2 = O_f^2 O_s^2$; fast time-scale order parameter, O_f^2; slow time-scale internal correlation time, $\tau_{e,s}$; fast time-scale internal correlation time τ_e; chemical shift anisotropy, CSA; chemical exchange contribution, R_{ex}. The MF models are:

M1 = $\{O_f^2\}$

M2 = $\{O_f^2, \tau_e\}$

M3 = $\{O_f^2, \tau_e, \text{CSA}\}$

M4 = $\{O_f^2, \tau_e, R_{ex}\}$

M5 = $\{O_f^2, \tau_e, \text{CSA}, R_{ex}\}$

M6 = $\{O_f^2, \tau_e; \text{fixed } R_{ex}\}$

M7 = $\{O_f^2, \tau_e, \text{CSA}; \text{fixed } R_{ex}\}$

M8 = $\{O^2, O_f^2, \tau_{e,s}\}$

M9 = $\{O^2, O_f^2, \tau_{e,s}, \text{CSA}\}$

M10 = $\{O^2, O_f^2, \tau_{e,s}, R_{ex}\}$

M11 = $\{O^2, O_f^2, \tau_{e,s}, \text{CSA}, R_{ex}\}$

M12 = $\{O^2, O_f^2, \tau_{e,s}; \text{fixed } R_{ex}\}$

M13 = $\{O^2, O_f^2, \tau_{e,s}, \text{CSA}; \text{fixed } R_{ex}\}$

In models M1–M7, $O_s^2 = 1$, and in models M8–M13, $\tau_e = 0$, see eqn (9). In models M1, M2, M4, M6, M8, M10, and M12, the CSA parameter was fixed to generalized CSA[30] calculated from the values determined by solid state NMR for the R- and S-galectin-3C complexes (Kalabekova, Quinn, Gronenborn, Akke & Polenova, to be published). Models M6, M7, M12, and M13 include an R_{ex} contribution that is fixed to the value estimated by ^{19}F CPMG relaxation dispersion experiments. Grid searches were performed for 10^5 randomly selected grid points per model, distributed over the intervals O^2 = [0.3, 1.0], τ_e = [0, 500] ps, CSA = [0, 150] ppm, R_{ex} = [0, 500] s^{-1}/ω_F^2. Fitted R_{ex} values are assumed to scale quadratically with the static magnetic field strength. Optimization was performed using MATLAB's quasi-Newton algorithm built into the fminunc unrestrained minimization function, which was allowed to run until a local minimum was found. The use of an un-restrained optimization routine allowed the standard error (SE) of each fitted variable to be estimated as the square roots of the diagonal of the inverse Hessian matrix, which contrasts with constrained optimization routines, where the estimated Hessian can be inaccurate[1]. However, due to the un-restrained optimization routine, model selection was done via assessment of the physical plausibility of the parameter set for a given model and using the Bayesian information criterion as defined for model-free optimization by d'Auvergne and Gooley.[31]

Molecular dynamics simulations and analysis. MD simulations of R- and S-galectin-3C complexes have been reported previously.[14] Briefly, MD simulations were performed using the Amber14 software suite.[32] Two different conformers of the S ligand observed in the crystal structure (denoted S_A and S_B) were simulated separately. Each complex (R, S_A, S_B) was run as 10 separate 10-ns long trajectories with different starting conditions.[14]

We used the *ccptraj* program[33] to analyze the MD trajectories with respect to the intramolecular dynamics of the ^{19}F–^{1}H interactions. We analyzed both the angular and radial variation in the F–H internuclear vectors to determine the overall order parameter of each vector as the product of the angular (O_θ^2) and radial (O_r^2) order parameters:[34] $O^2 \approx O_\theta^2 \cdot O_r^2$, where O_θ^2 was determined using the iRED protocol[35,36] and O_r^2 was calculated as:[34]

$$O_r^2 = \langle 1/r^3 \rangle^2 / \langle 1/r^6 \rangle \qquad (10)$$

where r is the distance between the F and H atoms, and the angular brackets denote averaging over the MD trajectory. The relative contribution to the ^{19}F dipolar relaxation from each hydrogen within 5 Å from a given fluorine was calculated.

RESULTS AND DISCUSSION

We used solution-state ^{19}F spin relaxation experiments to characterize the dynamics of two diastereomeric ligands bound to galectin-3C using the model-free formalism.[16,17] Each ligand contains two fluorine atoms, one at each end of the molecule, which serve as suitable probes for studying the conformational dynamics of the ligands bound to galectin-3C. The molecular structures of ligands R and S are shown in Figure 1 together with an overview of the superimposed three-dimensional structures of the two complexes, as well as the solution-state ^{19}F NMR spectra of the ligands in the protein-bound state. The two diastereomers have virtually identical structure in the bound complex with galectin-3C, except that the peripheral, solvent-exposed part of the ligand deviates somewhat as a consequence of the different geometries at the stereocenter and also exhibit greater flexibility in the X-ray crystal structure (Fig. 1B). In the protein-bound state, the chemical shifts of the two fluorines in both R and S are different (Fig. 1C), whereas the difference is much smaller in the unbound state (peaks at –112.76 ppm and –112.78 ppm for both R and S; data not shown). Both ligands give rise to a broad peak at δ = –113.48 ppm (both R and S) and a sharper peak at δ = –113.02 ppm (R) and δ = –112.84 ppm (S). Given the highly similar structures of the complexes at the site of the buried fluorine, we expect that the broad peak

originates from this site. The peripheral fluorine, on the other hand, has slightly different average environment in the two complexes (Fig. 1B), which likely is reflected by the different chemical shifts of the sharper peaks (Fig. 1C).

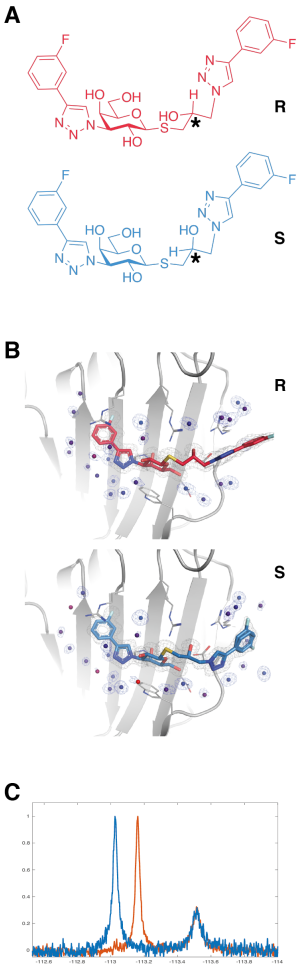


Figure 1. Overview of the R and S ligands. (A) Chemical structures of R (red) and S (blue). The stereocenter is marked with an asterisk. (B) X-ray crystal structures of the ligands bound in the carbohydrate binding site of galectin-3C; PDB-IDs: 6QGF (R) and 6QGE (S). Carbon atoms are colored red (R) or blue (S). Non-carbon atoms are colored: N, dark blue; O, red; S, yellow; F, pale blue. Ligand-coordinating protein side chains are shown in stick representation, while the protein backbone is shown as a ribbon diagram. Water oxygens are shown as spheres. The buried end ("in", left-hand side) of the ligand is virtually identical in the R- and S-complexes. The solvent-exposed end ("out", right-hand side) of the ligand differs between the R- and S-complexes. The X-ray crystal structure of the S-complex reveals two orientations of the fluoro-phenyl ring at the solvent-exposed ("out") end of the ligand. (C) ^{19}F NMR spectra of R (red) and S (blue) ligands bound to galectin-3C. The sharp signal (left, "out") has different chemical shifts in the two complexes, whereas the broad signal (right, "in") has identical chemical shift.

Indeed, solid-state MAS NMR 2D ^{13}C-^{19}F and hCH HETCOR spectra reveal that the broad peak exhibits a greater number of dipolar couplings with ^{1}H nuclei in the protein, including several site-specifically assigned resonances, thereby providing firm support for the assignment of the broad peak to the buried fluorine atom (data not shown; Kalabekova, Quinn, Gronenborn, Akke & Polenova, to be published). Hereafter the buried ^{19}F nucleus will be denoted 'in', while the peripheral is called 'out'; i.e., the different nuclei in the two ligands are referred to as R-in, R-out, S-in, and S-out.

^{19}F chemical shift anisotropy. As a first step toward interpreting ^{19}F relaxation data in terms of intramolecular dynamics, we used CSA values of the two fluorine atoms in each ligand previously determined by solid-state MAS NMR (Kalabekova, Quinn, Gronenborn, Akke & Polenova, to be published). Table 1 shows the chemical shift anisotropies and asymmetries of the two fluorine sites in R and S. The values of $\Delta\delta$, as well as η, are similar for R-in and S-in, and for R-out and S-out, which is expected given the high similarities of the three-dimensional structures of the R- and S-complex (Fig. 3B). The chemical shielding tensor is highly asymmetric, as expected from previous work on meta-fluoro-substituted phenylalanine, where $\Delta\delta$ and η range between 81–84 ppm and 0.65–0.76, respectively.[37] The CSA data for R and S deviate from these model values, a result that is explained by the specific interactions between the fluorine and functional groups in the protein. We assume that the CSA values determined for the micro-crystalline state represent the CSA in solution as well, although differences may occur due to interactions in the crystalline state (e.g., crystal contacts) not present in solution, or differences in conformational dynamics and vibrational averaging effects.[38]

Table 1. Chemical shift anisotropies of ^{19}F nuclei in R- and S-galectin-3C complexes.

Site [a]	$\Delta\delta$ (ppm) [b]	η [c]	$\Delta\delta_{eff}$ (ppm) [d]
R-out	53.2 ± 0.5	0.9	59.9 ± 0.5
R-in	73.6 ± 0.6	0.7	79.4 ± 0.6
S-out	51.9 ± 0.3	0.9	58.5 ± 0.3

S-in	75.0 ± 0.5	0.6	79.4 ± 0.5

[a] "in" denotes buried fluorine atom; "out" denotes peripheral fluorine atom. [b] $\Delta\delta = \delta_{zz}-(\delta_{xx}+\delta_{yy})/2$. [c] $\eta = 3(\delta_{xx}-\delta_{yy})/(2\Delta\delta)$. [d] $\Delta\delta_{eff} = \Delta\delta(1+\eta^2/3)^{1/2}$.

^{19}F relaxation experiments. We acquired ^{19}F relaxation data at three static magnetic field strengths: 11.7 T, 14.1 T, and 16.4 T, using ^{1}H-decoupled ^{19}F inversion recovery R_1 relaxation experiments and CPMG R_2 relaxation experiments. Determining R_2 by fitting eqn (5) to the CPMG decay curves was straightforward, whereas fitting eqn (4) to the inversion recovery curves to determine R_1 and σ_{FH} posed some problems. We initially performed fits using the full equation (5), but σ_{FH} shows great covariance with ϕ (when the latter is included as a fitting parameter) and is generally poorly determined (see SI for further discussion and results). By contrast, R_1 is well determined and does not show great variation between different fitting protocols (with or without σ_{FH} and/or ϕ). We note that σ_{FH} in many cases fits to very small values near zero (SI Table S1), suggesting that spin diffusion among ^{1}H spins renders ^{19}F–^{1}H cross relaxation inefficient, as has been observed for other proteins.[10]

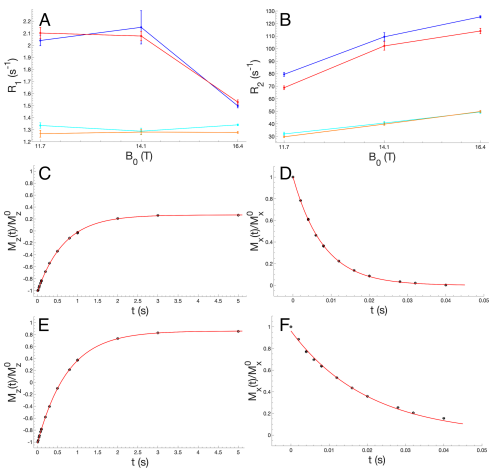


Figure 2. ^{19}F relaxation data. (A) R_1 and (B) R_2 relaxation rate constants plotted versus static magnetic field strength, B_0. Red, R-in; blue, S-in; orange, R-out; cyan, S-out. (C–F) Representative relaxation data measured at B_0 = 16.4 T. (C, E) R_1 ^{1}H-decoupled inversion recovery experiments. (D, F) R_2 CMPG experiments. (C, D) Data for S-in. (E, F) Data for S-out.

Thus, we proceeded to fit the inversion recovery curves to a simple mono-exponential function of R_1. Figure 2A, B show the resulting R_1 and R_2 relaxation rate constants as a function of B_0 for both ligands and Table 2 details the fitted values.

As can be seen, the two different fluorine positions (e.g., comparing S-in with S-out) on the ligand exhibit dramatically different relaxation behavior (Fig. 2A, B), whereas the same positions in the two different ligands (e.g., comparing R-in with S-in) show similar relaxation rate constants. The largest difference is observed for the R_2 values of the buried fluorine atoms at 16.4 T: R_2 = 125.3 ± 0.9 s^{-1} for S-in and R_2 = 114 ± 2 s^{-1} for R-in. This difference cannot be explained solely by millisecond exchange contributions to R_2, because CPMG relaxation dispersion experiments acquired at 16.4 T result in significant and roughly equal exchange rates, R_{ex} = 13 ± 4 s^{-1} for S-in and 11 ± 4 s^{-1} for R-in (see SI: Fig. S2 and Tables S3,4). In the absence of exchange, R_2 is dominated by the CSA, which contributes 70%–85% of the total R_2 between 11.7 T and 16.4 T. However, R-in and S-in have very similar CSA values (Table 1) that cannot explain the difference. Instead, there might exist differences between R-in and S-in in fast exchange contributions to R_2 that are not quenched by the CPMG refocusing field. We return to this issue below.

Table 2. Relaxation rate constants of ^{19}F nuclei in R- and S-galectin-3C complexes.

Site	B_0 (T)	R_1 (s^{-1})	R_2 (s^{-1})
R-out	11.7	1.27 ± 0.03	29.7 ± 0.6
R-out	14.1	1.28 ± 0.02	39.8 ± 0.9
R-out	16.4	1.28 ± 0.01	50 ± 1
R-in	11.7	2.10 ± 0.05	69 ± 1
R-in	14.1	2.08 ± 0.04	102 ± 3
R-in	16.4	1.53 ± 0.02	114 ± 2
S-out	11.7	1.34 ±0.02	32 ± 2
S-out	14.1	1.29 ± 0.03	41 ± 1
S-out	16.4	1.34 ± 0.01	50 ± 1
S-in	11.7	2.04 ± 0.04	79 ± 2
S-in	14.1	2.2 ± 0.1	109 ± 3
S-in	16.4	1.50 ± 0.02	125 ± 1

The R_1 relaxation rate constant is dominated by ^{19}F–^{1}H dipolar relaxation, while the CSA makes only a minor contribution to the overall relaxation. The CSA contribution to R_1 is less than 5% at 11.7 T and 10% at 16.4 T, as calculated from eqns (6–7). Therefore, R_1 should decrease monotonously with B_0, cf. eqn (6). Thus, the similar values measured at 11.7 T and 14.1 T for the buried ^{19}F spins (Fig. 2A; Table 2) are unexpected, possibly indicating that one of the data points suffers from systematic errors. Most likely, the data measured at 14.1 T is incorrect, possibly due to poor

temperature calibration; see further below under the next section.

Two compounding factors contribute to the difference in relaxation rate constants between the buried and peripheral sites. First, the buried ^{19}F has many more dipole interactions with surrounding ^{1}H spins in the protein, which increase the relaxation rate constants, whereas the peripheral ^{19}F has relatively few protein ^{1}H spins within close range. Second, the buried site is likely to be more rigidly held in the binding pocket, whereas the peripheral ^{19}F has little steric hindrance from protein atoms restricting its conformational dynamics. The large difference in relaxation rates between the buried and peripheral sites suggests significant differences in their dynamics. Nonetheless, assessing the relative motional amplitudes experienced by the two sites requires careful analysis of the differences in proton density around the ^{19}F nuclei.

Model-free analysis. We performed MF analysis using the B_0 dependent R_1 and R_2 data as input. Galectin-3C in complex with either R or S has a nearly isotropic diffusion tensor, $D_{||}/D_{\perp} = 1.1$.[14] Here we make the approximations that the global rotational diffusion is isotropic. With this approximation it is permissible to use an effective CSA, $\Delta\delta_{eff}$ in eqns (6–7), as described by Dayie et al.[30] The effective CSA in solution could conceivably be different from that in the crystalline form due to intermolecular contacts, particularly for the peripheral (out) fluorine, which extends towards the neighboring protein molecule in the crystal. In addition, conformational dynamics and vibrational averaging effects might differ between the solution and crystalline states.[38] To account for these caveats, we included a subset of MF models where the CSA was treated as a free fitting parameter. In addition to using different MF models, we also performed separate fits using F–H distances derived from either the X-ray crystal structure or the MD simulation. In the latter case we also included MD-derived order parameters specific to each F–H_i vector; importantly, this approach implies that the fitted O^2 value serves as an additional scaling factor, rather than describing the actual order parameter of the ligand. Furthermore, we included a subset of MF fits using exchange-free R_2 rate constants, i.e., where the estimated R_{ex} contribution had been subtracted from R_2 prior to MF fitting (see SI for details on this procedure). Finally, due to the potential problems with the R_1 data measured at 14.1 T (described above), we performed separate MF analyses using only data obtained at 11.7 T and 16.4 T, as well as data obtained at all three fields. However, the resulting MF parameters are very similar, and we center the discussion below on the three-field data.

All MF fits involved extensive grid searches prior to optimization. Our use of an unconstrained local optimization routine can result in optimized parameters that correspond to unphysical values (e.g., $O^2 > 1$); all such results were discarded.

Table 3 presents the final, best-fit MF parameters obtained using F–H distances derived from the X-ray crystal structures. Below, the results of the MF fits will be discussed in more detail for each individual fluorine position.

Table 3. Model-free parameters of ^{19}F nuclei in R- and S-galectin-3C complexes based on the X-ray structure.

Site	M[a]	O^2	τ_e (ps)	$\Delta\delta_{eff}$ (ppm)	R_{ex}/ω_F^2 (s/rad^2)
R-out	M3	0.58± 0.01	290±40	81.5±0.3	
R-in	M4	0.65± 0.02	35±4	79.4	330±20
S-out	M3	0.56± 0.01	152±9	84.9±0.9	
S-in	M4	0.57± 0.02	40±3	79.4	437±12

[a] Selected model yielding the best fit.

In all cases, the best fit is obtained for the dataset created by subtracting from R_2 the R_{ex} value determined by relaxation dispersion. It turns out that O^2 appears to be relatively independent of which model is selected, when comparing different models that give similar residuals, χ^2. In all cases, M3 or M4 yield the best fits. In M4, where $\Delta\delta_{eff}$ is fixed to the value determined by solid-state NMR, there is a contribution from R_{ex} that exceeds $\Delta\delta_{eff}$ in magnitude. This result is conspicuous given that the expected R_{ex} contributions were subtracted from the R_2 rates prior to MF fitting. To further investigate this issue, we performed exploratory calculations to determine what ranges of MF parameters result in the observed B_0 dependence of R_1 and R_2 (cf. Fig. 2A).

These calculations make it clear that the R_2 data for R-in and S-in cannot be explained without a significant R_{ex} term of approximately 40 s^{-1}, which is considerably greater than that determined by CPMG relaxation dispersion (R_{ex} = 11–13 s^{-1}, see above). Thus, it appears there is an additional exchange process (on top of the exchange between free and bound states) that must be faster than the CPMG refocusing frequency. This conclusion agrees with previous observations that the side chain of R144, which sits directly on top of the fluoro-phenyl ring in the conventional X-ray structure and forms the closest approach of any protein atom to the fluorine, actually undergoes extensive conformational exchange: the guanidino group of R144 is unobservable in the ^{1}H-^{15}N HSQC spectrum due to exchange broadening, and the ensemble-refined X-ray

diffraction data shows large-scale fluctuations of this side chain.[14]

Furthermore, the calculations indicate that the R_1 data at 11.7 T and 16.4 T for R-in and S-in can be explained by $O^2 \approx 0.75–0.8$ and $\tau_e \approx 10$ ps, but only if the dipolar interactions with protons in the R144 side chain are reduced below the values expected from the conventional X-ray structure or MD simulation — again pointing towards extensive conformational fluctuations of the R144 side chain.

The corresponding calculations for R-out and S-out indicate that the experimental R_1 and R_2 data can be reproduced by $O^2 \approx 0.4–0.55$ and τ_e on the order of 100 ps, without the addition of R_{ex} terms beyond those detected by CPMG relaxation dispersion. These results are in reasonable agreement with the MF optimization results (Table 3).

In the case of the peripheral fluorine, the fitted value of $\Delta\delta_{eff}$ (Table 3) differs from that determined by solid-state NMR (Table 1), but agrees perfectly with that obtained for model compounds, suggesting that the CSA determined by solid-state NMR is affected by crystal contacts between the peripheral end of the ligand and the neighboring protein in the crystal, whereas in solution the peripheral ^{19}F does not form any interaction with protein atoms. In contrast, the buried fluorine fits best using a fixed CSA value determined by solid-state NMR, in agreement with the expectation that the environment of this end of the ligand is nearly identical in the crystalline and solution states.

In addition, we performed additional MF fits using F–H distances and F-H order parameters derived from the MD trajectories. The final, best fitted of these models are presented in table 4.

Table 4. Model-free parameters of ^{19}F nuclei in R- and S-galectin-3C complexes based on the MD trajectories.

Site	M[a]	O^2	τ_e (ps)	$\Delta\delta_{eff}$ (ppm)	R_{ex}/ω_F^2 (s/rad^2)
R-out	M3	0.54±0.01	350±30	85.2±0.4	
R-in	M4	0.82±0.03	81±6	79.4	250±20
S-out	M3	0.38[b]	826.8[b]	99.55[b]	
S-in	M4	0.78±0.02	107±6	79.4	338±15

[a] Selected model yielding the best fit. [b] Estimated uncertainty <1e-4

The fitted O^2 parameters correspond well to the expected difference between in- and out- fluorine positions. However, their physical interpretation differ, as these O^2 values represent the additional scaling of the O^2 values estimated via MD. This can, to some extent, point to the under sampling of molecular motions by the MD simulations.

Concluding remarks. Quantitative analysis of ^{19}F relaxation data is challenging as the scarcity of previous reports attest. In the present study, we encountered several challenges that required careful examination of what data to include in the analysis. Taking into account previous knowledge of the dynamic behavior of the two ligand–galectin-3C complexes, obtained from both NMR and ensemble-refined X-ray diffraction data, we are able to rationalize the experimental ^{19}F relaxation data in terms of model-free parameters. While the model-free optimizations based on the conventional X-ray structures suggest only minor differences in the conformational dynamics of the buried and peripheral ends of the protein-bound ligands, our model calculations that take into account the flexibility of the R144 side chain indicate clear differences between the buried and peripheral fluorine atoms in their dynamics, with order parameters of approximately 0.75–0.8 and 0.4–0.55, respectively. In summary, careful examination of the all data available has enabled us to reach a consistent picture of ligand dynamics in the protein-bound state. Future work will be aimed to refine and improve on these studies.

ASSOCIATED CONTENT

Supporting Information. Tables of R_1 and σ_{FH}, sample viscosity and corrected τ_c, This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

* Corresponding author: Mikael Akke, e-mail: mikael.akke@bpc.lu.se.

Present Addresses

† Present address: Medicinal biochemistry and Biophysics, Molecular Structural Biology, Karolinska Institute, Stockholm, Sweden.

‡ Present address: Institute of Physics, Biophysics, Martin-Luther-University Halle-Wittenberg, D-06120 Halle (Saale), Germany.

ACKNOWLEDGMENT

We gratefully acknowledge the Swedish NMR Center at the University of Gothenburg for 500 MHz and 700 MHz NMR measurement time and Ulrika Brath for assistance with these experiments. Protein production was carried out by the Lund Protein Production Platform at Lund University. We thank Roza Kalabekova, Caitlin M. Quinn, Angela Gronenborn and Tatyana Polenova for communicating ^{19}F CSA values prior to publication, and Majda Misini Ignjatovic and Ulf Ryde for providing MD trajectories. This work was supported by the Knut and Alice Wallenberg Foundation (KAW 2013.022) and the Swedish Research Council (2018-4995 to MA).

REFERENCES

(1) Akke, M. Conformational Dynamics and Thermodynamics of Protein-Ligand Binding Studied by NMR Relaxation. *Biochemical Society Transactions* **2012**, *40*, 419–423.

(2) Akke, M.; Brüschweiler, R.; Palmer, A. G. NMR Order Parameters and Free Energy: An Analytical Approach and Its Application to Cooperative Ca2+ Binding by Calbindin D9k. *J Am Chem Soc* **1993**, *115*, 9832–9833.

(3) Diehl, C.; Engström, O.; Delaine, T.; Håkansson, M.; Genheden, S.; Modig, K.; Leffler, H.; Ryde, U.; Nilsson, U. J.; Akke, M. Protein Flexibility and Conformational Entropy in Ligand Design Targeting the Carbohydrate Recognition Domain of Galectin-3. *J Am Chem Soc* **2010**, *132*, 14577–14589.

(4) Frederick, K. K.; Marlow, M. S.; Valentine, K. G.; Wand, A. J. Conformational Entropy in Molecular Recognition by Proteins. *Nature* **2007**, *448*, 325–330.

(5) Tzeng, S.-R.; Kalodimos, C. G. Dynamic Activation of an Allosteric Regulatory Protein. *Nature* **2009**, *462*, 368–372.

(6) Richardson, P. Fluorination Methods for Drug Discovery and Development. *Expert Opinion on Drug Discovery* **2016**, *11* (10), 983–999. https://doi.org/10.1080/17460441.2016.1223037.

(7) Jhoti, H.; Cleasby, A.; Verdonk, M.; Williams, G. Fragment-Based Screening Using X-Ray Crystallography and NMR Spectroscopy. *Curr Opin Chem Biol* **2007**, *11*, 485–493.

(8) Gronenborn, A. Small, but Powerful and Attractive: 19F in Biomolecular NMR. *Structure* **2022**, *30*, 6–14.

(9) Lu, M.; Ishima, R.; Polenova, T.; Gronenborn, A. M. 19F NMR Relaxation Studies of Fluorosubstituted Tryptophans. *Journal of Biomolecular NMR* **2019**, *73* (8–9), 401–409. https://doi.org/10.1007/s10858-019-00268-y.

(10) Luck, LindaA.; Vance, JosephE.; O'Connell, ThomasM.; London, RobertE. 19F NMR Relaxation Studies on 5-Fluorotryptophan- and Tetradeutero-5-Fluorotryptophan-Labeled E. Coli Glucose/Galactose Receptor. *Journal of Biomolecular NMR* **1996**, *7* (4). https://doi.org/10.1007/BF00200428.

(11) Shi, P.; Wang, H.; Xi, Z.; Shi, C.; Xiong, Y.; Tian, C. Site-Specific ^{19} F NMR Chemical Shift and Side Chain Relaxation Analysis of a Membrane Protein Labeled with an Unnatural Amino Acid. *Protein Science* **2011**, *20* (1), 224–228. https://doi.org/10.1002/pro.545.

(12) Hoang, J.; Prosser, R. S. Conformational Selection and Functional Dynamics of Calmodulin: A ^{19} F Nuclear Magnetic Resonance Study. *Biochemistry* **2014**, *53* (36), 5727–5736. https://doi.org/10.1021/bi500679c.

(13) Peng, J. W. Cross-Correlated 19F Relaxation Measurements for the Study of Fluorinated Ligand-Receptor Interactions. *Journal of Magnetic Resonance* **2001**, *153*, 32–47. https://doi.org/10.1006/jmre.2001.2422.

(14) Verteramo, M. L.; Stenström, O.; Ignjatović, M. M.; Caldararu, O.; Olsson, M. A.; Manzoni, F.; Leffler, H.; Oksanen, E.; Logan, D. T.; Nilsson, U. J.; Ryde, U.; Akke, M. Interplay between Conformational Entropy and Solvation Entropy in Protein-Ligand Binding. *J Am Chem Soc* **2019**, *141*, 2012–2026. https://doi.org/10.1021/jacs.8b11099.

(15) Cavanagh, J.; Fairbrother, W. J.; Palmer, A. G.; III; Skelton, N. J. *Protein NMR Spectroscopy: Principles and Practice*, 2nd ed.; Academic Press, 1995.

(16) Halle, B.; Wennerström, H. Interpretation of Magnetic Resonance Data from Water Nuclei in Heterogeneous Systems. *Journal of Chemical Physics* **1981**, *75*, 1928–1943.

(17) Lipari, G.; Szabo, A. Model-Free Approach to the Interpretation of Nuclear Magnetic Resonance Relaxation in Macromolecules. I. Theory and Range of Validity. *J Am Chem Soc* **1982**, *104*, 4546–4559.

(18) Halle, B. The Physical Basis of Model-Free Analysis of NMR Relaxation Data from Proteins and Complex Fluids. *Journal of Chemical Physics* **2009**, *131*, 224507.

(19) Clore, G. M.; Szabo, A.; Bax, A.; Kay, L. E.; Driscoll, P. C.; Gronenborn, A. M. Deviations from the Simple Two-Parameter Model-Free Approach to the Interpretation of Nitrogen-15 Nuclear Magnetic Relaxation of Proteins. *J Am Chem Soc* **1990**, *112*, 4989–4991.

(20) Diehl, C.; Engström, O.; Delaine, T.; Håkansson, M.; Genheden, S.; Modig, K.; Leffler, H.; Ryde, U.; Nilsson, U. J.; Akke, M. Protein Flexibility and Conformational Entropy in Ligand Design Targeting the Carbohydrate Recognition Domain of Galectin-3. *J Am Chem Soc* **2010**, *132*, 14577–14589.

(21) Diehl, C.; Genheden, S.; Modig, K.; Ryde, U.; Akke, M. Conformational Entropy Changes upon Lactose Binding to the Carbohydrate Recognition Domain of Galectin-3. *Journal of Biomolecular NMR* **2009**, *45*, 157–169. https://doi.org/10.1007/s10858-009-9356-5.

(22) Zhou, Z.; Kümmerle, R.; Qiu, X.; Redwine, D.; Cong, R.; Taha, A.; Baugh, D.; Winniford, B. A New Decoupling Method for Accurate Quantification of Polyethylene Copolymer Composition and Triad Sequence Distribution with 13C NMR. *J. Magn. Res.* **2007**, *187*, 225–233.

(23) Palmer, A. G.; Skelton, N. J.; Chazin, W. J.; Wright, P. E.; Rance, M. Suppression of the Effects of Cross-Correlation between Dipolar and Anisotropic Chemical Shift Relaxation Mechanisms in the Measurement of Spin-Spin Relaxation Rates. *Molecular Physics* **1992**, *75*, 699–711.

(24) Kay, L. E.; Nicholson, L. K.; Delagio, F.; Bax, A.; Torchia, D. A. Pulse Sequences for Removal of the Effects of Cross Correlation between Dipolar and Chemical-Shift Anisotropy Relaxation Mechanisms on the Measurement of Heteronuclear T1 and T2 Values in Proteins. *Journal of Magnetic Resonance* **1992**, *97*, 359–375.

(25) Loria, J. P.; Rance, M.; Palmer, A. G. A Relaxation-Compensated Carr-Purcell-Meiboom-Gill Sequence for Characterizing Chemical Exchange by NMR Spectroscopy. *J Am Chem Soc* **1999**, *121*, 2331–2332.

(26) Delaglio, F.; Grzesiek, S.; Vuister, G. W.; Zhu, G.; Pfeifer, J.; Bax, A. NMRPipe: A Multidimensional Spectral Processing System Based on UNIX Pipes. *Journal of Biomolecular NMR* **1995**, *6*, 277–293.

(27) Carver, J. P.; Richards, R. E. A General Two-Site Solution for the Chemical Exchange Produced Dependence of T2 upon the Carr-Purcell Pulse Separation. *Journal of Magnetic Resonance* **1972**, *6*, 89–105.

(28) Davis, D. G.; Perlman, M. E.; London, R. E. Direct Measurements of the Dissociation-Rate Constant for Inhibitor-Enzyme Complexes via the T1r and T2 (CPMG) Methods. *Journal of Magnetic Resonance, Series B* **1994**, *104*, 266–275.

(29) Wallerstein, J.; M, A. Minute Additions of DMSO Affect Protein Dynamics Measurements by NMR Relaxation Experiments through Significant Changes in Solvent. *ChemPhysChem* **2019**, *20*, 326–332.

(30) Dayie, K. T.; Wagner, G. Carbonyl Carbon Probe of Local Mobility in 13C, 15N-Enriched Proteins Using High-Resolution Nuclear Magnetic Resonance. *J Am Chem Soc* **1997**, *119*, 7797–7806.

(31) d'Auvergne, E. J.; Gooley, P. R. The Use of Model Selection in the Model-Free Analysis of Protein Dynamics. *Journal of Biomolecular NMR* **2003**, *25*, 25–39.

(32) Case, D. A.; Berryman, J. T.; Betz, R. M.; Cerutti, D. S.; Cheatham, T. E., III; Darden, T. A.; Duke, R. E.; Giese, T. J.; Gohlke, H.; Goetz, A. W.; Homeyer, N.; Izadi, S.; Janowski, P.; Kaus, J.; Kovalenko, A.; Lee, T E.; LeGrand, S.; Li, P.; Luchko, T.;, P. A. AMBER. University of California: San Francisco 2015.

(33) Roe, D. R.; Cheatham III, T. E. PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Synamics Trajectory Data. *J Chem Theory Com* **2013**, *9*, 3084–3095. https://doi.org/10.1021/ct400341p.

(34) Brüschweiler, R.; Roux, B.; Blackledge, M.; Griesinger, C.; Karplus, M.; Ernst, R. R. Influence of Rapid Intramolecular Motion on NMR Cross-Relaxation Rates. A Molecular Dynamics Study of Antamanide in Solution. *J Am Chem Soc* **1992**, *114* (7), 2289–2302.

(35) Prompers, J. J.; Brüschweiler, R. Reorientational Eigenmode Dynamics: A Combined MD/NMR Relaxation Analysis Method for Flexible Parts in Globular Proteins. *J Am Chem Soc* **2001**, *123*, 7305–7313.

(36) Prompers, J. J.; Brüschweiler, R. General Framework for Studying the Dynamics of Folded and Nonfolded Proteins by NMR Relaxation Spectroscopy and MD Simulation. *J Am Chem Soc* **2002**, *124*, 4522–4534.

(37) Dürr, U. H. N.; Grage, S. L.; Witter, R.; Ulrich, A. S. Solid State 19F NMR Parameters of Fluorine-Labeled Amino Acids. Part I: Aromatic Substituents. *Journal of Magnetic Resonance* **2008**, *191*, 7–15. https://doi.org/10.1016/j.jmr.2007.11.017.

(38) Tang, S.; Case, D. A. Vibrational Averaging of Chemical Shift Anisotropies in Model Peptides. *Journal of Biomolecular NMR* **2007**, *38* (3), 255–266. https://doi.org/10.1007/s10858-007-9164-8.

Supplementary Information

Conformational Dynamics of Designed Drug-like Ligands Bound to Proteins Determined by ^{19}F NMR Relaxation and Molecular Dynamics Simulations

Sven Wernersson[1], Olof Stenström[1], Simon Birgersson[1], Magdalena Riad [1], Ulrich Weininger[1§], Göran Carlström[2], Ulf J. Nilsson[2], Mikael Akke[1,*]

[1] Division of Biophysical Chemistry, Center for Molecular Protein Science, Department of Chemistry, Lund University, Sweden; [2] Centre for Analysis and Synthesis, Department of Chemistry, Lund University, Sweden

Table S1. R_1 and σ_{FH} rate constants of ^{19}F nuclei in R- and S-galectin-3C complexes

Site	B_0 (T)	R_1 (s^{-1})	σ_{FH} (s^{-1})
S-out	11.7	1.33 ± 0.02	0.005 ± 0.02
S-out	14.1	1.29 ± 0.03	–0.26 ± 0.02
S-out	16.4	1.34 ± 0.01	–0.177 ± 0.004
R-out	11.7	1.27 ± 0.03	0.006 ± 0.02
R-out	14.1	1.28 ± 0.02	–0.24 ± 0.02
R-out	16.4	1.28 ± 0.01	–0.124 ± 0.005
S-in	11.7	2.04 ± 0.04	0.007 ± 0.02
S-in	14.1	2.2 ± 0.1	–1.8 ± 0.1
S-in	16.4	1.50 ± 0.02	–1.03 ± 0.02
R-in	11.7	2.10 ± 0.05	–0.05 ± 0.02
R-in	14.1	2.08 ± 0.04	–1.61 ± 0.04
R-in	16.4	1.53 ± 0.02	–1.03 ± 0.02

Rotational correlation time, τ_c

This study uses rotational correlation times estimated from previous model-free (MF) analysis of ^{15}N relaxation data.[1] The NMR samples used to determine the rotational diffusion tensor included DMSO at concentrations of 1.2% and 4.3% for ligand R and S, respectively. In the present ^{19}F relaxation study, the DMSO concentration was adjusted to 5% in both samples. The viscosity of water-DMSO mixtures is highly dependent on the volume fraction of DMSO: in the regime 0–20%, the viscosity increases by a factor of 3.[2] To correct for the difference in DMSO concentration and hence in viscosity, we performed a linear interpolation of temperature-corrected literature data following published protocols.[2] We then calculated corrected values of τ_c for the NMR samples used in the present study using the ratio of the estimated viscosities. The resulting viscosity and τ_c values are reported in Table S2.

Table S2. Calculated sample viscosities (η) and corrected rotational correlation times (τ_c)

Sample	R–galectin-3C	S–galectin-3C
^{15}N, [DMSO] (v/v)	1.2%	4.3%
^{15}N, η (mPa s)	0.9241	0.9936
^{15}N, τ_c (ns)	7.45	8.04
^{19}F, [DMSO] (v/v)	5.0%	5.0%
^{19}F, η (mPa s)	1.0098	1.0098
η_{19F}/η_{15N}	1.093	1.016
^{19}F, τ_c (ns)	8.14	8.17

^{19}F CPMG relaxation dispersion data analysis

The ^{19}F CPMG relaxation dispersion data acquired at a magnetic field strength of 16.4 T were fitted to the Carver-Richards equation[3] as presented in Davis et al.[4] Fitting was performed using various combinations of

free and fixed exchange populations (p_B) and chemical shift differences ($\Delta\omega$), under the assumption that the chemical exchange is dominated by equilibrium exchange between free and protein-bound states. Exchanging populations were estimated from the ligand and protein concentrations in the NMR sample, given previously measured dissociation constants.[5] Chemical shifts differences were calculated from the chemical shifts of the ^{19}F peaks in protein-free and protein saturated NMR samples measured under otherwise identical sample conditions. As the ^{19}F NMR peaks in the protein-free NMR samples are unassigned, the CPMG data were fitted using both of the two possible chemical shift differences and the best-fitting alternative was chosen. The fitted parameters are presented in Table S3. The fitted dispersion curves obtained using fixed populations and chemical shift differences are presented in Figure S1.

Table S3 R-in. ^{19}F exchange parameters from CPMG relaxation dispersion experiments for R-in

Fixed parameter	χ^2_{red}	k_{ex} (10^2 s^{-1})	p_B	$\Delta\omega$ (ppm)	$R_{2,0}$ (s^{-1})
none	0.96	33 ± 15	0.87 ± 0.03	0.19 ± 0.05	120 ± 2
p_B	0.76	15 ± 4.6	0.985	0.66 ± 0.2	120 ± 2
$\Delta\omega$	0.66	0.51 ± 0.9	0.7 ± 0.2	0.77	121 ± 1
p_B + $\Delta\omega$	0.74	14 ± 4	0.985	0.77	120 ± 2

Table S3 R-out. ^{19}F exchange parameters from CPMG relaxation dispersion experiments for R-out

Fixed parameter	χ^2_{red}	k_{ex} (10^2 s^{-1})	p_B	$\Delta\omega$ (ppm)	$R_{2,0}$ (s^{-1})
none	0.43	11±5	0.98±0.01	0.40±0.07	56±0.3
p_B	1.61	15±4	0.985	0.55±0.04	56±0.3
$\Delta\omega$	0.40	10±4	0.98±0.005	0.412	56±0.3
p_B + $\Delta\omega$	4.8	16±4	0.985	0.412	57±0.3

Table S3 S-in. ^{19}F exchange parameters from CPMG relaxation dispersion experiments for S-in

Fixed parameter	χ^2_{red}	k_{ex} (10^2 s^{-1})	p_B	$\Delta\omega$ (ppm)	$R_{2,0}$ (s^{-1})
none	0.14	5.2±7	0.96±0.2	0.54±0.2	123±2
p_B	0.46	22±13	0.985	0.73±0.1	122±3
$\Delta\omega$	0.16	0.51±0.5	0.5±0.2	0.751	123±1
p_B + $\Delta\omega$	0.43	21±8	0.985	0.751	122±2

Table S3 S-out. ^{19}F exchange parameters from CPMG relaxation dispersion experiments for S-out

Fixed parameter	χ^2_{red}	k_{ex} (10^2 s^{-1})	p_B	$\Delta\omega$ (ppm)	$R_{2,0}$ (s^{-1})
none	0.48	12±1	0.50±0.01	0.09±0.002	56±0.3
p_B	4.78	15±4	0.985±0	0.50±0.03	55±0.4
$\Delta\omega$	0.62	4.5±2	0.94±0.02	0.280±0	56±0.3
p_B + $\Delta\omega$	16.1	13±4	0.985±0	0.280±0	58±0.3

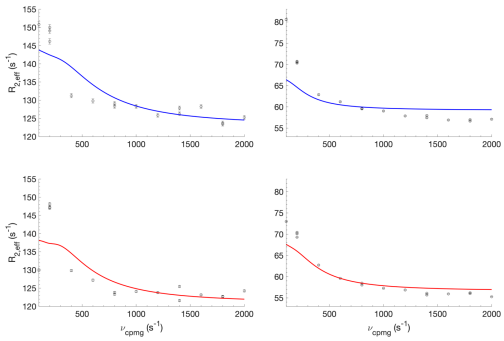


Figure S1. Fitted ^{19}F CPMG relaxation dispersions. Exchange populations and chemical shift differences have been frozen under the assumption that the total exchange is dominated by the protein-ligand binding equilibrium. Top row: Ligand S, bottom row: Ligand R. Left column: Buried Fluorine position (in), right column: peripheral fluorine position (out).

Estimating chemical exchange contributions to R_2

The chemical exchange contributions (R_{ex}) to R_2 at each static magnetic field strength were estimated from the fitted exchange rates to the Carver Richards equation using fixed populations and chemical shift differences (SI table 6). The difference between R_{ex} terms at ν_{cpmg} = [500Hz, 10kHz] was calculated. R_2-experiments all use ν_{cpmg} = 500Hz, and ν_{cpmg} = 10kHz represents $R_{2,0}$. Uncertainties in R_{ex} were propagated using an analytical expression, as Monte Carlo based propagation was found to give unreliable results. The estimated R_{ex} terms are presented in supplementary table 6, and in a static magnetic field strength (B_0) independent unit in table 7, where it is assumed to scale quadratically with B_0.

Table 6: Fitted R_2 relaxation rate constants at CPMG refocusing frequencies (ν_{CPMG}) 500Hz, 10kHz. The Chemical Exchange (R_{ex}) contribution at ν_{CPMG}=500Hz is estimated as the difference between these R_2 rates. The estimation is based on fitting of CT ^{19}F CPMG relaxation dispersion data.

Table 6a, Estimated R_{ex} contributions at a static magnetic field strength of 11.7T.

Ligand	R_2, ν_{CPMG}=500Hz	R_2, ν_{CPMG}=10kHz	R_{ex} (s^{-1}), ν_{CPMG}=500Hz
S-in	129.7±0.8	122±0.3	7.7±0.8
S-out	58.9±0.4	58.0±0.3	0.92±0.5
R-in	126±1	120±0.3	6±1
R-out	59.2±0.5	57.0±0.3	2.2±0.6

Table 6b, Estimated R_{ex} contributions at a static magnetic field strength of 14.1T.

Ligand	R_2, ν_{CPMG}=500Hz	R_2, ν_{CPMG}=10kHz	R_{ex} (s^{-1}), ν_{CPMG}=500Hz
S-in	133±1	122±0.3	11±1
S-out	59.3±0.4	58.0±0.3	1.3±0.5
R-in	129±2	120±0.3	9±2
R-out	60.2±0.6	57.0±0.3	3.2±0.7

Table 6c, Estimated R_{ex} contributions at a static magnetic field strength of 16.4T.

Fluorine	R_2, ν_{CPMG}=500Hz	R_2, ν_{CPMG}=10kHz	R_{ex} (s^{-1}), ν_{CPMG}=500Hz
S-in	135±4	122±2	13±4
S-out	59.8±0.5	58.0±0.3	1.8±0.6
R-in	131±3	120±2	11±4
R-out	61.2±0.7	57.0±0.3	4.2±0.7

Table 7: Estimated Chemical Exchange (R_{ex}) contributions to R_2 expressed in a B_0 independent unit, assuming that R_{ex} scales quadratically with B0.

Ligand	S-in	S-in	S-in	S-in
R_{ex} (s/rad^2), ν_{CPMG}=500Hz	87±26	12±4	71±24	28±5

Estimation of antiphase contributions to R_2

The effective relaxation rate (Γ^*) in a relaxation period with a CPMG-element, starting from transverse single quantum magnetization (S_x) under evolution from heteronuclear coupling can be described as [6]

$$\Gamma* = 0.5\left(1+\frac{\sin(2\pi J_{IS}\tau)}{2\pi J_{IS}\tau}\right)\Gamma_S + 0.5\left(1-\frac{\sin(2\pi J_{IS}\tau)}{2\pi J_{IS}\tau}\right)\Gamma_{IS}$$

Where J_{IS} is the J-coupling constant, τ is half the time between two 180-degree pulses, Γ_S is the in-phase magnetization relaxation constant, and Γ_{IS} is the antiphase magnetization relaxation constant. If the term $\sin(2\pi J_{IS}\tau)/2\pi J_{IS}\tau$ is close to 1, the contribution to the effective relaxation constant from antiphase magnetization is negligible. Assuming the HF J-couplings on the aromatic rings are similar to previous data on fluorines situated in an ortho-position [7], we have at most $J \approx 12$ Hz. For all R_2 datasets, the refocusing delay τ was set to 1ms, which means $\sin(2\pi J_{IS}\tau)/2\pi J_{IS}\tau$ = 0.999. This is likely sufficient to minimize any antiphase contributions from evolution of the heteronuclear couplings. In addition, all transverse NMR relaxation experiments used decoupling during the relaxation period, which should suppress any build-up of antiphase magnetization from transverse cross-correlated DD-CSA relaxation (η_{xy}).

iRED analysis of Fluorine order parameters

We analyzed existing Molecular dynamics (MD) simulations [1], of Galectin 3C (Gal3C) in complex with either the R or S ligand . The S-ligand has two conformers in the crystal structure, S_A and S_B. These conformers were used for separate MD simulations, meaning that there are three sets of trajectories, S_A, S_B and R, consisting of 10 individual trajectories 10ns long each, with varied solvation boxes and starting trajectories. Further details of the MD simulations are found in Verteramo et al [1]. We analyzed the angular order parameter of the F-C vector using the iRED [1] protocol. Each trajectory was analyzed separately, and the order parameters from all trajectories in each set averaged. The average and standard deviation for each order parameter is listed in table 8.

Supplementary Table 8: Order parameters (O^2) for the F-C bond vector, as estimated from MD simulations via the iRED protocol for each ligand and fluorine position. The sub-labels A, B for ligand S refers to different starting conformations used in the MD simulations. Each value refers to the average and standard deviation of ten independent trajectories.

Starting conformation	O^2 (in-F)	O^2 (out-F)
GAL3C - S_A (PDB-ID: 6QGE)	0.915±0.01	0.587±0.25
GAL3C - S_B (PDB-ID: 6QGE)	0.918±0.02	0.536±0.31
GAL3C - R (PDB-ID: 6QGF)	0.921±0.02	0.756±0.14

Estimation of the ^{1}H-contribution to dipolar relaxation

The squared dipolar coupling constant (d^2) can be defined as:

$d^2 = ((\mu_0 \hbar \gamma_F \gamma_H)/4\pi)^2 \cdot f(r_{HF})$, with either $f(r_{HF}) = \overline{1/r^6}$ or $f(r_{HF}) = \left(\overline{1/r^3}\right)^2$

Where $\mu 0$ is the vacuum permeability constant, $\hbar$ is the reduced Planck constant, and γ_F, γ_H are the gyromagnetic ratios of fluorine and hydrogen respectively. The different averages of the HF-distances (r_{HF}) in the denominator represents time averaging in X-ray crystallography and NMR spectroscopy respectively. Values of d^2 for the different fluorine positions are presented in table 9. These are calculated using either the distances in the x-ray crystal structure directly, the average distance in the MD-simulation, or the average distance in the MD simulation weighted with the effective order parameter of each HF-vector (these order parameters are shown in figure 2).

Supplementary Table 9: Effective dipolar coupling constants calculated for each fluorine position. The x-ray columns refer to d^2-values calculated using distances from the respective x-ray crystal structure, MD refers to d^2-values calculated using average distances from the MD simulations, and MD+HF-O^2 refers to average MD distances weighted with the order parameter of the individual HF vector.

Fluorine	d^2, x-ray (10^7 rad^2/s^2)	d^2, MD (10^7 rad^2/s^2)	d^2, MD+HF-O^2 (10^7 rad^2/s^2)
S-in	1.151	0.9953	0.7880
S-out	0.4377	0.3533	0.2215
R-in	1.149	1.023	0.8698
R-out	0.3608	0.3455	0.2842

The following tables (10-15) presents the contributions from individual proton atoms to the d^2 constants calculated from MD for the different fluorine hydrogen distances on average within 5Å of the fluorine of interest. In addition to these tables, figure 2 shows the relative strength of the dipolar coupling constant contribution from remote protons as a function of distance to the fluorine atom.

Table 10: Calculated d^2 constants for the R-in fluorine, sorted by average distance to the fluorine atom. Residue 251 is the ligand. The second column contains mean distance, $\bar{r}$. The third column contains d^2 calculated from the sixth power of the average distance. The fourth column contains d^2 calculated from the third power of the average squared distance

Residue-Atom	$\bar{r}$ (Å)	$d2 \propto \overline{1/r^6}$ (10^7 rad^2/s^2)	$d^2 \propto \left(\overline{1/r^3}\right)^2$ (10^7 rad^2/s^2)
251-HI	2.58	176.18	174.55
251-HG	2.59	169.81	168.19
144-HB2	2.76	142.15	131.21
144-HG3	3.00	111.89	103.27
144-HD2	3.21	94.51	86.15
160-HD22	3.39	43.60	39.21
160-HD21	3.41	46.97	40.62
146-HB3	3.50	54.09	46.01
146-HB1	3.60	42.64	35.75
145-H	3.69	25.67	23.34
145-HA	3.75	22.89	21.27
146-HB2	3.84	36.61	30.50
146-H	3.98	15.37	14.37
144-HG2	4.10	26.74	24.75

238-HA2	4.11	15.32	13.40
238-HA3	4.14	15.65	13.64
144-HD3	4.20	21.17	19.35
144-HB3	4.31	8.43	8.19
144-HE	4.48	9.23	8.22
251-HF	4.50	6.11	6.10
146-HA	4.67	5.17	5.05
144-HA	4.68	5.38	5.16
251-HA	4.79	4.22	4.20

Table 11: Calculated d^2 constants for the R-out fluorine, sorted by average distance to the fluorine atom. Residue 251 is the ligand. The second column contains mean distance, $\bar{r}$. The third column contains d^2 calculated from the sixth power of the average distance. The fourth column contains d^2 calculated from the third power of the average squared distance.

Residue-Atom	$\bar{r}$ (Å)	$d2 \propto \overline{1/r^6}$ (10^7 rad^2/s^2)	$d^2 \propto \left(\overline{1/r^3}\right)^2$ (10^7 rad^2/s^2)
251-HI'	2.58	174.42	172.83
251-HG'	2.60	168.21	166.58
251-HF'	4.51	6.05	6.04

Table 12: Calculated d^2 constants for the S_A-in fluorine, sorted by average distance to the fluorine atom. Residue 251 is the ligand. The second column contains mean distance, $\bar{r}$. The third column contains d^2 calculated from the sixth power of the average distance. The fourth column contains d^2 calculated from the third power of the average squared distance.

Residue-Atom	$\bar{r}$ (Å)	$d2 \propto \overline{1/r^6}$ (10^7 rad^2/s^2)	$d^2 \propto \left(\overline{1/r^3}\right)^2$ (10^7 rad^2/s^2)
251-HI	2.58	175.02	173.39
251-HG	2.59	169.63	168.00
144-HB2	2.84	127.27	114.53
144-HG3	2.87	122.05	110.42
144-HD2	3.16	88.47	74.48
146-HB1	3.57	48.77	37.65
160-HD22	3.58	33.38	29.35
145-HA	3.60	28.83	26.52
146-HB2	3.63	44.74	34.43
146-HB3	3.64	46.51	35.15
160-HD21	3.64	34.15	28.33
145-H	3.70	26.42	23.29
146-H	3.84	19.00	17.62
238-HA2	3.90	21.42	18.17
238-HA3	3.97	20.16	17.05
144-HG2	4.29	17.47	11.83
144-HB3	4.37	10.07	8.38
144-HE	4.37	10.72	9.11
144-HD3	4.48	10.19	8.10
251-HF	4.50	6.10	6.09
146-HA	4.64	5.35	5.22
144-HA	4.69	5.63	5.21
144-HH11	4.74	6.80	5.69
251-HA	4.80	4.17	4.14
144-NH1	4.97	4.76	4.16
237-HB3	4.99	4.08	3.70

Table 13: Calculated d^2 constants for the S_A-out fluorine, sorted by average distance to the fluorine atom. Residue 251 is the ligand. The second column contains mean distance, $\bar{r}$. The third column contains d^2 calculated from the sixth power of the average distance. The fourth column contains d^2 calculated from the third power of the average squared distance.

Residue-Atom	$\bar{r}$ (Å)	$d2 \propto \overline{1/r^6}$ (10^7 rad^2/s^2)	$d^2 \propto \left(\overline{1/r^3}\right)^2$ (10^7 rad^2/s^2)
251-HI'	2.57	179.28	177.58
251-HG'	2.60	168.62	167.03
251-HF'	4.50	6.07	6.06
251-HA'	4.77	4.38	4.35

Table 14: Calculated d^2 constants for the S_B-in fluorine, sorted by average distance to the fluorine atom. Residue 251 is the ligand. The second column contains mean distance, $\bar{r}$. The third column contains d^2 calculated from the sixth power of the average distance. The fourth column contains d^2 calculated from the third power of the average squared distance.

Residue-Atom	$\bar{r}$ (Å)	$d2 \propto \overline{1/r^6}$ (10^7 rad^2/s^2)	$d^2 \propto \left(\overline{1/r^3}\right)^2$ (10^7 rad^2/s^2)
251-HI	2.58	175.25	173.60
251-HG	2.60	168.92	167.29
144-HB2	2.89	124.90	111.43
144-HG3	3.00	115.44	101.21

144-HD2	3.18	88.53	75.73
160-HD22	3.55	34.63	30.49
145-HA	3.59	28.67	26.54
146-HB1	3.61	47.07	36.75
146-HB3	3.62	46.90	35.78
160-HD21	3.63	35.28	29.12
145-H	3.65	31.54	26.82
146-HB2	3.67	43.12	32.80
146-H	3.84	18.81	17.52
238-HA2	3.92	21.32	17.97
238-HA3	3.96	20.45	17.44
144-HE	4.27	11.93	10.14
144-HB3	4.29	16.53	12.39
144-HG2	4.32	11.27	9.34
144-HD3	4.46	11.07	8.61
251-HF	4.50	6.10	6.09
144-HA	4.62	9.06	7.19
146-HA	4.64	5.38	5.24
144-HH11	4.78	6.08	5.22
251-HA	4.80	4.18	4.15
144-NH1	4.97	4.55	4.06

Table 15: Calculated d^2 constants for the S_B-out fluorine, sorted by average distance to the fluorine atom. Residue 251 is the ligand. The second column contains mean distance, $\bar{r}$. The third column contains d^2 calculated from the sixth power of the average distance. The fourth column contains d^2 calculated from the third power of the average squared distance.

Residue-Atom	$\bar{r}$ (Å)	$d2 \propto \overline{1/r^6}$ (10^7 rad^2/s^2)	$d^2 \propto \left(\overline{1/r^3}\right)^2$ (10^7 rad^2/s^2)
251-HI'	2.58	172.12	173.41
251-HG'	2.60	168.09	166.47
251-HF'	4.50	6.07	6.06

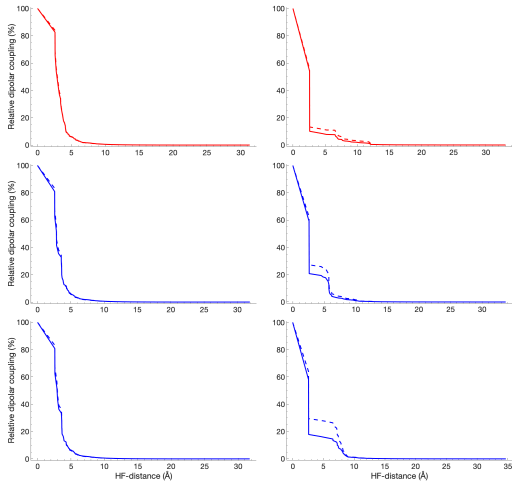


Figure 2: Relative contribution of remote protons to the total dipolar coupling constant to each fluorine as a function of distance. The relative strength at a given distance is the sum of all contributions at that distance and further out. Left: in-position fluorines, right: out-position fluorines. Top: ligand R, middle: ligand S_A, bottom: ligand S_B. Solid lines: HF distances averaged as $1/\langle r^6\rangle$, dotted lines: HF distances averaged as $1/\langle r^2\rangle^3$.

Order parameters of the HF vectors from MD

The radial and angular order parameters of each HF vector with a maximum distance of 5Å was calculated using the methods outlined above. The average and standard deviation of each ligand is presented in tables 16 and 17 for the fluorine in- and out-position respectively. The corresponding distributions of order parameters are plotted as histograms in figures 3 and 4.

Table 16: Average and standard deviation of the radial (O_r^2), angular (O_a^2) and product ($O_r^2 \cdot O_a^2$) HF order parameters for the fluorine in-position of each ligand. Ligands S_A, S_B refers to the different starting conformations of the S-ligand.

Ligand	average O_r^2	average O_a^2	average $O_r^2 \cdot O_a^2$
R	0.92±0.05	0.89±0.14	0.82±0.13
SA	0.87±0.08	0.84±0.18	0.73±0.17
SB	0.87±0.08	0.84±0.18	0.73±0.17

Table 17: Average and standard deviation of the radial (O_r^2), angular (O_a^2) and product ($O_r^2 \cdot O_a^2$) HF order parameters for the fluorine out-position of each ligand. Ligands S_A, S_B refers to the different starting conformations used for the MD simulations of the S-ligand.

Ligand	average O_r^2	average O_a^2	average $O_r^2 \cdot O_a^2$

R	0.99±0.00	0.86±0.10	0.86±0.10
SA	0.99±0.00	0.69±0.16	0.69±0.16
SB	0.99±0.00	0.63±0.11	0.62±0.10

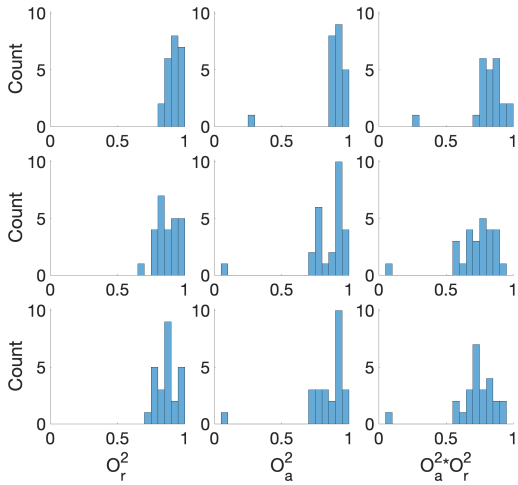


Figure 3: Histograms of radial (O_r^2), angular (O_a^2) and product ($O_r^2 \cdot O_a^2$) HF order parameter distributions for the fluorine in-position. Top: ligand R, middle: Ligand S_A, bottom: ligand S_B. Ligands S_A, S_B refers to the different starting conformations of the S-ligand.

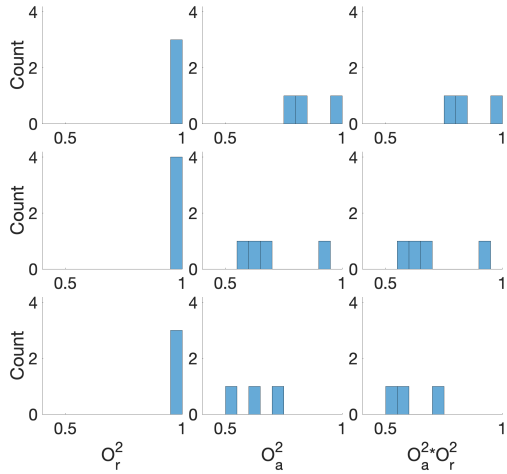


Figure 4: Histograms of radial (O_r^2), angular (O_a^2) and product ($O_r^2 \cdot O_a^2$) HF order parameter distributions for the fluorine out-position. Top: ligand R, middle: Ligand S_A, bottom: ligand S_B. Ligands S_A, S_B refers to the different starting conformations of the S-ligand.

Model free model fitting results

The following tables present the complete set of fitted MF models for fluorine S-in (tables 18-21), S-out (tables 22-25), R-in (tables 26-29), R-out (tables 30-33).

Table 18: Fluorine S-in MF fitting results, using X-ray derived HF-distances (PDB-ID: 6QGE (2)), including all protons within 5Å of the fluorine.

Model	χ^2	BIC	O^2	τ (ps)	CSA (ppm)	R_{ex}/ω_F^2 (s^3/rad^2)	O_f^2	τ_s (ps)
1	3450	3450	1.02±4e-3	-	79.40	-	-	-
2	257	261	1660	30e6	79.40	-	-	-
3	61.6	67.0	0.57±0.02	13±1	136±2	-	-	-
4	61.5	69	0.57±0.02	39±3	79.40	500±10	-	-
5	72.8	80.0	0.34±0.03	1080±70	0±16	785±5	-	-
6	257	261	1660	30e6	79.40	87.86	-	-
7	61.6	67.0	0.57±0.02	13±1	136±2	87.86	-	-
8	577	583	1.5	-	79.40	-	0.74±0.01	830±40
9	68.9	76.0	0.40 [a]	-	147.90 [a]	-	0.59 [a]	7248 [a]
10	67.1	74.3	0.44±0.04	-	79.40	540±20	0.71±0.02	2500±1000
11	61.4	70.4	0.57±0.02	-	81±20	500±100	1.00 [a]	37.50 [a]
12	577	583	1.46±0.01	-	79.40	87.86	0.74±0.01	830±40
13	68.9	76.0	0.40 [a]	-	147.90[a]	87.86	0.59 [a]	7248 [a]

[a]The estimated uncertainty is <0.0001

Table 19: Fluorine S-in MF fitting results, using X-ray derived HF-distances (PDB-ID: 6QGE (2)), including all protons within 5Å of the fluorine. Estimated chemical exchange contributions to R2 removed prior to analysis.

Model	χ^2	BIC	O^2	τ (ps)	CSA (ppm)	R_{ex}/ω_F^2 (s^3/rad^2)	O_f^2	τ_s (ps)
1	1276	1279	0.85±0.005	-	79.40	-	-	-
2	958	962	0.90±0.003	-320±30	79.40	-	-	-
3	45.5	50.9	0.57±0.02	15±1	130±2	-	-	-
4	45.6	51.0	0.57±0.02	40±3	79.40	437±12	-	-
5	45.3	52.5	0.57±0.02	24±14	100±30	300±300	-	-
6	958	962	0.90±0.003	-320±30	79.40	0.00	-	-
7	45.5	50.9	0.57±0.02	15±1	130±2	0.00	-	-
8	409	414	1.25±0.01	-	79.40	-	0.74±0.01	830±60
9	51.8	59.0	0.34 [a]	-	145.56 [a]	-	0.59 [a]	8442.96 [a]
10	51.0	58.1	0.4±0.2	-	79.40	500±30	0.67±0.05	5000±4000
11	50.6	59.6	0.40 [a]	-	135.55 [a]	70.25 [a]	0.61 [a]	5417.18 [a]
12	409	414	1.25 [a]	-	79.40	0.00	0.74±0.01	830±60
13	51.8	59.0	0.34 [a]	-	145.56 [a]	0.00	0.59 [a]	8443 [a]

[a]The estimated uncertainty is <0.0001

Table 20: Fluorine S-in MF fitting results, using MD derived HF-distances, including all protons within 5Å of the fluorine.

Model	χ^2	BIC	O^2	τ (ps)	CSA (ppm)	R_{ex}/ω_F^2 (s^3/rad^2)	O_f^2	τ_s (ps)
1	1410	1420	1.28±0.01	-	79.40	-	-	-
2	328	331	1.58±0.01	1290±70	79.40	-	-	-
3	61.6	67.0	0.80±0.02	39±2	115±2	-	-	-
4	61.8	67.2	0.79±0.02	107±6	79.40	400±10	-	-
5	61.4	68.5	0.80±0.02	60±30	100±20	200±200	-	-
6	328	331	1.58±0.01	1290±70	79.40	87.86	-	-
7	61.6	67.0	0.80±0.02	39±2	115±2	87.86	-	-
8	294	299	1.55±0.01	-	79.40	-	1.07±0.01	950±90
9	67.	74.6	0.63±0.09	-	123±4	-	0.84±0.03	4200±3100
10	61.2	68.3	0.81±0.02	-	79.40	395±10	55.53[a]	0.30[a]
11	65.8	74.7	0.71±0.03	-	110±20	200±200	0.84±0.05	607[a]
12	294	299	1.55±0.01	-	79.40	87.86	1.07±0.01	954±90
13	67.4	74.6	0.63±0.09	-	123±4	87.86	0.84±0.03	4200±3100

[a]The estimated uncertainty is <0.0001

Table 21: Fluorine S-in MF fitting results, using MD derived HF-distances, including all protons within 5Å of the fluorine. Estimated chemical exchange contributions to R_2 removed prior to analysis.

Model	χ^2	BIC	O^2	τ (ps)	CSA (ppm)	R_{ex}/ω_F^2 (s^3/rad^2)	O_f^2	τ_s (ps)
1	522	524	1.12±0.01	-	79.40	-	-	-
2	225	228	1.38±0.01	1300±100	79.40	-	-	-
3	45.6	50.9	0.79±0.02	45±3	110±2	-	-	-
4	45.8	51.1	0.78±0.02	107±6	79.40	338±15	-	-
5	49.85	57.01	0.59±0.07	2000±1000	65±15	500±100	-	-
6	225	228	1.38±0.01	1300±100	79.40	0.00	-	-
7	45.56	50.93	0.79±0.02	45±3	110±2	0.00	-	-
8	209	214	1.37±0.01	-	79.40	-	1.05±0.01	1000±100
9	48.9	56.1	0.70±0.02	-	117±3	-	0.82±0.02	657[a]
10	50.0	57.1	0.58±0.09	-	79.40	400±20	0.94±0.04	3000±2000
11	49.9	58.9	0.57±0.09	-	70±60	500±400	1.0±0.2	3000±2000
12	209	214	1.37±0.01	-	79.40	0.00	1.05±0.01	1000±100
13	48.9	56.1	0.70±0.02	-	117±3	0.00	0.82±0.02	657[a]

[a]The estimated uncertainty is <0.0001

Table 22: Fluorine S-out MF fitting results, using X-ray derived HF-distances (PDB-ID: 6QGE (2)), including all protons within 5Å of the fluorine.

Model	χ^2	BIC	O^2	τ (ps)	CSA (ppm)	R_{ex}/ω_F^2 (s^3/rad^2)	O_f^2	τ_s (ps)
1	1251	1252	1.40±0.01	-	58.50	-	-	-
2	1992	1996	0.46±0.01	291±5	58.50	-	-	-
3	10.2	15.5	0.57±0.01	147±9	86±1	-	-	-
4	86.8	92.2	-0.20±0.03	380±10	58.50	384±9	-	-
5	8.95	16.1	0.65±0.02	60±50	120±40	-400±400	-	-
6	1992	1996	0.46±0.01	291±5	58.50	11.64	-	-
7	10.2	15.5	0.57±0.01	147±9	86±1	11.64	-	-
8	229	235	0.95±0.01	-	58.50	-	429.68 [a]	0.12 [a]
9	16.1	23.3	0.28 [a]	-	122.35 [a]	-	0.57 [a]	493 [a]
10	9.56	16.7	0.59±0.02	-	58.50	159±8	66.70 [a]	1.20 [a]
11	20.9	29.8	0.02 [a]	-	266.91 [a]	30±100	0.22 [a]	1956.71 [a]
12	229	235	0.95±0.01	-	58.50	11.64	429.68 [a]	0.12 [a]
13	16.1	23.3	0.28 [a]	-	122.35 [a]	11.64	0.57 [a]	493.36 [a]

[a]The estimated uncertainty is <0.0001

Table 23: Fluorine S-out MF fitting results, using X-ray derived HF-distances (PDB-ID: 6QGE (2)), including all protons within 5Å of the fluorine. Estimated chemical exchange contributions to R_2 removed prior to analysis.

Model	χ^2	BIC	O^2	τ (ps)	CSA (ppm)	R_{ex}/ω_F^2 (s^3/rad^2)	O_f^2	τ_s (ps)
1	1301	1303	1.4±0.01	-	58.50	-	-	-
2	1845	1848	0.44±0.01	291±5	58.50	-	-	-
3	9.98	15.36	0.56±0.01	152±9	84.9±0.9	-	-	-
4	86.22	91.59	-0.20±0.03	380±10	58.50	372±9	-	-
5	8.69	15.85	0.65±0.02	62±50	120±40	-400±400	-	-
6	1845	1848	0.44±0.01	291±5	58.50	0.00	-	-
7	9.98	15.36	0.56±0.01	152±9	85±1	0.00	-	-
8	196	201	0.92±0.01	-	58.50	-	303[a]	0.18 [a]
9	8.53	15.70	0.65±0.02	-	79±1	-	105 [a]	0.46 [a]
10	9.28	16.45	0.59±0.02	-	58.50	148±8	47.1 [a]	1.71 [a]
11	32.87	41.83	0.11 [a]	-	140.21 [a]	31.30 [a]	0.62 [a]	1823 [a]
12	196	201	0.92±0.01	-	58.50	0.00	303 [a]	0.18 [a]
13	8.53	15.70	0.65±0.02	-	79±1	0.00	105 [a]	0.46 [a]

[a]The estimated uncertainty is <0.0001

Table 24: Fluorine S-out MF fitting results, using MD derived HF-distances, including all protons within 5Å of the fluorine.

Model	χ^2	BIC	O^2	τ (ps)	CSA (ppm)	R_{ex}/ω_F^2 (s^3/rad^2)	O_f^2	τ_s (ps)
1	3510	3510	2.11±0.01	-	58.50	-	-	-
2	4150	4150	0.05±0.01	303±4	58.50	-	-	-
3	18.8	24.1	0.44±0.04	700±100	97±3	-	-	-
4	199	204	-0.81±0.05	519±18	58.50	535±2	-	-
5	21.0	28.2	0.1±0.1	690±100	83±2	232±6	-	-
6	4150	4150	0.05±0.01	303±4	58.50	11.64	-	-
7	18.8	24.1	0.44±0.04	700±100	97±3	11.64	-	-
8	8.90	14.3	1.22±0.01	-	58.50	-	3.99 [a]	33.05 [a]
9	22.9	30.1	0.21±0.02	-	130±10	-	0.80±0.1	1328[a]
10	8.75	15.9	1.24±0.04	-	58.50	-7±11	33.34 [a]	2.74 [a]
11	19.5	28.5	0.17±0.2	-	118 [a]	100±200	0.72 [a]	1000±400
12	8.90	14.3	1.22±0.01	-	58.50	11.64	3.99 [a]	33.05 [a]
13	22.9	30.1	0.21±0.02	-	125±12	11.64	0.80±0.1	1328 [a]

[a]The estimated uncertainty is <0.0001

Table 25: Fluorine S-out MF fitting results, using MD derived HF-distances, including all protons within 5Å of the fluorine. Estimated chemical exchange contributions to R_2 removed prior to analysis.

Model	χ^2	BIC	O^2	τ (ps)	CSA (ppm)	R_{ex}/ω_F^2 (s^3/rad^2)	O_f^2	τ_s (ps)
1	3623	3630	2.11±0.01	-	58.50	-	-	-
2	3877	3880	0.02±0.01	304±4	58.50	-	-	-
3	22.3	27.7	0.38[a]	826.8[a]	99.55 [a]	-	-	-
4	198	204	-0.81±0.05	520±20	58.50	520±10	-	-
5	11.5	18.6	0.91±0.02	95±50	180±10	-1900±200	-	-
6	3880	3880	0.02±0.01	304±4	58.50	0.00	-	-
7	22.3	27.7	0.38 [a]	827 [a]	99.55 [a]	0.00	-	-
8	9.01	14.4	1.18±0.01	-	58.50	-	2.54 [a]	75.35 [a]
9	16.2	23.3	0.50 [a]	-	90.38 [a]	-	1.05 [a]	528 [a]
10	8.50	15.7	1.24±0.04	-	58.50	-18±10	13.29 [a]	7.33 [a]
11	18.6	27.5	0.2±0.3	-	98.79 [a]	100±200	0.89 [a]	800±400
12	9.01	14.4	1.18±0.01	-	58.50	0.00	2.54 [a]	75.35 [a]
13	16.2	23.3	0.50 [a]	-	90.38 [a]	0.00	1.05 [a]	528[a]

[a]The estimated uncertainty is <0.0001

Table 26: Fluorine R-in MF fitting results, using X-ray derived HF-distances (PDB-ID: 6QGF (2)), including all protons within 5Å of the fluorine.

Model	χ^2	BIC	O^2	τ (ps)	CSA (ppm)	R_{ex}/ω_F^2 (s^3/rad^2)	O_f^2	τ_s (ps)
1	913	915	0.86±0.01	-	79.40	-	-	-
2	690	694	0.91±0.002	-301±33	79.40	-	-	-
3	65.75	71.12	0.66±0.02	13±2	122±2	-	-	-
4	65.72	71.10	0.66±0.02	35±4	79.40	400±10	-	-
5	65.18	72.35	0.55±0.08	158±90	32±7	680±30	-	-
6	690	694	0.91±0.002	-301±30	79.40	71.4	-	-
7	65.75	71.12	0.66±0.02	13±2	122±2	71.4	-	-
8	283	289	1.27±0.01	-	79.40	-	0.78±0.01	830±70
9	63.69	70.86	0.50±0.09	-	132±6	-	0.67±0.03	6000±4000
10	62.88	70.05	0.4±0.2	-	79.40	480±40	0.70±0.03	8000±5000
11	62.77	71.73	0.4±0.2	-	70±100	600±700	0.7±0.1	6000±4000
12	283	289	1.27±0.01	-	79.40	71.39	0.78±0.01	829±71
13	63.69	70.86	0.50±0.09	-	132±6	71.39	0.67±0.03	5800±4000

Table 27: Fluorine R-in MF fitting results, using X-ray derived HF-distances (PDB-ID: 6QGF (2)), including all protons within 5Å of the fluorine. Estimated chemical exchange contributions to R_2 removed prior to analysis.

Model	χ^2	BIC	O^2	τ (ps)	CSA (ppm)	R_{ex}/ω_F^2 (s^3/rad^2)	O_f^2	τ_s (ps)
1	503	504	0.83±0.01	-	79.40	-	-	-
2	357	360	0.91±0.003	-348±47	79.40	-	-	-
3	62.3	67.7	0.66±0.02	16±2	115±2	-	-	-
4	62.3	67.6	0.65±0.02	35±4	79.40	330±20	-	-
5	62.6	69.7	0.45±0.04	1100±100	0±27	648±9	-	-
6	357	360	0.91±0.003	-350±50	79.40	0.00	-	-
7	62.3	67.7	0.66±0.02	16±2	115±2	0.00	-	-
8	214	219	1.14±0.02	-	79.40	-	0.78±0.01	829±91
9	60.0	67.1	0.54±0.07	-	124±5	-	0.68±0.03	3000±3000
10	58.2	65.3	0.30±0.2	-	79.40	420±40	0.69±0.03	9000±6000
11	58.0	67.0	0.3±0.3	-	68±180	500±1000	0.70±0.12	9000±6000
12	214	219	1.14±0.02	-	79.40	0.00	0.78±0.01	829±91
13	60.0	67.1	0.54±0.07	-	124±5	0.00	0.68±0.03	3000±3000

Table 28: Fluorine R-in MF fitting results, using MD derived HF-distances, including all protons within 5Å of the fluorine.

Model	χ^2	BIC	O^2	τ (ps)	CSA (ppm)	R_{ex}/ω_F^2 (s^3/rad^2)	O_f^2	τ_s (ps)
1	515	516	1.05±0.01	-	79.40	-	-	-
2	183	187	1.36±0.01	940±80	79.40	-	-	-
3	65.7	71.1	0.83±0.02	34±3	108±2	-	-	-
4	65.6	70.9	0.83±0.03	81±6	79.40	330±20	-	-
5	63.1	70.3	0.60±0.05	3700±200	0±40	716±9	-	-
6	183	187	1.36±0.01	940±80	79.40	71.39	-	-
7	65.7	71.1	0.83±0.02	34±3	108±2	71.39	-	-
8	183	189	1.36±0.01	-	79.40	-	1.00±0.01	910±120
9	63.9	71.1	0.67±0.09	-	116±4	-	0.85±0.03	4000±300
10	63.0	70.2	0.48±0.2	-	79.40	410±40	0.87±0.04	8000±5000
11	62.6	71.5	0.35±0.3	-	40±200	700±600	0.90±0.1	9000±7000
12	183	189	1.36±0.01	-	79.40	71.39	1.00±0.01	900±100
13	63.9	71.1	0.67±0.09	-	116±4	71.39	0.85±0.03	4000±3000

Table 29: Fluorine R-in MF fitting results, using MD derived HF-distances, including all protons within 5Å of the fluorine. Estimated chemical exchange contributions to R_2 removed prior to analysis.

Model	χ^2	BIC	O^2	τ (ps)	CSA (ppm)	R_{ex}/ω_F^2 (s^3/rad^2)	O_f^2	τ_s (ps)
1	259	261	1.0±0.1	-	79.40	-	-	-
2	139	143	1.22±0.02	790±90	79.40	-	-	-
3	62.2	67.6	0.83±0.02	41±3	102±2	-	-	-
4	62.1	67.5	0.82±0.03	81±6	79.40	250±20	-	-
5	58.9	66.1	0.58±0.05	3600±200	-0±30	640±10	-	-
6	139	143	1.22±0.02	790±90	79.40	0.00	-	-
7	62.2	67.6	0.83±0.02	41±3	102±2	0.00	-	-
8	139	144	1.22±0.02	-	79.40	-	0.99±0.01	800±100
9	59.4	66.6	0.63±0.1	-	112±5	-	0.85±0.04	5000±4000
10	58.3	65.5	0.40±0.3	-	79.40	360±50	0.86±0.04	9000±5000
11	58.1	67.1	0.5±0.4	-	40±200	600±900	0.9±0.2	6000±4000
12	139	144	1.22±0.02	-	79.40	0.00	0.99±0.01	800±100
13	59.4	66.6	0.63±0.1	-	112±5	0.00	0.85±0.04	5000±4000

Table 30: Fluorine R-out MF fitting results, using X-ray derived HF-distances (PDB-ID: 6QGF (2)), including all protons within 5Å of the fluorine.

Model	χ^2	BIC	O^2	τ (ps)	CSA (ppm)	R_{ex}/ω_F^2 (s^3/rad^2)	O_f^2	τ_s (ps)
1	1754	1756	1.46±0.01	-	59.90	-	-	-
2	3072	3075	0.60±0.01	289±5	59.90	-	-	-
3	9.60	14.98	0.59±0.01	200±20	84.7±0.7	-	-	-
4	54.43	59.80	-0.16±0.04	400±20	59.90	380±10	-	-
5	8.16	15.32	0.73±0.03	120±50	110±20	-300±200	-	-
6	3072	3075	0.60±0.01	289±5	59.90	27.62	-	-
7	9.60	14.98	0.59±0.01	200±20	84.7±0.7	27.62	-	-
8	74.61	79.98	1.07±0.01	-	59.90	-	22.91 [a]	2.32 [a]
9	9.38	16.55	0.59 [a]	-	84.80 [a]	-	0.98[a]	228 [a]
10	8.50	15.66	0.73±0.03	-	59.90	114±10	14.29 [a]	5.45 [a]
11	17.27	26.23	0.22±0.1	-	115.98 [a]	93±90	0.59 [a]	620±90
12	74.61	79.98	1.07±0.01	-	59.90	27.62	22.91 [a]	2.32 [a]
13	9.38	16.55	0.59 [a]	-	84.80 [a]	27.62	0.98[a]	228[a]

[a]The estimated uncertainty is <0.0001

Table 31: Fluorine R-out MF fitting results, using X-ray derived HF-distances (PDB-ID: 6QGF (2)), including all protons within 5Å of the fluorine. Estimated chemical exchange contributions to R_2 removed prior to analysis.

Model	χ^2	BIC	O^2	τ (ps)	CSA (ppm)	R_{ex}/ω_F^2 (s^3/rad^2)	O_f^2	τ_s (ps)
1	1976	1977	1.45±0.01	-	59.90	-	-	-
2	2423	2426	0.51±0.01	290±5	59.90	-	-	-
3	11.1	16.5	0.58±0.01	290±40	81.5±0.3	-	-	-
4	49.9	55.3	-0.16±0.04	400±20	59.90	350±10	-	-
5	6.76	13.9	0.72±0.04	130±50	110±20	-300±200	-	-
6	2423	2426	0.51±0.01	290±5	59.90	0.00	-	-
7	11.1	16.5	0.58±0.01	290±40	81.5±0.3	0.00	-	-
8	45.0	50.4	0.98±0.01	-	59.90	-	80 [a]	0.72 [a]
9	12.2	19.3	0.33 [a]	-	109[a]	-	0.67 [a]	500[a]
10	7.2	14.3	0.73±0.03	-	59.90	88±10	12.21 [a]	6.45 [a]
11	6.79	15.8	0.75±0.05	-	90±20	-100±300	1.29 [a]	85±24
12	45.0	50.4	0.98±0.01	-	59.90	0.00	80 [a]	0.72 [a]
13	12.2	19.3	0.33 [a]	-	109.31 [a]	0.00	0.67 [a]	500 [a]

[a]The estimated uncertainty is <0.0001

Table 32: Fluorine R-out MF fitting results, using MD derived HF-distances, including all protons within 5Å of the fluorine.

Model	χ^2	BIC	O^2	τ (ps)	CSA (ppm)	R_{ex}/ω_F^2 (s^3/rad^2)	O_f^2	τ_s (ps)
1	2890	2890	1.62±0.01	-	59.90	-	-	-
2	4280	4280	0.49±0.01	291±4	59.90	-	-	-
3	11.7	17.1	0.57±0.01	340±30	86.9±0.4	-	-	-
4	82.7	88.1	-0.36±0.04	450±20	59.90	430±10	-	-
5	23.8	30.9	0.31±0.07	808.10 [a]	89±4	120±70	-	-
6	4280	4280	0.49±0.01	291±4	59.90	27.62	-	-
7	11.7	17.1	0.57±0.01	340±30	87±0.4	27.62	-	-
8	20.5	25.9	1.12±0.01	-	59.90	-	312.32 [a]	0.21 [a]
9	18.7	25.9	0.24 [a]	-	129.16 [a]	-	0.61 [a]	800±100
10	8.31	15.5	0.95±0.04	-	59.90	60±10	9.92 [a]	8.53 [a]
11	21.9	30.8	0.12 [a]	-	109.68 [a]	164.12 [a]	0.77 [a]	1173 [a]
12	20.5	25.9	1.12±0.01	-	59.90	27.62	312.32 [a]	0.21 [a]
13	18.7	25.9	0.24 [a]	-	129.16 [a]	27.62	0.61 [a]	800±100

[a]The estimated uncertainty is <0.0001

Table 33: Fluorine R-out MF fitting results, using MD derived HF-distances, including all protons within 5Å of the fluorine. Estimated chemical exchange contributions to R2 removed prior to analysis.

Model	χ^2	BIC	O^2	τ (ps)	CSA (ppm)	R_{ex}/ω_F^2 (s^3/rad^2)	O_f^2	τ_s (ps)
1	3120	3120	1.62±0.01	-	59.90	-	-	-
2	3400	3400	0.39±0.01	292±4	59.90	-	-	-
3	9.85	15.2	0.54±0.01	350±30	85.2±0.4	-	-	-
4	77.5	82.9	-0.36±0.04	450±20	59.90	410±10	-	-
5	20.3	27.4	0.12±0.07	1132 [a]	89.91 [a]	170±60	-	-
6	3400	3400	0.39±0.01	292±4	59.90	0.00	-	-
7	9.85	15.2	0.54±0.01	350±30	85.22±0.4	0.00	-	-
8	10.46	15.8	1.04±0.01	-	59.90	-	8.68 [a]	9.31 [a]
9	17.50	24.7	0.14 [a]	-	150.9 [a]	-	0.51 [a]	1145 [a]
10	7.01	14.2	0.93±0.04	-	59.90	34±10	3.06 [a]	37.60 [a]
11	18.27	27.2	0.08 [a]	-	128 [a]	140±50	0.62 [a]	1262 [a]
12	10.46	15.8	1.04±0.01	-	59.90	0.00	8.68 [a]	9.31 [a]
13	17.50	24.7	0.14 [a]	-	151 [a]	0.00	0.51 [a]	1145 [a]

[a]The estimated uncertainty is <0.0001

REFERENCES

(1) Verteramo, M. L.; Stenström, O.; Ignjatović, M. M.; Caldararu, O.; Olsson, M. A.; Manzoni, F.; Leffler, H.; Oksanen, E.; Logan, D. T.; Nilsson, U. J.; Ryde, U.; Akke, M. Interplay between Conformational Entropy and Solvation Entropy in Protein-Ligand Binding. *Journal of the American Chemical Society* **2019**, *141* (5), 2012–2026. https://doi.org/10.1021/jacs.8b11099.

(2) Wallerstein, J.; Akke, M. Minute Additions of DMSO Affect Protein Dynamics Measurements by NMR Relaxation Experiments through Significant Changes in Solvent Viscosity. *ChemPhysChem* **2019**, *20* (2), 326–332. https://doi.org/10.1002/cphc.201800626.

(3) Carver, J. P.; Richards, R. E. A General Two-Site Solution for the Chemical Exchange Produced Dependence of T2 upon the Carr-Purcell Pulse Separation. *Journal of Magnetic Resonance (1969)* **1972**, *6* (1), 89–105. https://doi.org/10.1016/0022-2364(72)90090-X.

(4) Davis, D. G.; Perlman, M. E.; London, R. E. Direct Measurements of the Dissociation-Rate Constant for Inhibitor-Enzyme Complexes via the T1ϱ and T2 (CPMG) Methods. *Journal of Magnetic Resonance, Series B* **1994**, *104* (3), 266–275. https://doi.org/10.1006/jmrb.1994.1084.

(5) Verteramo, M. L.; Stenström, O.; Ignjatović, M. M.; Caldararu, O.; Olsson, M. A.; Manzoni, F.; Leffler, H.; Oksanen, E.; Logan, D. T.; Nilsson, U. J.; Ryde, U.; Akke, M. Interplay between Conformational Entropy and Solvation Entropy in Protein-Ligand Binding. *J Am Chem Soc* **2019**, *141*, 2012–2026. https://doi.org/10.1021/jacs.8b11099.

(6) Palmer, A. G.; Skelton, N. J.; Chazin, W. J.; Wright, P. E.; Rance, M. Suppression of the Effects of Cross-Correlation between Dipolar and Anisotropic Chemical Shift Relaxation Mechanisms in the Measurement of Spin-Spin Relaxation Rates. *Molecular Physics* **1992**, *75* (3), 699–711. https://doi.org/10.1080/00268979200100511.

(7) Peng, J. W. Cross-Correlated 19F Relaxation Measurements for the Study of Fluorinated Ligand-Receptor Interactions. *Journal of Magnetic Resonance* **2001**, *153* (1), 32–47. https://doi.org/10.1006/jmre.2001.2422.

Paper III

The cosolvent dimethylsulfoxide influences protein–ligand binding kinetics via solvent viscosity effects revealing the success rate of complex formation following diffusive protein–ligand encounter

Sven Wernersson,‡ Simon Birgersson,‡ Mikael Akke*

Division of Biophysical Chemistry, Center for Molecular Protein Science, Department of Chemistry, Lund University, P. O. Box 124, SE-221 00 Lund, Sweden

ABSTRACT: Protein–ligand exchange kinetics determines the duration of biochemical signals and consequently plays an important role in drug design. Binding studies commonly require solubilization of designed ligands in solvents like dimethyl sulfoxide (DMSO), resulting in residual amounts of DMSO following titration of solubilized ligands into aqueous protein samples. Therefore, it is critical to establish whether DMSO influences protein–ligand binding. Here we address the general and indirect effect of DMSO on protein–ligand binding caused by the solvent viscosity, which is strongly dependent on the relative concentrations of DMSO and water. As a model system, we studied the binding of a drug-like ligand to the carbohydrate recognition domain of galectin-3 in the presence of variable amounts of DMSO. We used isothermal titration calorimetry to characterize binding thermodynamics and ^{15}N NMR relaxation to monitor kinetics. The binding enthalpy is not affected, but we observe a subtle trend of increasingly unfavorable entropy of binding, and consequently decreased affinity, with increasing DMSO concentration. Increasing concentration of DMSO results in a reduced association rate of binding, whereas the dissociation rate is less affected. The observed association rate is inversely proportional to the viscosity of the DMSO-water mixture, as expected from theory, but significantly reduced from the diffusion-controlled limit. By comparing the viscosity dependence of the observed association rate with that of the theoretical diffusion-controlled association rate, we estimate the success rate of productive complex formation following initial encounter of protein and ligand, showing that only 1 out of several hundred binding 'attempts' are successful.

Understanding molecular recognition between proteins and ligands is central to the physical and life sciences and a key aspect of drug design. Ligand binding kinetics has come into focus in the last decade, because of its role in determining the lifetime of the ligand–protein complex, which in turn governs the duration of a biochemical signal or its inhibition.[1–5] Furthermore, the lifetime of the complex, which is equal to the inverse of the dissociation (off-) rate constant, is often found to be a superior predictor of in vivo efficacy compared to the equilibrium binding constant.[5–7] These observations have spawned initiatives to optimize binding kinetics.[5] We have previously investigated the binding kinetics of two series of congeneric ligands designed to inhibit the carbohydrate recognition domain of galectin-3 (Gal3C).[8] Notably, the two ligand series showed different linear free-energy relationships between the off-rate constants and the equilibrium affinity, suggesting that the ligand structure affects the position of the transition state along the generalized reaction coordinate of the binding process.[8] Furthermore, previous studies have indicated that the association (on-) rate depends on the nature of the initial encounter complex.[9]

Gal3C is a member of the galectin family of carbohydrate binding proteins, which have a highly conserved carbohydrate recognition domain. Gal3C is implicated in numerous cellular functions, including cell differentiation, cell cycle regulation, and apoptosis, making it a target for treatment of inflammation and cancer.[10–12] The ligand binding site in Gal3C is located in a shallow and water-exposed groove across a 6-stranded β-sheet, where a number of hydrophilic residues are poised to coordinate ligand oxygen atoms arranged in a sugar-like pattern.[13] The relatively high solubility of natural galectin ligands in aqueous solution results in low affinity and makes it challenging to design high-affinity synthetic ligands. Nonetheless, compounds with nanomolar dissociation constants have been developed towards Gal3C by successfully increasing hydrophobicity while maintaining polar interactions with the canonical ligand-coordinating protein side chains.[14,15] In most cases, the increased hydrophobicity of these ligands reduces their solubility in water.

Indeed, designed organic compounds targeting proteins are commonly poorly soluble in water. For this reason, in vitro binding studies involving such compounds often require mixed solvents to achieve the desired solubility. Dimethyl sulfoxide (DMSO) is one of the most commonly used organic solvents, because it is completely miscible with water and has low chemical reactivity. Ligands can thus be solubilized at high concentration in DMSO and subsequently be added to an aqueous protein solution, e.g., to determine the

binding constant by titration. The resulting protein–ligand solution typically contains a few percent DMSO, and rarely more than ten percent, which is not expected to affect protein structure or stability to any greater extent,[16–18] although contrasting results have been reported in some cases[19] and cell-based screening methods often have a DMSO tolerance of less than 2%.[20] Long-lived interactions between DMSO and proteins that can perturb protein structure and function seem to require suitable binding pockets or clefts,[21] whereas transient interactions with the protein surface are not sufficient in this regard.[22,23] However, the viscosity of DMSO–water mixtures is strongly dependent on the amount of DMSO present, particularly in the dilute regime, where the viscosity increases linearly by a factor of 3 as the volume fraction of DMSO increases to 20%.[24–26] Thus, it is conceivable that DMSO indirectly affects protein function via the change in viscosity. We have previously demonstrated that this effect on viscosity, when uncorrected for, can lead to erroneous conclusions in studies of protein conformational dynamics.[23] Given the importance of characterizing binding kinetics, it is critical to understand how cosolvents might affect the kinetic rate constants.

Here we address the question of whether DMSO might influence the binding affinity and kinetics. As a model system we use Gal3C and a designed, drug-like compound with micromolar affinity and sufficiently high solubility in water that it does not require the addition of DMSO to reach the concentrations used in our in vitro binding studies.[8] We observe subtle variation in binding affinity, which are driven by changes in entropy. Furthermore, the kinetic on-rate for binding varies with DMSO concentration in the manner expected from the viscosity of the DMSO–water mixtures. In contrast, the off-rate for binding shows weaker dependence on viscosity. We take advantage of the linear variation of the on-rate with the inverse of the viscosity to determine the success rate of ligand binding following protein–ligand encounter. We find that each diffusive encounter between ligand and protein has less than 1% chance of forming a productive complex before the encounter complex dissociates.

MATERIALS AND METHODS

Protein expression and purification. Galectin-3C was expressed and purified following published protocols,[27,28] yielding a protein stock solution of 16 mg/ml in buffer consisting of 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, 140 mM NaCl, 2.7 mM KCl, pH 7.3, 2 mM ethylenediaminetetraacetic acid (EDTA), 4 mM tris(2-carboxyethyl)phosphine hydrochloride (TCEP), and 150 mM lactose. The protein stock solution was stored at 278 K.

Ligand synthesis and purification. The ligand orthofluoro-phenyltriazolyl-galactosylthiolglucoside has been described before and shown to be fully water soluble.[8,29]

Isothermal titration calorimetry. Gal3C was prepared by extensive dialysis against 5 mM 4-(2-hydroxyethyl)-1-piperazinethanesulfonic acid (HEPES) buffer pH 7.4 to remove all lactose, followed by centrifugation at 14,000 rpm to remove any aggregates. The protein concentration in each sample was determined using UV absorbance.[28] Ligand and protein solutions at specified DMSO concentrations were prepared using a 50% stock solution of DMSO in 5 mM HEPES buffer to yield samples with 2%, 6% or 10% DMSO and pH 7.4. The protein concentration was 0.74 mM.

ITC experiments were performed on a MicroCal PEAQ–ITC instrument (Malvern Panalytical Ltd) at a temperature of 301 K by titrating the ligand at a concentration of 1.5 mM into the cell containing the protein at a concentration of 150 μM. The DMSO concentrations in the cell and the syringe were carefully matched to minimize the heat of dilution. Three replicate experiments were performed for each condition (DMSO concentration), with an initial injection volume of 0.4 μl followed by 9 injections of 4 μl each using a spinning speed of 750 rpm, reference power of 10 μcal/s, duration of injections of 0.8 s for the first injection and 8 s for the subsequent injections. Each of the triplicate experiments were carried out by concatenating two runs, each with 10 injections, as described above, to generate a single thermogram comprising 18 injections (after subtracting the two initial injections). Additional measurements were carried out for the 0% and 6% DMSO samples in the same manner, but with a factor of two lower injection volume to acquire a greater number of data points in the initial phase of the titrations and thereby improve the definition of the baseline; these runs were performed in duplicate.

Individual thermograms were concatenated and corrected for baseline differences. Peak integration was done using NITPIC.[30] A single-site binding model was fitted simultaneously to the 3 titration curves using SEDPHAT[31] to yield the binding enthalpy ($\Delta H°$), fraction of binding-competent protein (n), and dissociation constant (K_d).

$$\Delta Q_i = Q_i - Q_{i-1} + (V_i/V_0)[Q_i - Q_{i-1}]/2 + Q_{off} \quad (1)$$

where V_i is the volume of the ith injection, V_0 is the cell volume, Q_{off} is an offset parameter that accounts for the heat of mixing, and Q_i is the heat function following the ith injection:

$$Q_i = (\Delta H V_0/2)\left[\alpha - \sqrt{\alpha^2 - 4n\,M_i X_i}\right] \quad (2)$$

where $\alpha = n\,P_i + L_i + K_d$, and P_i and L_i are the total concentrations of the protein and the ligand, respectively, in the cell at any given point of the titration. The free energy and entropy of binding were subsequently determined using the relationships $\Delta G° = RT\ln(K_d)$ and $-T\Delta S° = \Delta G° - \Delta H°$. Although SEDPHAT reports asymmetric error estimates, the present analysis resulted in nearly symmetric errors, which we report as the average of the upper and lower error bounds. Graphical representations of thermograms and isotherms were prepared using GUSSI.[31]

NMR sample preparation. Samples were prepared for NMR with a target saturation of 95% in mind, resulting in samples of 0.4 mM ^{15}N-labelled Gal3C, 0.42 mM of ligand and with 0%, 2%, 6% or 10% DMSO (v/v) in 5 mM HEPES buffer pH 7.5. Thus, the protein concentration in the NMR samples is only a factor of 2.7 greater than that used in the ITC experiments, making the conditions highly similar.

NMR relaxation dispersion experiments. Backbone amide ^{15}N CPMG relaxation dispersion experiments were performed at static magnetic field strengths of 11.7 T, using a Varian/Agilent VNMRS DirectDrive spectrometer equipped with a room-temperature triple-resonance probe, and 14.1 T, using a Bruker Avance NEO spectrometer equipped with a 5 mm HPCN QCI cryo-probe. All experiments were

performed at a temperature of 301 K. Temperature calibration was performed prior to each series of relaxation experiments using a neat protonated methanol sample.[32,33] The sample pH was adjusted immediately before each series of relaxation experiments and checked after each series to make sure the sample pH had not drifted. Constant-time relaxation-compensated CPMG experiments[34,35] were performed at 11.7 T using a 40 ms constant-time relaxation period with CPMG refocusing frequencies ν_{cpmg} of [3×0, 50, 100, 150, 200, 300, 400, 500, 650, 800, 950] Hz and interleaved sampling of t_1 points with different values of ν_{cpmg}. Experiments were acquired with a 2 s recovery delay, 80 scans for each 2D plane, and spectral windows in (t_1, t_2) of (1620, 7023) Hz, sampled over (128, 2210) points. Experiments performed at 14.1 T used ν_{cpmg} = [2×0, 50, 100, 2×300, 400, 500, 600, 700, 800, 900, 1000, 1100] Hz, 2 s recovery delay, 24 scans for each 2D plane, and spectral windows in (t_1, t_2) of (2129, 9615) Hz, sampled over (128, 2306) points.

NMR relaxation data analysis. All spectra were processed using NMRPipe.[36] The processing protocol included squared cosine-bell window functions in both dimensions, a solvent filter, zero-filling to twice the size rounded to the nearest power of two, and polynomial baseline correction in the direct dimension. Linear prediction to twice the number of datapoints was applied in the indirect dimension. Peak volumes were extracted using PINT, which employs line-shape fitting to resolve overlapped peaks.[37,38] Peak intensities were evaluated using a weighted sum of Lorentzian and Gaussian line shapes. The uncertainties of the fitted peak volumes were estimated from the baseplane noise.

The relaxation dispersion data were analyzed using in-house Matlab scripts. Relaxation dispersion curves were fitted to the Carver-Richards two-state exchange model:[39,40]

$$R_{2eff} = R_{20} + R_{ex}(1/\tau) \quad (3)$$

in which

$$R_{ex}(1/\tau) = \frac{1}{2}\left\{k_{ex} - \frac{1}{\tau}cosh^{-1}\left(D_+cosh(\eta_+) - D_-cosh(\eta_-)\right)\right\} \quad (4)$$

$$D_\pm = \frac{1}{2}\{\pm 1 + (\psi + 2\Delta\omega^2)/(\psi^2 + \zeta^2)^{1/2}\} \quad (5)$$

$$\eta_\pm = \frac{\tau}{\sqrt{2}}[\{\pm\psi + (\psi^2 + \zeta^2)^{1/2}\}]^{1/2} \quad (6)$$

and $\psi = k_{ex}^2 - \Delta\omega^2$, $\zeta = -2\Delta\omega\, k_{ex}(1-2p_F)$; $k_{ex} = k_1 + k_{-1}$ is the sum of the forward and reverse rate constants, corresponding to $k_{on}[L]$ and k_{off} in the present case; $\Delta\omega$ is the chemical shift difference between the exchanging free and ligand bound states; R_{20} is the average limiting value of the relaxation rate constant for processes other than chemical exchange; p_F is the population of the (less populated) free state, which is related to the bound state by $p_F = 1 - p_B$; and $\tau = 1/2\,\nu_{cpmg}$ is the spacing between refocusing pulses in the CPMG pulse train.

In fitting the exchange model to the data, we fixed $\Delta\omega$ to the value derived from the chemical shifts measured in spectra of the free and fully saturated states. We performed two separate sets of fits in which p_F was either included as a free parameter of the fit, or fixed at the values calculated from K_d and the total concentrations of protein and ligand in the sample. The statistical significance of each fit was assessed by also fitting the data to a constant R_{20} value (i.e. modelling a flat dispersion profile, indicating the absence of exchange), and the *F*-test was used to discriminate between models by rejecting the simpler model at the level $p < 0.001$. Errors in the fitted parameters were estimated from 1000 synthetic data sets created using Monte-Carlo simulations.[9]

THEORY

The diffusion controlled on-rate constant describing the encounter of two spherical molecules, ligand (*L*) and protein (*P*), is given by:[41,42]

$$k_{on,D} = 4\pi(R_L + R_P)(D_L + D_P) \quad (7)$$

where $R_L + R_P$ is the contact distance between the centers of the two spheres (sum of the radii of *L* and *P*), and $D_L + D_P$ is the relative diffusion coefficient of the ligand–protein pair. The diffusion coefficient is given by the Stokes-Einstein equation, $D_i = k_B T/(6\pi R_i \eta)$, where η is the solvent viscosity. Thus, $k_{on,D}$ can be expressed as:

$$k_{on,D} = \frac{2k_BT}{3\eta}\left(2 + \frac{R_L}{R_P} + \frac{R_P}{R_L}\right) \quad (8)$$

For the system studied here, the radii are approximately R_L = 6 Å and R_P = 16 Å, yielding an approximate value of $k_{on,D} \approx 3.36\, k_BT/\eta$.

In reality, each encounter does not lead to productive binding, because the encounter complex can dissociate before the ligand has had time to diffuse across the protein surface into the binding site. Thus, the binding process involves a pre-equilibrium of the encounter complex and a diffusive search over the protein surface from the point of first contact to the binding site:[41]

$$L + P \underset{k_{dis}}{\overset{k_{on,D}}{\rightleftarrows}} L*P \xrightarrow{k_s} LP$$

where *L*P* and *LP* denote the encounter complex and final complex, respectively, k_{dis} is the rate constant for dissociation of the encounter complex, and k_s is the rate constant for the search process. The effective on-rate constant is then given by:

$$k_{on} = \frac{k_{on,D}k_s}{k_{dis}+k_s} \quad (9)$$

Assuming that viscosity mainly affects $k_{on,D}$ and much less so k_{dis} and k_s, the slope of k_{on} versus $1/\eta$ can be compared with the slope expected for a diffusion controlled reaction, i.e., the slope of $k_{on,D}$ versus $1/\eta$ as given in eq. (8), to yield the ratio $k_s/k_{dis} = \rho$, which provides an estimate of the 'success'-rate of productive complex formation in the binding site following the initial encounter between protein and ligand:

$$\frac{k_{on}}{k_{on,D}} = \frac{k_s}{k_{dis}+k_s} = \frac{\rho}{1+\rho} \quad (10)$$

so that $\rho = (k_{on}/k_{on,D})/(1-k_{on}/k_{on,D})$. Thus, if $k_{on} << k_{on,D}$, $\rho \approx (k_{on}/k_{on,D})$, which is shown to be valid in the present case (see Results).

RESULTS AND DISCUSSION

We performed ITC experiments and NMR relaxation dispersion experiments at four different sample conditions:

without DMSO and with 3 different concentrations of DMSO, 2% (v/v), 6%, and 10%.

ITC experiments reveal subtle effects of DMSO on binding affinity. We used ITC to determine the thermodynamic fingerprint of ligand binding to Gal3C (Fig. 1; Table 1; Fig. S1). The enthalpy of binding does not depend on the DMSO concentration, in agreement with previous work showing that DMSO does not interact with the ligand binding site of Gal3C.[23] In contrast, there is a modest trend towards increasingly unfavorable entropy of binding, resulting in a slight increase in K_d, with increasing DMSO concentration. Given that DMSO does not inhibit the binding site, the observed effect of DMSO on the binding affinity must therefore arise from indirect effects.

Table 1. Thermodynamics of ligand binding

c(DMSO) (v/v %)	$-\Delta H°$ (kJ/mol)	$-T\Delta S°$ (kJ/mol)	$-\Delta G°$ (kJ/mol)	K_d (10^{-6} M)
0%	49.1 ± 2.1	18.8 ± 2.2	30.3 ± 0.6	5.6 ± 1.3
2%	49.8 ± 1.0	19.8 ± 1.1	29.9 ± 0.4	6.5 ± 0.9
6%	49.8 ± 1.1	20.2 ± 1.1	29.6 ± 0.2	7.3 ± 0.7
10%	49.4 ± 3.8	21.3 ± 4.0	28.2 ± 1.0	13 ± 5

NMR relaxation dispersion experiments reveal effects of DMSO on binding kinetics. We used NMR spectroscopy to investigate how the binding kinetics vary with DMSO concentration. We performed ^{15}N CMPG relaxation dispersion experiments at each DMSO concentration on samples that were very similar to those resulting from the ITC measurements: the protein concentration in the NMR sample was a factor of 2.7 higher than in the ITC experiments and the protein was saturated with ligand to approximately 95%. Under these conditions, the system is undergoing equilibrium exchange between the free (with a relative population of p_F = 0.05) and bound (p_B = 0.95) states, which gives rise to an additional contribution to the transverse relaxation rates for those protein residues that experience different chemical shifts in the two states, see eq (3).

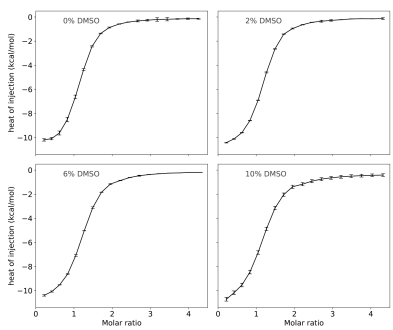


Figure 1. Representative isotherms from ITC measurements of ligand binding to Gal3C at four DMSO concentrations: 0%, 2%, 6%, and 10% v/v. Figure S1 shows all binding isotherms.

The resulting relaxation data show significant dispersion in all 4 samples for 3 residues: I145, L147 and E185, which are all located in or close to the binding site (Fig. 2a). Several other residues also show relaxation dispersions, but are partially overlapped or very broadened in one or several of the samples. Figure 2 shows relaxation dispersion data for I145 and E185, the two residues that exhibit the largest dispersion step (i.e., the largest chemical shift difference between the free and bound states). In fitting the two-state exchange model to the data, we fixed the chemical shift difference, $\Delta\omega$, to the value calculated from the peak positions in HSQC spectra of the free protein and fully saturated (p_B > 99%) protein. The population of the bound state, p_B, can be calculated from K_d and the reactant concentrations in the NMR sample. However, in order to allow for minor systematic errors in these values and variation between samples, we did not only perform fits with fixed populations, but also included fits with the populations included as free parameters. We performed a global fit of the exchange data for all 3 residues at each DMSO concentration; however, the fits are largely governed by I145 and E185 due to the superior data quality obtained for these residues. The gradual increase in R_{20}, i.e., the limiting value reached for high ν_{cpmg} (Fig. 2), with increasing DMSO concentration is fully explained by the increase in the correlation time for rotational diffusion.[23]

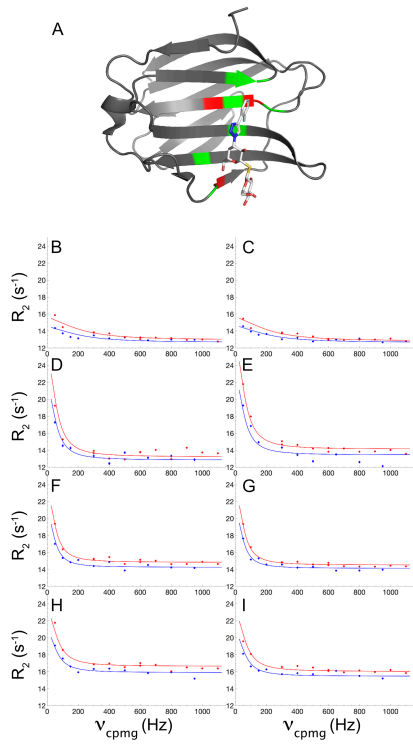


Figure 2. Ligand binding kinetics measured by NMR relaxation dispersion. (A) Crystal structure of the Gal3C–ligand complex with the protein and ligand shown in ribbon and stick representation, respectively. Protein residues showing significant

relaxation dispersion at all 4 DMSO concentrations are colored red: I145, L147 and E185. Protein residues with backbone atoms within 5 Å from the ligand are colored green. The ligand atoms are colored white (carbon), red (oxygen), blue (nitrogen), yellow (sulfur), pale blue (fluorine). (B–I): ^{15}N CPMG relaxation dispersion profiles of residues I145 (B, D, F, H) and E185 (C, E, G, I) measured in 0% DMSO (B, C), 2% DMSO (D, E), 6% DMSO (F, G), and 10% DMSO (H, I). CPMG relaxation dispersions were acquired at static magnetic field strengths of 11.7 T (blue) and 14.1 T (red). Panel A was prepared using PDB-ID: 6RZF[43] and the Pymol software package.[44]

The determined exchange parameters are listed in Table 2. The fits including the relative population of the bound state p_B as a free parameter yield values that vary between 0.90–0.94 in the different samples, in good agreement with the target value of 0.95 calculated from K_d and the reactant concentrations in the NMR sample. Based on the fitted parameters k_{ex} and p_B, we obtain $k_{off} = k_{ex}/(1-p_B)$, and then calculate $k_{on} = k_{off}/K_d$, where K_d is taken from the ITC measurements; given that the sample conditions are very nearly the same in the ITC and NMR experiments, we are confident these value sof K_d are valid.

Table 2. Ligand binding exchange parameters

	c(DMSO) (v/v %)			
Parameters (p_B free)	0%	2%	6%	10%
k_{ex} (s^{-1})	1420 ± 6	353 ± 3	276 ± 4	379 ± 2
p_B [a]	0.93	0.90	0.93	0.94
k_{off} (s^{-1}) [b]	96.6 ± 0.6	36.7 ± 0.4	18.5 ± 0.3	23.0 ± 0.1
k_{on} (10^6 M^{-1} s^{-1}) [c]	17 ± 4	5.6 ±0.8	2.5 ±0.3	1.8 ±0.7
Parameters (p_B fixed)				
k_{ex} (s^{-1})	855 ± 2	375 ± 7	353 ± 5	377 ± 2
p_B	0.95	0.95	0.95	0.95
k_{off} (s^{-1}) [b]	38.5 ± 0.1	18.7 ± 0.3	17.6 ± 0.2	18.8 ± 0.1
k_{on} (10^6 M^{-1} s^{-1}) [c]	7 ± 2	2.9 ± 0.4	2.4 ± 0.2	1.5 ± 0.6

[a] errors in p_B are in the third decimal in all cases; [b] calculated from k_{ex} and p_F, $k_{off} = k_{ex}/p_F$; [c] calculated from k_{off} and K_d.

The two different approaches of fitting the relaxation dispersion data, viz. using p_B as a free fitting parameter or using a fixed value of p_B, produce slightly different exchange rates (Table 2). Naturally, in those cases where the fitted value of p_B is similar to the calculated value the two approaches yield similar results. The extracted value of k_{off} does not differ by more than a factor of 2.5 (at 0% DMSO). Similarly, k_{on} is broadly similar and differs at most by a factor of 2.4 (again at 0% DMSO) and is identical at the two highest DMSO concentrations. Moreover, k_{on} follows the same trend in both cases (Figure 3), decreasing monotonously with increasing DMSO concentration. In fact, plotting k_{on} versus $1/\eta$ reveals a linear relationship as expected from eqs. (7–9), see Figure 3. The linear fit yields a Pearson's correlation coefficient of 0.98 and p-value of 0.02, indicating that the likelihood of obtaining this level of correlation by chance is low.

The slope of k_{on} versus $1/\eta$ is 0.061 Pa M^{-1} for the fit using fixed p_B and 0.020 Pa M^{-1} for the free fit, which can be compared with the value expected for a diffusion limited on-rate constant, $k_{on,D}$ = 8.3 Pa M^{-1}, as determined from eq (8). Thus, a significant reduction in the slope is observed compared to the diffusion-controlled case, despite the variation among the results obtained from the two different fitting protocols.

The ratio $k_{on}/k_{on,D}$ provides an estimate of $\rho = k_s/k_{dis}$, see eq. (10), which is a measure of the success rate of complex formation. The range of values determined here yields a success rate in the range $\rho \approx$ 0.2% to 0.7%. In other words, 135–415 transient encounters between ligand and protein occur on average for each successful binding event.

While it is generally difficult to calculate the effective on-rate constant from first principles, it is arguably straightforward to estimate the diffusion controlled on-rate describing the first encounter between ligand and protein. The simplified treatment used here neglects any effects of molecular shape by assuming that both ligand and protein can be treated as spheres, eqs. (7–8). Nonetheless, we believe that our approach provides a reasonable estimate of the ligand binding 'success rate', showing that several hundred binding attempts are required for each successful complex formation of Gal3C with the ligand. It should be noted that in the present case the ligand is highly water soluble. Presumably, a more hydrophobic ligand might show less tendency to dissociate from the protein surface and therefore exhibit a greater 'success rate' than what is observed here. Similarly, electrostatic interactions between ligand and protein—which do not play any significant role in the present case, since the ligand is uncharged—should result in an improved success rate.[45] We expect that the ligand-binding success rate is highly system dependent.

Concluding remarks. We conclude that the addition of up to 10% DMSO does not affect the binding thermodynamics to any appreciable extent in the studied model system. While there is a subtle trend toward decreasing binding

affinity (i.e., increasing K_d) with increasing DMSO concentration, the effect is within a factor of 2, which translates into a few kJ/mol in standard free energy. The subtle change in standard free energy of binding originates entirely from an entropic effect, whereas the enthalpy of binding does not vary with DMSO concentration, which is expected since DMSO does not inhibit the binding site of galectin-3.[23] However, it is entirely possible that DMSO might act as a competitive inhibitor in other systems.

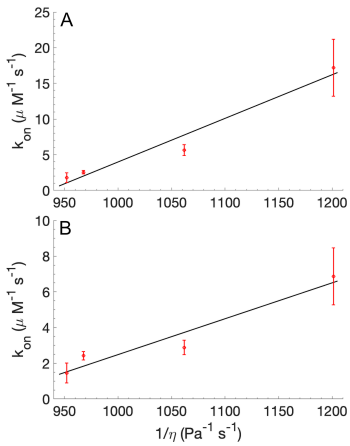


Figure 3. Effective on-rate constants (k_{on}; red symbols) plotted against the inverse dynamic viscosity ($1/\eta$) of the DMSO-water mixture. The black line shows the fitted linear regression model. k_{on} was determined from relaxation dispersion fits including p_B as a free parameter (A), or a fixed parameter (B). Error bars indicate ±1 standard deviation.

Our study reveals a statistically significant effect on the effective on-rate constant of binding, which decreases with increasing DMSO concentration. This effect is expected from the increase in viscosity with increasing DMSO concentration, which acts to decrease the diffusion coefficient. The slope of k_{on} versus $1/\eta$ is significantly reduced from the value expected for a diffusion controlled reaction, indicating that several hundred protein–ligand encounters occur for each productive binding event. In other words, each diffusive encounter between ligand and protein has less than 1% chance of resulting in a productive complex before the encounter complex dissociates. We anticipate that the approach introduced here should be widely applicable to studies of protein-ligand binding kinetics aiming to unravel ligand binding mechanisms and should serve as a useful complement to methods for characterizing encounter complexes.[46,47]

ASSOCIATED CONTENT

Supporting Information. Figures showing all ITC binding isotherms. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

* mikael.akke@bpc.lu.se

Author Contributions

‡These authors contributed equally.

ACKNOWLEDGMENT

This work was supported by the Swedish Research Council (2018-4995) and the Knut and Alice Wallenberg Foundation (2013.022). We thank Kristoffer Peterson and Ulf J. Nilsson for providing the ligand, Olof Stenström and Johan Wallerstein for advice on ITC measurements, and Kristofer Modig for helpful discussions.

ABBREVIATIONS

CPMG, Carr-Purcell-Meiboom-Gill; DMSO, dimethylsulfoxide; Gal3C, the carbohydrate binding domain of galectin-3; ITC, isothermal titration calorimetry.

REFERENCES

(1) Tummino, P. J.; Copeland, R. A. Residence Time of Receptor - Ligand Complexes and Its Effect on Biological Function. *Biochemistry* **2008**, *47*, 5481–5492. https://doi.org/10.1021/bi8002023.

(2) Tonge, P. J. Drug-Target Kinetics in Drug Discovery. *ACS Chemical Neuroscience* **2018**, *9*, 29–39. https://doi.org/10.1021/acschemneuro.7b00185.

(3) Bernetti, M.; Masetti, M.; Rocchia, W.; Cavalli, A. Kinetics of Drug Binding and Residence Time. *Annual Review of Physical Chemistry* **2019**, *70*, 143–171. https://doi.org/10.1146/annurev-physchem-042018-052340.

(4) Nunes-Alves, A.; Kokh, D. B.; Wade, R. C. Recent Progress in Molecular Simulation Methods for Drug Binding Kinetics. *Current Opinion in Structural Biology* **2020**, *64*, 126–133. https://doi.org/10.1016/j.sbi.2020.06.022.

(5) Pan, A. C.; Borhani, D. W.; Dror, R. O.; Shaw, D. E. Molecular Determinants of Drug-Receptor Binding Kinetics. *Drug Discovery Today* **2013**, *18*, 667–673. https://doi.org/10.1016/j.drudis.2013.02.007.

(6) Copeland, R. A. The Drug-Target Residence Time Model: A 10-Year Retrospective. *Nat. Rev. Drug Discov.* **2016**, *15*, 87–95.

(7) Yin, N.; Pei, J.; Lai, L. A Comprehensive Analysis of the in Uence of Drug Binding Kinetics on Drug Action at Molecular and Systems Levels. *Mol. Biosyst.* **2013**, *9*, 1381–1389.

(8) Stenström, O.; Diehl, C.; Modig, K.; Nilsson, U. J.; Akke, M. Mapping the Energy Landscape of Protein–Ligand Binding via Linear Free Energy Relationships Determined by Protein NMR Relaxation Dispersion. *RSC Chemical Biology* **2021**, *2*, 259–265. https://doi.org/10.1039/d0cb00229a.

(9) Gabdoulline, R. R.; Wade, R. C. On the Protein-Protein Diffusional Encounter Complex. *Journal of Molecular Recognition* **1999**, *12* (4), 226–234. https://doi.org/10.1002/(SICI)1099-1352(199907/08)12:4<226::AID-JMR462>3.0.CO;2-P.

(10) Capone, E.; Iacobelli, S.; Sala, G. Role of Galectin 3 Binding Protein in Cancer Progression: A Potential Novel Therapeutic Target. *J Transl Med* **2021**, *19*, 405.

(11) Hara, A.; Niwa, M.; Noguchi, K.; Kanayama, T.; Niwa, A.; Matsuo, M.; Hatano, Y.; Tomita, H. Galectin-3 as a Next-Generation Biomarker for Detecting Early Stage of Various Diseases. *Biomolecules* **2020**, *10*, 389.

(12) Liu, F.-T.; Rabinovich, G. A. Galectins as Modulators of Tumour Progression. *Nature Reviews Cancer* **2005**, *5*, 29–41. https://doi.org/10.1038/nrc1527.

(13) Saraboji, K.; Håkansson, M.; Genheden, S.; Diehl, C.; Qvist, J.; Weininger, U.; Nilsson, U. J.; Leffler, H.; Ryde, U.; Akke, M.; Logan, D. T. The Carbohydrate-Binding Site in Galectin-3 Is Preorganized to Recognize a Sugar-like Framework of Oxygens: Ultra-High Resolution Structures and Water Dynamics. *Biochemistry* **2012**, *51*, 296–306.

(14) Delaine, T.; Collins, P.; MacKinnon, A.; Sharma, G.; Stegmayr, J.; Rajput, V. K.; Mandal, S.; Cumpstey, I.; Larumbe, A.; Salameh, B. A.; Kahl-Knutsson, B.; van Hattum, H.; van Scherpenzeel, M.; Pieters, R. J.; Sethi, T.; Schambye, H.; Oredsson, S.; Leffler, H.; Blanchard, H.; Nilsson, U. J. Galectin-3-Binding Glycomimetics That Strongly Reduce Bleomycin-Induced Lung Fibrosis and Modulate Intracellular Glycan Recognition. *ChemBioChem* **2016**, *17*, 1759–1770. https://doi.org/10.1002/cbic.201600285.

(15) Zetterberg, F. R.; Peterson, K.; Johnsson, R. E.; Brimert, T.; Håkansson, M.; Logan, D. T.; Leffler, H.; Nilsson, U. J. Monosaccharide Derivatives with Low-Nanomolar Lectin Affinity and High Selectivity Based on Combined Fluorine–Amide, Phenyl–Arginine, Sulfur–π, and Halogen Bond Interactions. *ChemMedChem* **2018**, *13*, 133–137. https://doi.org/10.1002/cmdc.201700744.

(16) Arakawa, T.; Kita, Y.; Timasheff, S. N. Protein Precipitation and Denaturation by Dimethyl Sulfoxide. *Biophysical Chemistry* **2007**, *131* (1–3), 62–70. https://doi.org/10.1016/j.bpc.2007.09.004.

(17) Voets, I. K.; Cruz, W. A.; Moitzi, C.; Lindner, P.; Areas, E. P. G.; Schurtenberger, P. DMSO-Induced Denaturation of Hen Egg White Lysozyme. *Journal of Physical Chemistry B* **2010**, *114* (36), 11875–11883. https://doi.org/10.1021/jp103515b.

(18) Batista, A. N. L.; Batista, J. M.; Bolzani, V. S.; Furlan, M.; Blanch, E. W. Selective DMSO-Induced Conformational Changes in Proteins from Raman Optical Activity. *Physical Chemistry Chemical Physics* **2013**, *15* (46), 20147–20152. https://doi.org/10.1039/c3cp53525h.

(19) Tjernberg, A.; Markova, N.; Griffiths, W. J.; Hallen, D. DMSO-Related Effects in Protein Characterization. *Journal of Biomolecular Screening* **2006**, *11* (2), 131–137. https://doi.org/10.1177/1087057105284218.

(20) Johnston, P. A.; Johnston, P. A. Cellular Platforms for HTS: Three Case Studies. *Drug Discovery Today* **2002**, *7* (6), 353–363. https://doi.org/10.1016/S1359-6446(01)02140-7.

(21) Papaneophytou, C. P.; Grigoroudis, A. I.; McInnes, C.; Kontopidis, G. Quantification of the Effects of Ionic Strength, Viscosity, and Hydrophobicity on Protein–Ligand Binding Affinity. *ACS Medicinal Chemistry Letters* **2014**, *5* (8), 931–936. https://doi.org/10.1021/ml500204e.

(22) Johannesson, H.; Denisov, V. P.; Halle, B. Dimethyl Sulfoxide Binding to Globular Proteins. A Nuclear Magnetic Relaxation Dispersion Study. *Protein Science* **1997**, *6*, 1756–1763.

(23) Wallerstein, J.; M, A. Minute Additions of DMSO Affect Protein Dynamics Measurements by NMR Relaxation Experiments through Significant Changes in Solvent. *ChemPhysChem* **2019**, *20*, 326–332.

(24) Cowie, J. M. G.; Toporowski, P. M. Association in the Binary Liquid System Dimethyl Sulphoxide - Water. *Canadian Journal of Chemistry* **1961**, *39*, 2240–2243. https://doi.org/10.1139/v61-296.

(25) Schichman, S. A.; Amey, R. L. Viscosity and Local Liquid Structure in Dimethyl Sulfoxide-Water Mixtures. *Journal of Physical Chemistry* **1971**, *75*, 98–102. https://doi.org/10.1021/j100671a017.

(26) Catalan, J.; Diaz, C.; Garcia-Blanco, F. Characterization of Binary Solvent Mixtures of DMSO with Water and Other Cosolvents. *Journal of Organic Chemistry* **2001**, *66*, 5846–5852. https://doi.org/10.1021/jo010415i.

(27) Diehl, C.; Engström, O.; Delaine, T.; Håkansson, M.; Genheden, S.; Modig, K.; Leffler, H.; Ryde, U.; Nilsson, U. J.; Akke, M. Protein Flexibility and Conformational Entropy in Ligand Design Targeting the Carbohydrate Recognition Domain of Galectin-3. *J Am Chem Soc* **2010**, *132*, 14577–14589.

(28) Diehl, C.; Genheden, S.; Modig, K.; Ryde, U.; Akke, M. Conformational Entropy Changes upon Lactose Binding to the Carbohydrate Recognition Domain of Galectin-3. *Journal of Biomolecular NMR* **2009**, *45*, 157–169. https://doi.org/10.1007/s10858-009-9356-5.

(29) Peterson, K.; Kumar, R.; Stenström, O.; Verma, P.; Verma, P. R.; Håkansson, M.; Kahl-Knutsson, B.; Zetterberg, F.; Leffler, H.; Akke, M.; Logan, D. T.; Nilsson, U. J. Systematic Tuning of Fluoro-Galectin-3 Interactions Provides Thiodigalactoside Derivatives with Single-Digit NM Affinity and High Selectivity. *Journal of Medicinal Chemistry* **2018**, *61*, 1164–1175. https://doi.org/10.1021/acs.jmedchem.7b01626.

(30) Keller, S.; Vargas, C.; Zhao, H.; Piszczek, G.; Brautigam, C. A.; Schuck, P. High-Precision Isothermal Titration Calorimetry with Automated Peak-Shape Analysis. *Analytical Chemistry* **2012**, *84*, 5066–5073. https://doi.org/10.1021/ac3007522.

(31) Brautigam, C. A.; Zhao, H.; Vargas, C.; Keller, S.; Schuck, P. Integration and Global Analysis of Isothermal Titration Calorimetry Data for Studying Macromolecular Interactions. *Nature Protocols* **2016**, *11*, 882–894. https://doi.org/10.1038/nprot.2016.044.

(32) Van Geet, A. L. Calibration of Methanol Nuclear Magnetic Resonance Thermometer at Low Temperature. *Analytical Chemistry* **1970**, *42* (6), 679–680. https://doi.org/10.1021/ac60288a022.

(33) Raiford, D. S.; Fisk, C. L.; Becker, E. D. Calibration of Methanol and Ethylene Glycol Nuclear Magnetic Resonance Thermometers. *Analytical Chemistry* **1979**, *51* (12), 2050–2051. https://doi.org/10.1021/ac50048a040.

(34) Loria, J. P.; Rance, M.; Palmer, A. G. A Relaxation-Compensated Carr-Purcell-Meiboom-Gill Sequence for Characterizing Chemical Exchange by NMR Spectroscopy. *J Am Chem Soc* **1999**, *121*, 2331–2332.

(35) Mulder, F. A. A.; Skrynnikov, N. R.; Hon, B.; Dahlquist, F. W.; Kay, L. E. Measurement of Slow (Microseconds-Milliseconds) Time Scale Dynamics in Protein Side Chains by 15N Relaxation Dispersion NMR Spectroscopy: Application to Asn and Gln Residues in a Cavity Mutant of T4 Lysozyme. *J Am Chem Soc* **2001**, *123*, 967–975.

(36) Delaglio, F.; Grzesiek, S.; Vuister, G. W.; Zhu, G.; Pfeifer, J.; Bax, A. NMRPipe: A Multidimensional Spectral Processing System Based on UNIX Pipes. *Journal of Biomolecular NMR* **1995**, *6*, 277–293.

(37) Ahlner, A.; Carlsson, M.; Jonsson, B. H.; Lundström, P. PINT: A Software for Integration of Peak Volumes and Extraction of Relaxation-rates. *Journal of Biomolecular NMR* **2013**, *56*, 191–202. https://doi.org/10.1007/s10858-013-9737-7.

(38) Niklasson, M.; Otten, R.; Ahlner, A.; Andresen, C.; Schlagnitweit, J.; Petzold, K.; Lundström, P. Comprehensive Analysis of NMR Data Using Advanced Line Shape Fitting. *Journal of Biomolecular NMR* **2017**, *69*, 93–99. https://doi.org/10.1007/s10858-017-0141-6.

(39) Carver, J. P.; Richards, R. E. A General Two-Site Solution for the Chemical Exchange Produced Dependence of T2 upon the Carr-Purcell Pulse Separation. *Journal of Magnetic Resonance* **1972**, *6*, 89–105.

(40) Davis, D. G.; Perlman, M. E.; London, R. E. Direct Measurements of the Dissociation-Rate Constant for Inhibitor-Enzyme Complexes via the T1r and T2 (CPMG) Methods. *Journal of Magnetic Resonance, Series B* **1994**, *104*, 266–275.

(41) Schreiber, G.; Haran, G.; Zhou, H. X. Fundamental Aspects of Protein - Protein Association Kinetics. *Chemical Reviews* **2009**, *109*, 839–860. https://doi.org/10.1021/cr800373w.

(42) Smoluchowski, M. v. Versuch Einer Mathematischen Theorie Der Koagulationskinetik Kolloider Lösungen. *Z. Phys. Chem.* **1917**, *92*, 129–168.

(43) Wallerstein, J.; Ekberg, V.; Misini Ignjatović, M.; Kumar, R.; Caldararu, O.; Peterson, K.; Leffler, H.; Logan, D. T.; Nilsson, U. J.; Ryde, U.; Akke, M. Entropy–Entropy Compensation Between the Conformational and Solvent Degrees of Freedom Fine-Tunes Affinity in Ligand Binding to Galectin-3C. *JACS Au* **2021**, *1*, 484–500. https://doi.org/10.1021/jacsau.0c00094.

(44) *The PyMOL Molecular Graphics System*; Schrödinger, LLC.

(45) Schreiber, G.; Haran, G.; Zhou, H. X. Fundamental Aspects of Protein - Protein Association Kinetics. *Chemical Reviews* **2009**, *109*, 839–860. https://doi.org/10.1021/cr800373w.

(46) Clore, G. M.; Iwahara, J. Theory, Practice, and Applications of Paramagnetic Relaxation Enhancement for the Characterization of Transient Low-Population States of Biological Macromolecules and Their Complexes. *Chemical Reviews* **2009**, *109* (9), 4108–4139.

(47) Ubbink, M. The Courtship of Proteins: Understanding the Encounter Complex. *FEBS Letters* **2009**, *583*, 1060–1066.

Paper IV

Magn. Reson., 2, 571–587, 2021
https://doi.org/10.5194/mr-2-571-2021

Open Access
MAGNETIC RESONANCE

Rapid measurement of heteronuclear transverse relaxation rates using non-uniformly sampled $R_{1\rho}$ accordion experiments

Sven Wernersson[1], Göran Carlström[2], Andreas Jakobsson[3], and Mikael Akke[1]

[1]Biophysical Chemistry, Center for Molecular Protein Science, Department of Chemistry, Lund University, P.O. Box 124, 22100 Lund, Sweden
[2]Centre for Analysis and Synthesis, Department of Chemistry, Lund University, P.O. Box 124, 22100 Lund, Sweden
[3]Department of Mathematical Statistics, Lund University, P.O. Box 118, 22100 Lund, Sweden

Correspondence: Mikael Akke (mikael.akke@bpc.lu.se)

Received: 12 May 2021 – Discussion started: 20 May 2021
Revised: 25 June 2021 – Accepted: 28 June 2021 – Published: 12 July 2021

Abstract. Multidimensional, heteronuclear NMR relaxation methods are used extensively to characterize the dynamics of biological macromolecules. Acquisition of relaxation datasets on proteins typically requires significant measurement time, often several days. Accordion spectroscopy offers a powerful means to shorten relaxation rate measurements by encoding the "relaxation dimension" into the indirect evolution period in multidimensional experiments. Time savings can also be achieved by non-uniform sampling (NUS) of multidimensional NMR data, which is used increasingly to improve spectral resolution or increase sensitivity per unit time. However, NUS is not commonly implemented in relaxation experiments, because most reconstruction algorithms are inherently nonlinear, leading to problems when estimating signal intensities, relaxation rate constants and their error bounds. We have previously shown how to avoid these shortcomings by combining accordion spectroscopy with NUS, followed by data reconstruction using sparse exponential mode analysis, thereby achieving a dramatic decrease in the total length of longitudinal relaxation experiments. Here, we present the corresponding transverse relaxation experiment, taking into account the special considerations required for its successful implementation in the framework of the accordion-NUS approach. We attain the highest possible precision in the relaxation rate constants by optimizing the NUS scheme with respect to the Cramér–Rao lower bound of the variance of the estimated parameter, given the total number of sampling points and the spectrum-specific signal characteristics. The resulting accordion-NUS $R_{1\rho}$ relaxation experiment achieves comparable precision in the parameter estimates compared to conventional CPMG (Carr–Purcell–Meiboom–Gill) R_2 or spin-lock $R_{1\rho}$ experiments while saving an order of magnitude in experiment time.

1 Introduction

NMR relaxation offers a powerful means to study the dynamics of proteins and other biological macromolecules (Alderson and Kay, 2020; Mittermaier and Kay, 2009; Palmer, 2004). Most commonly, relaxation experiments on proteins are acquired as a series of two-dimensional (2D) spectra, in order to resolve as many resonances as possible, wherein relaxation rates are measured via their effect on the resonance intensities in a "third dimension" obtained by parametrically varying the length of a relaxation time period or the refocusing frequency of an applied radio-frequency field, or both. Thus, relaxation experiments often involve significant time requirements and may take up to several days. An ingenious alternative to these lengthy experiments is offered by the accordion approach originally developed by Bodenhausen and Ernst (Bodenhausen and Ernst, 1981, 1982) to study chemical exchange. In accordion spectroscopy, the third dimen-

sion is incremented synchronously with the second (indirect) dimension, with the result that the relaxation decay is encoded into the interferogram of the indirect evolution period. Consequently, the total experiment time is reduced significantly. More recent implementations include the constant-time accordion experiment (Carr et al., 1998; Mandel and Palmer, 1994), from which relaxation rate constants can be extracted using either time-domain analysis of the interferogram (Mandel and Palmer, 1994) or line-shape analysis of the Fourier transformed data (Chen and Tjandra, 2009; Harden and Frueh, 2014; Rabier et al., 2001).

Non-uniform sampling (NUS) of the indirect dimensions of multidimensional NMR data can greatly shorten the total experiment time (Gołowicz et al., 2020; Mobli and Hoch, 2014) and has become commonplace in the last decade. However, most spectral reconstruction algorithms suffer from nonlinearity of signal intensities, which limits "plug-and-play" use of NUS in quantitative experiments and requires careful consideration of both sampling schemes and data modeling to produce consistent results and reliable error estimates (East et al., 2021; Linnet and Teilum, 2016; Mayzel et al., 2017; Stetz and Wand, 2016; Urbańczyk et al., 2017). We recently introduced an approach that avoids these problems by combining accordion spectroscopy with NUS (Carlström et al., 2019) and analyzing the resulting data using DSURE (damped super-resolution estimator), a sparse reconstruction technique enabling maximum-likelihood estimation of the time-domain signal parameters from NUS data (Juhlin et al., 2018; Swärd et al., 2016). We stress the point that accordion spectroscopy encodes the desired relaxation rate constants in the interferogram of the multidimensional dataset; hence, the analysis does not rely on measuring intensities in multiple NUS datasets. Moreover, maximum likelihood estimation of model parameters makes it straightforward to derive reliable error bounds. Our approach leads to accumulated time savings through both the accordion and NUS methods. Compared to a conventional relaxation experiment, accordion reduces the experiment time by a factor of $M/2$, where M is the number of datasets included in the conventional approach, and NUS reduces the experiment time by a factor of N_{full}/N, where N_{full} is the number of data points sampled in the indirect dimension of the conventional experiment, and N is the number of points in the NUS scheme. We previously demonstrated this approach by measuring longitudinal relaxation rate constants (R_1) in proteins with time savings of up to a factor of 20 (Carlström et al., 2019). For example, using this approach we have successfully measured R_1 on protein samples with 10-fold lower concentration than normally used (Verteramo et al., 2021).

A number of considerations are of general importance when choosing the detailed sampling scheme for NUS, including the need to keep the total number of increments small in order to speed up data acquisition and to sample short t_1 values to optimize sensitivity and long t_1 values to optimize spectral resolution (Hyberts et al., 2014; Mobli and Hoch, 2014). Various NUS schemes have been developed to accommodate these different requirements, including the popular Poisson-gap scheme (Hyberts et al., 2010). However, in the context of relaxation experiments, the most important aspect is to achieve high precision in the estimated relaxation rate constants. To this end, we have previously developed a method to optimize the sampling scheme with respect to the Cramér–Rao lower bound (CRLB), which yields a lower bound on the achievable variance of the parameters, given the actual spectrum characteristics (i.e., the total number of component signals and their resonance frequencies, line widths, and intensities) and the number of sampling points (Carlström et al., 2019; Månsson et al., 2014; Swärd et al., 2018); similar implementations have followed (Jameson et al., 2019; Waudby et al., 2021).

Here, we introduce accordion-NUS $R_{1\rho}$ pulse sequences that complement the previously presented R_1 experiment (Carlström et al., 2019). The $R_{1\rho}$ relaxation experiment can be implemented using either on-resonance or off-resonance spin-lock fields, making it suitable for measurement of R_2 relaxation rate constants to characterize fast dynamics, as well as conformational/chemical exchange processes across a wide range of timescales (Akke and Palmer, 1996). The present paper addresses several issues concerning measurement of transverse relaxation rates in accordion mode combined with NUS. We validate the accordion-NUS method by extensive comparisons with data acquired using uniformly sampled accordion experiments, as well as conventional non-accordion experiments of both the $R_{1\rho}$ and CPMG (Carr–Purcell–Meiboom–Gill) R_2 types. In addition, exchange contributions to the transverse relaxation rates were characterized using CPMG relaxation dispersion experiments. Our results show that the accordion-NUS $R_{1\rho}$ experiment enables measurement of accurate R_2 relaxation rate constants with a relative uncertainty of only 2 %–3 % using a sampling density of 50 % in the indirect dimension. Lower sampling densities lead to progressively reduced precision, with 5 % relative uncertainty being obtained using less than 20 % sampling density.

2 Materials and methods

2.1 Constant-time accordion relaxation methodology

For completeness, here we briefly outline the salient features of the constant-time accordion method (Mandel and Palmer, 1994). Figure 1 shows the accordion $R_{1\rho}$ pulse sequence. The total relaxation delay is $T_\kappa = n \cdot 4 \cdot \tau$, where n is the sampled point number in t_1, $\tau = (\kappa \cdot \Delta t_1)/4$, Δt_1 is the dwell time, and κ is the accordion scaling factor. The total constant-time delay $T = \tau_1 + \tau_2 + \tau_3$, where $\tau_1 = (T - t_1)/2$, $\tau_2 = T/2 - \Delta$, $\tau_3 = \Delta + t_1/2$, $\Delta = 1/(4J)$, and J is the ^{1}H-^{15}N ^1J-coupling constant (~ 92 Hz). The constant-time evolution period leads to reduced signal-to-noise ratio compared to the non-constant time alternative but is still favorable due

to its superior resolution of closely spaced signals (Mandel and Palmer, 1994), which is of critical importance in accordion experiments (see Results section). In the forward accordion experiment, where T_κ is incremented together with t_1, the effective relaxation rate constant of the interferogram is given by $R_{2,\mathrm{fwd}} = R_{\mathrm{inh}} + \kappa \cdot R_{1\rho}$, where R_{inh} describes line broadening due to static magnetic field inhomogeneity. By comparison, in a non-constant time experiment, the line width would be given by $R_2 + R_{\mathrm{inh}} + \kappa \cdot R_{1\rho}$, i.e., essentially a factor of 2 greater than in the constant time version (Mandel and Palmer, 1994). In the reverse accordion experiment, where T_κ is decremented as t_1 is incremented, the effective relaxation rate constant of the interferogram is given by $R_{2,\mathrm{rev}} = R_{\mathrm{inh}} - \kappa \cdot R_{1\rho}$. The rotating-frame relaxation rate constant is calculated as $R_{1\rho} = (R_{2,\mathrm{fwd}} - R_{2,\mathrm{rev}})/(2\kappa)$. As an alternative, $R_{1\rho}$ can also be determined by subtracting from $R_{2,\mathrm{fwd}}$ the line width measured in an interferogram from a reference experiment with the relaxation period set to 0: $R_{1\rho} = (R_{2,\mathrm{fwd}} - R_{\mathrm{ref}})/\kappa$. In estimating the standard error of $R_{1\rho}$ by error propagation, the factor of 2 difference in the numerator of the alternative approaches offsets the higher signal-to-noise ratio in the reference experiment as compared to the reverse experiment, with the result that the error bars are actually lower in the forward–reverse accordion approach.

2.2 NMR sample preparation

Uniformly ^{15}N-enriched galectin-3C was expressed and purified as described previously (Diehl et al., 2009, 2010; Wallerstein et al., 2021). The NMR sample containing galectin-3C in complex with the ligand 3′-[4-(3-fluorophenyl)-1H-1,2,3-triazol-1-yl]-3′-deoxy-β-D-galactopyranosyl-1-thio-β-D-glucopyranoside was prepared as described (Wallerstein et al., 2021) to yield a final protein concentration of 0.9 mM in 5 mM HEPES buffer, pH 7.3.

2.3 NMR relaxation experiments

All pulse sequences were based on the ^{1}H-^{15}N HSQC format (Bodenhausen and Ruben, 1980). Conventional and accordion ^{15}N $R_{1\rho}$ experiments were acquired with uniform sampling (US) on an Agilent/Varian VNMRS 600 MHz instrument equipped with a 5 mm HCN triple-resonance room temperature probe. To allow for comparisons between the two different ways of estimating $R_{1\rho}$ from accordion data, accordion experiments were performed using a combination of forward and reverse accordion modes (i.e., incrementing or decrementing the relaxation delay in step with t_1) together with a reference experiment excluding the accordion relaxation period but including the alignment blocks (Fig. 1). Both the conventional and accordion $R_{1\rho}$ experiments were performed with two different methods for aligning the magnetization along the effective spin-lock field axis: either hard pulses and delays (Hansen and Kay, 2007) or adiabatic amplitude/frequency ramps with a tan/tanh profile of 1.8 ms duration (Mulder et al., 1998). In the former case, the scaling factor δ was set to 1.35 for optimum alignment of spins with offsets within $\pm\omega_{\mathrm{SL}}$ from the spin-lock carrier frequency. The ^{15}N dimension was acquired with a spectral width of 2006 Hz, which was sampled over 132 increments in the accordion $R_{1\rho}$ experiment utilizing adiabatic alignment, with 128 increments in the accordion experiment using hard-pulse alignment, and with 128 increments in both conventional $R_{1\rho}$ experiments. The ^{1}H dimension was acquired with a spectral width of 8446 Hz, which was sampled over 2028 complex data points, in all experiments. All accordion experiments (forward, reverse, and reference) were acquired interleaved. Conventional $R_{1\rho}$ experiments were acquired by interleaving the relaxation periods of (6, 12, 23.9, 2 × 47.9, 95.7, and 191.4) ms. All $R_{1\rho}$ experiments employed a spin-lock field strength of $\omega_{\mathrm{SL}}/(2\pi) = 1380$ Hz. The effective spin-lock field strength in the rotating frame is given by $\omega_{\mathrm{eff}} = (\omega_{\mathrm{SL}}^2 + \Omega^2)^{1/2}$, where Ω is the offset from the spin-lock carrier (Akke and Palmer, 1996; Davis et al., 1994). The transverse relaxation constant R_2 was extracted from the $R_{1\rho}$ relaxation rates by correcting for off-resonance effects using the relationship $R_{1\rho} = \cos^2(\theta)R_1 + \sin^2(\theta)R_2$, where θ is the tilt angle of the spin-lock field defined by $\tan(\theta) = \omega_{\mathrm{SL}}/\Omega$ and the previously determined R_1 rate constant.

Conventional R_1 and R_2 CPMG experiments (Farrow et al., 1995; Skelton et al., 1993) were acquired with uniform sampling on a Bruker NEO 600 MHz instrument equipped with a 5 mm HPCN QCI cryoprobe, using spectral widths of 2129 Hz, which was sampled over 128 increments, in the ^{15}N dimension and 9615 Hz, which was sampled over 2048 complex data points, in the ^{1}H dimension. The R_1 and R_2 relaxation periods were acquired interleaved with the t_1 increments using delays of (2 × 0.04, 0.08, 0.12, 0.2, 0.4, 2 × 0.6, 0.72, 2 × 0.8, 1.0, 1.2, 1.6, and 2.0) s and (0, 42.2, 2 × 84.5, 126.7, and 169.0) ms, respectively. The R_1 experiment utilized ^{1}H WALTZ decoupling during the relaxation period. The R_2 experiment employed CPMG pulse trains with a fixed refocusing frequency, $\nu_{\mathrm{CPMG}} = 1/(2\tau_{\mathrm{CPMG}}) = 625$ Hz, where τ_{CPMG} is the delay between 180° pulses in the CPMG train, and a 180° pulse length of 80 µs. The R_2 values were not corrected for off-resonance effects. We also recorded CPMG relaxation dispersion datasets at static magnetic field strengths of 11.7 and 14.1 T, using Agilent/Varian spectrometers equipped with 5 mm HCN triple-resonance room temperature probes. The relaxation dispersion experiment was run as a constant-time version (Mulder et al., 2001) of the relaxation-compensated CPMG pulse sequence (Loria et al., 1999), using 18 refocusing frequencies acquired interleaved with 128 t_1 increments covering spectral widths of 1550 Hz (2006 Hz) at 11.7 T (14.1 T). The experiments at 11.7 and 14.1 T employed refocusing frequencies ν_{CPMG} of (2 × 0, 2 × 50, 2 × 100, 2 × 150, 2 × 200, 250, 300, 400, 500, 600, 700, 750, and 950) Hz and (2 × 0, 2 × 50, 2 × 100, 2 × 150,

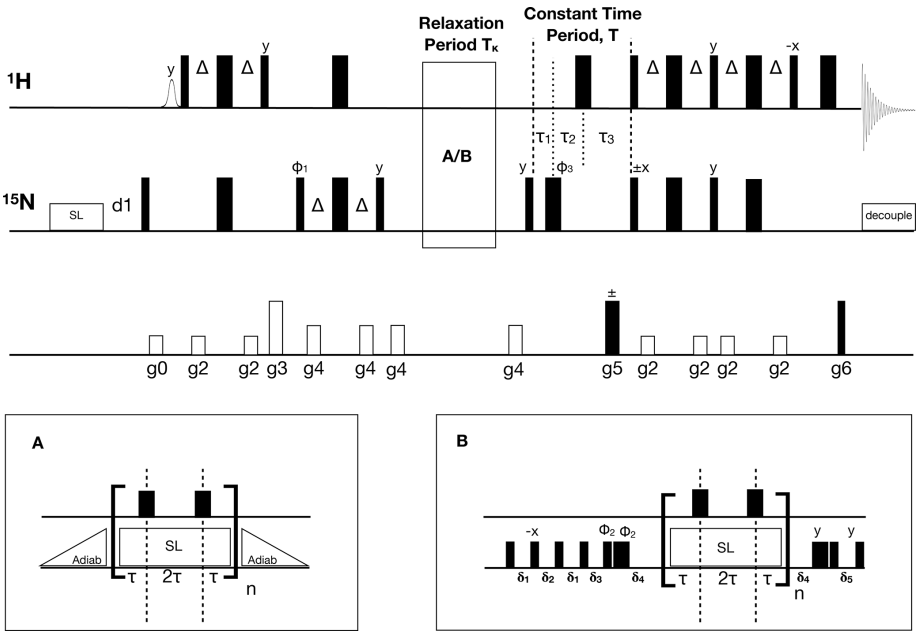


Figure 1. Pulse sequence for the ^{1}H-detected accordion-NUS ^{15}N $R_{1\rho}$ spin-lock experiment. Thin (thick) black bars correspond to 90° (180°) non-selective pulses. All pulses have phase x unless otherwise indicated. The spin-lock at the beginning of the sequence is a heat-compensation block (Wang and Bax, 1993). The open bell-shaped pulse at the beginning of the sequence is a selective pulse on the water resonance (Grzesiek and Bax, 1993). INEPT polarization transfer steps (Bodenhausen and Ruben, 1980; Morris and Freeman, 1979) use $\Delta = 1/(4J_{\mathrm{HN}})$, where J_{HN} is the one-bond scalar coupling constant. The relaxation period can be run with **(A)** adiabatic ramps or with **(B)** hard-pulse alignment blocks (Hansen and Kay, 2007). In both cases, 180° ^{1}H pulses are present at time points τ and 3τ to suppress cross-correlated relaxation (Massi et al., 2004). The hard pulse alignment delays are $\delta_1 = 1/(2\omega_{\mathrm{SL}}) - 2/\omega_N$, $\delta_2 = \delta/\omega_{\mathrm{SL}} - 2/\omega_N$, $\delta_3 = \delta/(2\omega_{\mathrm{SL}}) - 2/\omega_N$, $\delta_4 = 1/\omega_N$, and $\delta_5 = 1/(2\omega_{\mathrm{SL}}) - 2/\omega_N$, where ω_N is the field strength of the high-power ^{15}N 90° pulse, ω_{SL} is the spin-lock field strength, and $\delta = 1.35$ is a scaling factor optimized for alignment of spins with offsets within $\pm\omega_{\mathrm{SL}}$. The total relaxation delay is $T_\kappa = n \cdot 4 \cdot \tau$, where n is the sampled point number in t_1, and $\tau = (\kappa \cdot \Delta t_1)/4$, with $\kappa = 3$. The total constant-time period is $T = \tau_1 + \tau_2 + \tau_3 = 70$ ms, where $\tau_1 = (T - t_1)/2$, $\tau_2 = T/2 - \Delta$, and $\tau_3 = \Delta + t_1/2$. The phase cycle is $\varphi 1 = (x, -x)$, $\varphi 2 = (8(y), 8(-y))$, and $\varphi 3 = (x, x, y, y, -x, -x, -y, -y)$. The receiver phase cycle is rec $= (x, -x, -x, x, x, -x, -x, x)$ with relaxation block **(A)** and (rec, −rec) with block **(B)**. Gradient-enhanced PEP polarization transfer (Kay et al., 1992b; Palmer et al., 1991) is achieved by acquiring a second dataset with inversion of the phase of the ^{15}N 90° pulse indicated with $\pm x$ and gradient g5. The gradients g5 and g6 are used for coherence selection. The phase $\varphi 1$ and the receiver phase are inverted for each t_1 increment. The gradient times and levels are g0: 1 ms, 8.9 G cm^{-1}; g1: 1 ms, 8.9 G cm^{-1}; g2: 0.5 ms, 7.1 G cm^{-1}; g3: 1 ms, 44.4 G cm^{-1}; g4: 0.5 ms, 14.2 G cm^{-1}; g5: 1.25 ms, 53.2 G cm^{-1}; and g6: 0.125 ms, 53.8 G cm^{-1}.

2×200, 2×250, 350, 450, 550, 700, 850, and 1000) Hz, respectively.

2.4 Non-uniform sampling schemes

NUS schemes were generated by selecting data points from the uniformly sampled accordion $R_{1\rho}$ dataset. CRLB-optimized NUS schemes were obtained as described previously (Carlström et al., 2019). In practice, scheme optimization can be performed using modeled data constructed by taking known values of A_k, ω_k, and R_k obtained from DSURE estimation of an HSQC experiment, together with an estimate of the extra decay caused by relaxation during the accordion period (Carlström et al., 2019). Single-column CRLB-optimized (col-opt) (Carlström et al., 2019; Swärd et al., 2018) and sine-weighted Poisson-gap (Hyberts et al., 2010) sampling schemes were implemented for the accordion $R_{1\rho}$ dataset with adiabatic ramps, using in-house MATLAB scripts. Poisson-gap sampling schemes were generated by randomly varying the argument of the sinusoidal weighting function between 0 and $\pi/2$; see (Hyberts et al., 2010). The best sampling scheme was identified as the one having the lowest sum of the CRLB calculated over all columns containing peaks. In the case of the col-opt approach, the se-

lection was made among 97 different single-column CRLB-optimized schemes corresponding to each slice of the interferogram containing protein signals, whereas in the case of Poisson-gap sampling 1000 different schemes were compared. The sampling scheme was optimized individually for each of the reference, forward, and reverse accordion experiments. In each case, we generated individual datasets sampled with $N = 22$, 27, 32, 37, 42, 47, 52, 57, 62, 66, 67, 72, 77, 82, 87, 92, 97, 102, 107, 112, 117, 122, or 127 increments in the indirect dimension. The NUS datasets resulting from the different sampling schemes were subsequently reconstructed using the DSURE algorithm.

2.5 Data reconstruction, processing, and analysis

Non-accordion (i.e., conventional) datasets were processed using NMRPipe (Delaglio et al., 1995), with forward linear prediction to double the number of data points, cosine-squared window apodization, and zero-filling to twice the size rounded to the nearest power of two. $R_{1\rho}$, R_1, and R_2 relaxation rate constants were estimated from the conventional experiments by integrating the peak volumes using PINT (Ahlner et al., 2013; Niklasson et al., 2017), followed by fitting mono-exponential decay functions to the volumes using in-house MATLAB scripts. The fitted R_2 values determined by CPMG experiments were not adjusted for off-resonance effects (Korzhnev et al., 2000). Standard errors were estimated using jackknife resampling (Mosteller and Tukey, 1977) as implemented in PINT. Standard errors in R_2 derived from $R_{1\rho}$ experiments were estimated by Monte Carlo simulations using 10 000 samples drawn from normal distributions with widths corresponding to the standard errors of $R_{1\rho}$ and R_1 (Press et al., 1986).

The accordion datasets were processed and analyzed using the DSURE algorithm (Juhlin et al., 2018) implemented in MATLAB (MathWorks, Inc.). DSURE reconstruction was performed on individual t_1 interferograms, as described previously (Carlström et al., 2019). DSURE models interferograms as sums of exponentially decaying sinusoids:

$$A(t) = \sum_{k}^{K} A_k \exp[i\omega_k t - R_k t] + \varepsilon(t), \quad (1)$$

where A_k, ω_k, and R_k are the complex-valued amplitude, frequency, and decay rate of the kth signal, respectively, $\varepsilon(t)$ represents additive noise, and the sum runs over all K signals identified in a given interferogram. In reconstructing accordion data, the time domain data from the reverse mode was inverted and complex conjugated before estimation using DSURE. Standard errors of the estimated parameters were calculated as the CRLB, which is very close to the RMSE for statistically efficient estimators like DSURE. Explicit comparison of the RMSE, calculated from Monte Carlo simulations using 1000 samples, and the CRLB indicates that the two measures are in excellent agreement and deviate by at most a factor of 1.6 for the worst case (an interferogram containing three signal maxima) among our 50 % NUS data.

2.6 Statistical analysis

To compare the performance of the different approaches for measuring transverse relaxation rates, we used four different metrics. The relative difference and absolute deviation between datasets x and y are defined for a given residue i as $\Delta_{\mathrm{rel}} = 2(x_i - y_i)/(x_i + y_i)$ and $\Delta_{\mathrm{abs}} = |x_i - y_i|$, respectively. The root-mean-square deviation (RMSD) between two datasets is calculated pairwise over all residues (N_{res}), $\mathrm{RMSD} = [(\sum_i (x_i - y_i)^2 / N_{\mathrm{res}})]^{1/2}$. The mean relative uncertainty (MRU) of a given dataset is the mean, calculated over all residues, of the individual uncertainty in x_i (σ_{xi}; 1 standard deviation, as estimated by the DSURE algorithm) divided by x_i: $\mathrm{MRU} = (\sum_i \sigma_{xi}/x_i)/N_{\mathrm{res}}$.

3 Results and discussion

3.1 Pulse sequence design

The accordion-NUS $R_{1\rho}$ pulse sequence (Fig. 1) is based on our previous implementation to measure R_1 (Carlström et al., 2019), which included minor modifications of the original constant-time accordion experiment (Mandel and Palmer, 1994). In designing accordion-NUS versions of transverse relaxation experiments, it is necessary to consider the interplay between the minimum length of the relaxation block (A or B in Fig. 1), the number of sampled t_1 points, and the maximum attainable t_1 value. R_2 relaxation rate measurement is typically performed in one of two ways: using either CPMG pulse trains or a continuous spin-lock during the relaxation period so as to maintain in-phase magnetization and avoid significant evolution into anti-phase terms (Skelton et al., 1993), as well as reduce chemical/conformational exchange (R_{ex}) and magnetic-field inhomogeneity (R_{inh}) contributions to the effective transverse relaxation rate constant ($R_{2,\mathrm{eff}}$). Furthermore, it is necessary to suppress the effects of cross-correlated relaxation, which amounts to introducing additional relaxation delays (Kay et al., 1992a; Palmer et al., 1992) and to mitigate the effects of off-resonance effects and pulse imperfections (Korzhnev et al., 2000), leading to extended phase cycles (Yip and Zuiderweg, 2004). CPMG-type experiments for measuring chemical exchange involve extended spin-echo elements to average the relaxation rates of in-phase and anti-phase coherences (Loria et al., 1999). All in all, these requirements typically lead to relatively long relaxation blocks in CPMG-based experiments. Since the accordion experiment increments (or decrements) the relaxation period synchronously with the t_1 period, the minimum increment step for the relaxation period limits the maximum number of points that can be acquired in the t_1 dimension. In our initial testing of CPMG-based accordion experiments to measure R_2, we found that the maximum achievable length

of the t_1 dimension was 64 points, before the duty cycle and relaxation losses became serious concerns. While the resulting resolution in t_1 might suffice in certain cases, we opted instead for increased flexibility and designed the transverse relaxation experiment based on a spin-lock period. This strategy allows for significantly shorter increments of the relaxation period, and further enables facile adaptation to off-resonance $R_{1\rho}$ experiments for conformational exchange measurement. We implemented two types of pulse sequence elements to align the magnetization along the effective spin-lock field: adiabatic amplitude/frequency ramps (Mulder et al., 1998) or an element comprising hard pulses and delays (Hansen and Kay, 2007). The hard-pulse element is shorter than the adiabatic ramp (0.25–0.45 ms versus 1.8 ms in the present case) and in principle reduces relaxation losses, while the adiabatic ramp achieves superior alignment over a wider range of offsets, making it suitable for off-resonance $R_{1\rho}$ experiments used to characterize chemical exchange processes.

3.2 Comparison of DSURE-modeled accordion $R_{1\rho}$ relaxation data and conventional relaxation data

The ^{1}H-^{15}N 2D spectrum resulting from the accordion $R_{1\rho}$ relaxation data reconstructed using DSURE is shown in Fig. A1, together with representative examples of DSURE models of interferograms. We compared the performance of the accordion $R_{1\rho}$ experiments acquired with the two different alignment elements (adiabatic vs hard-pulse; see Sect. 2.3) and also compared the results obtained using the two different combinations of accordion modes (forward–reverse vs forward–reference; see Sect. 2.1). To validate the accordion $R_{1\rho}$ values determined using DSURE, we first compared these with rate constants determined from the conventional $R_{1\rho}$ experiment and the R_2 CPMG experiment (Fig. 2). Figure 2 shows the results obtained using the forward–reverse accordion data and adiabatic alignment, while the corresponding data obtained using hard-pulse alignment is highly similar and shown in Fig. A2. In general, the results are in very good agreement, with very few residues showing statistically significant deviations between experiments (Fig. 2a–d). Comparing the accordion $R_{1\rho}$ values with the conventional data, we obtain an RMSD of $0.45\,\mathrm{s}^{-1}$, whereas the comparison with the CPMG data yields an RMSD of $0.59\,\mathrm{s}^{-1}$. The CPMG data were not corrected for off-resonance effects (Korzhnev et al., 2000), leading to offset-dependent systematic errors of up to 5 % that might explain the somewhat poorer agreement in this case. The distributions of relative differences are centered around the mean values 0.02 and $0.00\,\mathrm{s}^{-1}$ and are sharper than normal distributions (Fig. 2e, f). The means of relative differences should be compared with the average relative uncertainties of the estimated $R_{1\rho}$ values, which are $0.14\,\mathrm{s}^{-1}$ for the conventional data and $0.21\,\mathrm{s}^{-1}$ for the accordion data, indicating that the accordion $R_{1\rho}$ experiment yields accurate data of comparable precision compared to the conventional experiment. The slight tendency towards higher R_2 values determined from the accordion experiment reflects small but noticeable differences for a subset of residues (viz. residues 151, 154, 181, 182, and 184). Visual inspection of these peaks indicates that these differences are likely due to overlap problems, which are exacerbated by the additional line broadening present in accordion spectra. In principle this problem could be mitigated by optimizing the accordion scaling factor κ or by acquiring data as a 3D experiment (Carr et al., 1998; Chen and Tjandra, 2009). For some residues, notably L219, the R_2 value from the reference CPMG dataset is considerably higher (Fig. 2b), reflecting the different levels of residual exchange contributions to the transverse relaxation rate resulting from the different refocusing frequencies. The L219 peak has an offset of −572 Hz from the ^{15}N spin-lock carrier, which results in an effective spin-lock field of $\omega_{\mathrm{eff}}/(2\pi) = 1494$ Hz, which is more than a factor of 2 greater than the effective CPMG refocusing field (625 Hz) in the reference R_2 dataset. L219 also shows clear signatures of fast conformational exchange in CPMG relaxation dispersion experiments (Fig. A3).

The two alignment variants yield highly similar results in the context of the accordion experiment, and the same is true for the two combinations of accordion modes (Fig. A4). The RMSD between the two alignment variants is $0.19\,\mathrm{s}^{-1}$ and the mean relative difference is $0.0\,\mathrm{s}^{-1}$, and the corresponding numbers for the two combinations of accordion modes are 0.24 and $0.0\,\mathrm{s}^{-1}$, respectively. In the following presentation of accordion-NUS experiments, we will base all analyses on the results obtained from the forward–reverse accordion approach using data acquired with adiabatic alignment.

3.3 Comparison of non-uniformly sampled and uniformly sampled accordion $R_{1\rho}$ relaxation data

Next, we tested the performance of the accordion $R_{1\rho}$ experiment acquired with NUS. We have previously evaluated the performance of various NUS schemes for the acquisition of accordion R_1 data, and found that superior results were obtained for schemes generated by column-wise optimization directly against the CRLB (denoted col-opt in the following) or schemes following the Poisson-gap distribution (Carlström et al., 2019). Therefore, we restrict our present comparisons to the performance of these two NUS schemes. Starting from the uniformly sampled accordion $R_{1\rho}$ dataset acquired with adiabatic ramps and the forward–reverse accordion mode (from here on denoted the US dataset), we generated two times two datasets, where we used either col-opt or Poisson-gap sampling schemes, both optimized for the forward accordion experiment alone (set F) or optimized individually for each of the forward and reverse experiments (set F+R). In a real case scenario, it is arguably more practical to perform the optimization on model datasets constructed by taking known values of A_k, ω_k, and R_k obtained from DSURE estimation of an HSQC experiment, together with

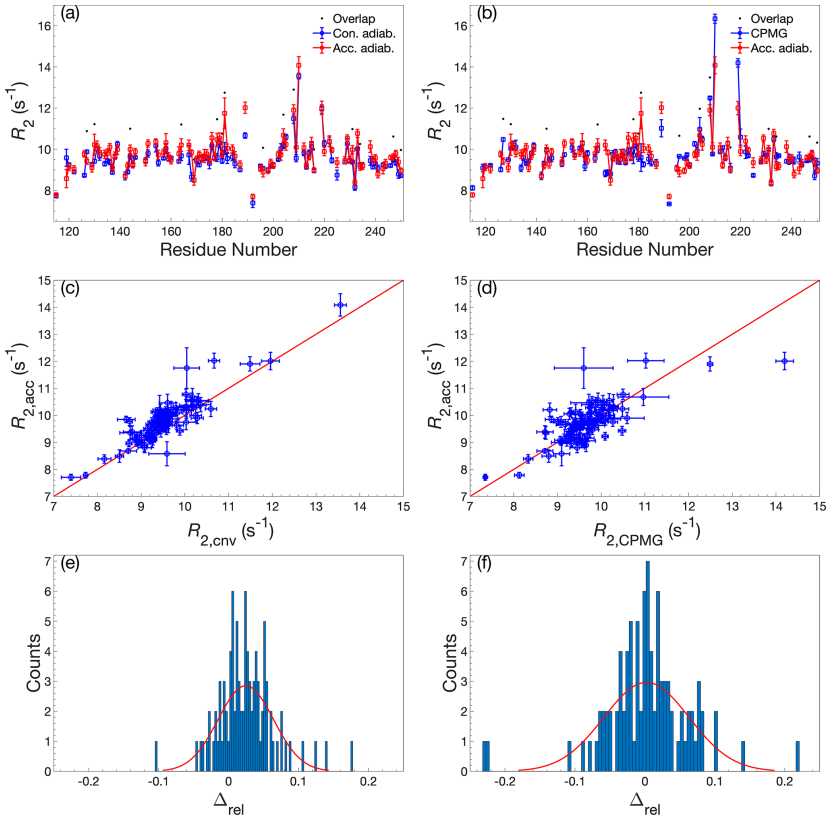


Figure 2. Comparison of R_2 determined by accordion $R_{1\rho}$ or conventional relaxation experiments. R_2 values determined by accordion $R_{1\rho}$ (red) spin-lock experiments compared with **(a, c, e)** R_2 determined by conventional $R_{1\rho}$ (blue) and **(b, d, f)** R_2 determined by conventional R_2 CPMG (blue). Both $R_{1\rho}$ experiments were acquired with adiabatic ramps. **(a, b)** R_2 plotted versus residue number. Black dots indicate residues showing significant overlap in the ^{1}H-^{15}N HSQC spectrum. **(c, d)** Covariance plot of R_2 datasets. **(e, f)** Histogram of the relative differences between datasets. The red curve describes the normal distribution that best fits the data. In panels **(a)**–**(d)**, error bars indicate ± 1 SD.

an estimate of the extra decay imparted by the accordion period (Carlström et al., 2019). Thus, there is no need to first record a US accordion dataset prior to optimizing the NUS scheme.

Figure 3 illustrates the performance of the accordion-NUS $R_{1\rho}$ experiment acquired with different NUS schemes (red and blue symbols) and optimization protocols (left- and right-hand columns). We compared the $R_{1\rho}$ values determined by accordion-NUS with those obtained from the accordion-US dataset. In general, the performance decreases with decreasing number of sampled points, as might be expected. The RMSD between the NUS and US datasets shows a clear trend towards higher values as the number of sampling points decrease, from less than 0.2 s^{-1} at N_{full} to 0.8–1.0 s^{-1} at $N = 22$ or 18 % sampling density (Fig. 3a, b). However, these plots show some degree of scatter, which reflects the random nature of the NUS schemes, where any given scheme with a lower number of points might yield lower RMSD than another scheme with higher number of points. By contrast, the mean relative uncertainty (MRU) in the estimated $R_{1\rho}$ parameter shows an essentially monotonous increase with decreasing number of points, from 2.2 % at N_{full} to 5 % at $N = 22$ (Fig. 3c, d). The increasing uncertainty is relatively modest down to about 50 % sampling ($N = 66$), where MRU is ca. 2.8 %, but beyond this point both the RMSD and the MRU start to increase more steeply. These results are rather similar for the Poisson-gap and col-opt optimized schemes, with a small advantage for col-opt schemes, espe-

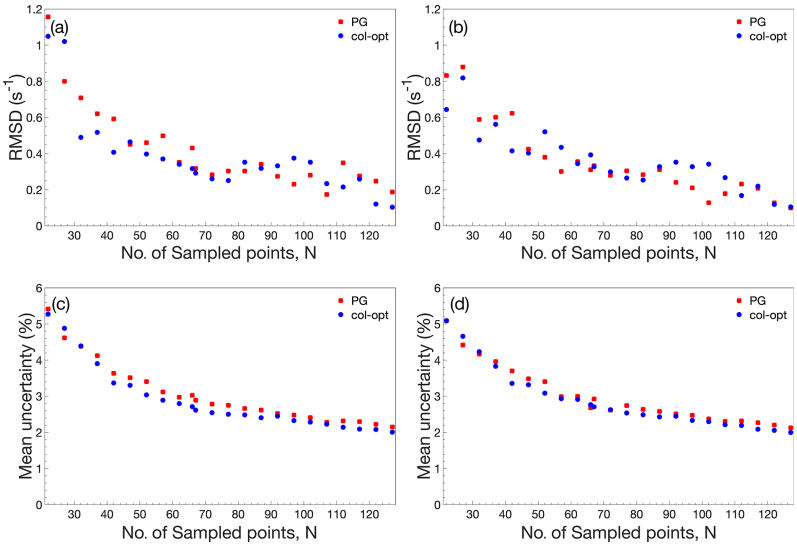


Figure 3. Performance comparison of accordion-NUS $R_{1\rho}$ experiments. **(a, b)** RMSD between the NUS dataset and the corresponding US dataset. **(c, d)** Mean relative uncertainty of the NUS $R_{1\rho}$ estimates. The left-hand side panels **(a)** and **(c)** show results obtained with NUS schemes optimized only for the forward accordion experiment, while the right-hand side panels **(b)** and **(d)** show the corresponding results obtained with schemes optimized individually for the forward and reverse accordion experiments. All data were acquired using adiabatic alignment.

cially in the case of the precision of the estimated parameters (Fig. 3c, d). However, greater improvements in precision are expected for relaxation rate constants of signals in interferograms whose sampling schemes have been individually optimized with respect to the CRLB, as shown previously (Carlström et al., 2019). Altogether, these results indicate that our present implementation of the accordion-NUS approach to measure $R_{1\rho}$ achieves equally good precision of the estimated relaxation rate constants as did our previously presented accordion-NUS R_1 experiment (Carlström et al., 2019). Furthermore, the sampling schemes optimized separately for each of the forward and reverse experiments (F+R) show a modest advantage in performance over F for low N, which might be expected (compare Fig. 3a with Fig. 3b and Fig. 3c with Fig. 3d).

3.4 Spectral characteristics affecting accuracy and precision of accordion-NUS $R_{1\rho}$ relaxation rate constants

Next, we investigated how various spectral characteristics affect the accuracy and precision of the estimated relaxation rate constant. We calculated the absolute deviation (Δ_{abs}) between the US estimate and the 50 % NUS estimate ($N = 66$), as well as the relative uncertainty, for each residue and plotted the results against signal intensity, resonance frequency offset from the spin-lock carrier, and the number of estimated signals present in the interferogram of the col-opt and F+R optimized data (Fig. 4). There is no obvious relationship between Δ_{abs} and signal intensity, although larger values of Δ_{abs} ($> 0.4\,\text{s}^{-1}$) are not observed for the most intense signals (Fig. 4a). However, there is a trend towards lower relative uncertainty with higher signal intensity (Fig. 4d), where a value of 1.5 % is observed for the strongest signals and 6 %–8 % at the other extreme. The results further reveal that the number of signals in the interferogram has an effect on both Δ_{abs} and the relative uncertainty, with a trend toward slightly larger errors as the number of signals increases (Fig. 4c, f); the relative uncertainty varies from 1.5 % for single signals to 6 %–8 % for the worst cases among interferograms containing seven signals. Reassuringly, the mean relative uncertainty increases only slightly from 2 % for single peaks to 3.2 % for seven peaks. Thus, there is no dramatic decrease in performance even at the highest number of signals. This effect of the number of signals also explains the apparent higher Δ_{abs} and higher relative uncertainty for residues with offsets around 0 and 500 Hz, because this region of the spectrum is the most crowded (Fig. 4b, e). Furthermore, this result mirrors the observations of deviations between the accordion and conventional data discussed above in connection with Fig. 2.

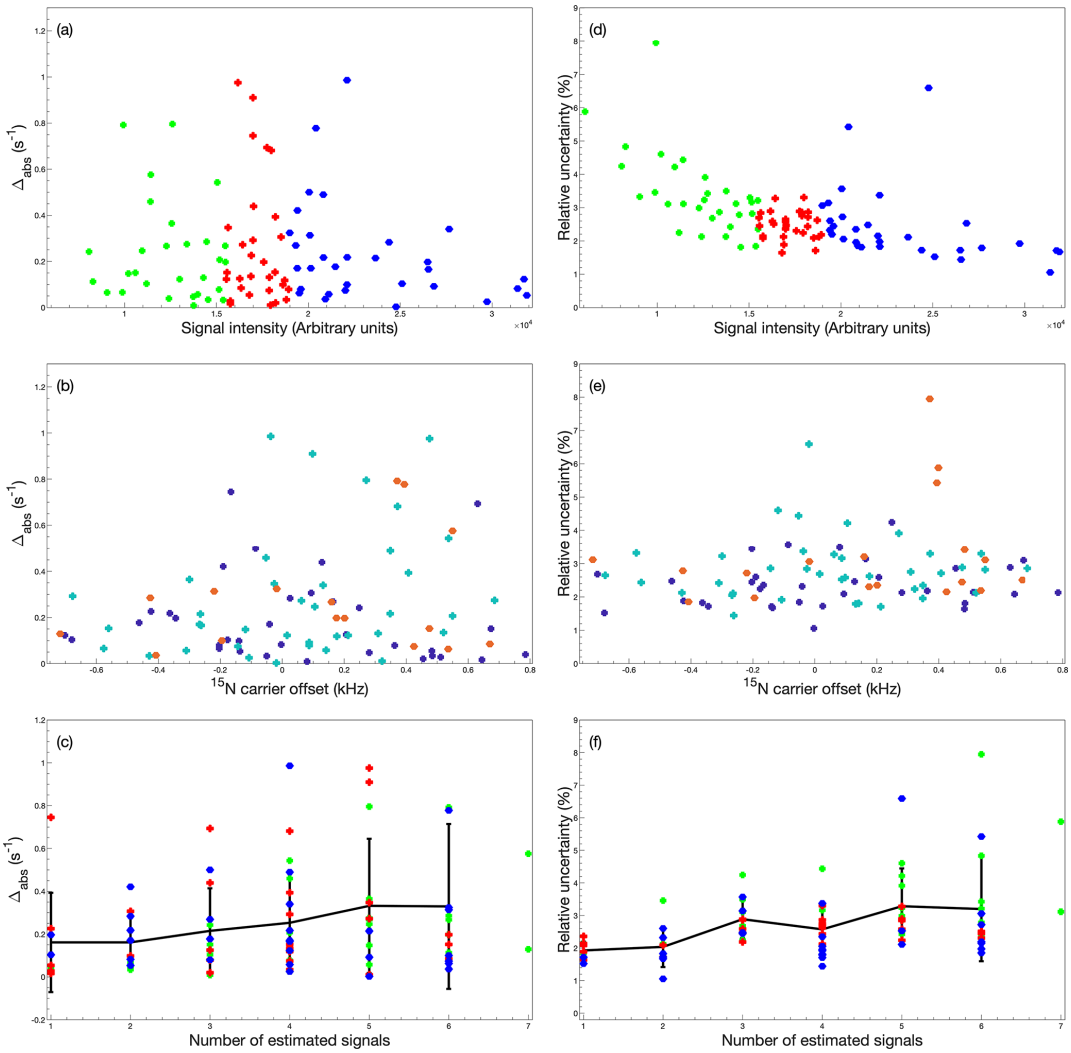


Figure 4. Dependence of accordion-NUS $R_{1\rho}$ accuracy and precision on spectrum characteristics. **(a–c)** Absolute deviation (Δ_{abs}) between $R_{1\rho}$ values obtained from US and 50 % NUS data and **(d–f)** relative uncertainty ($\sigma_i / R_{1\rho,i}$) of $R_{1\rho}$ values obtained from 50 % NUS data, plotted as a function of **(a, d)** signal intensity, **(b, e)** ^{15}N spin-lock carrier offset, and **(c, f)** number of estimated signals in the interferogram. In panels **(a)**, **(c)**, **(d)**, and **(f)**, the data are divided into tertiles according to signal intensity and color-coded as green, first tertile (lowest intensity); red, second tertile; and blue, third tertile. In panels **(b)** and **(e)**, the data are divided into subsets according to the number of estimated signals in the interferogram and color-coded as purple, 1, 2, or 3 signals; cyan, 4 or 5 signals; and orange, 6 or 7 signals. The black symbols with error bars in panels **(e)** and **(f)** represent the average and standard deviation of all data with a given number of peaks in the interferogram. All data were acquired using adiabatic alignment and determined using col-opt and F+R optimized NUS schemes.

4 Conclusions

We have described a non-uniformly sampled accordion $R_{1\rho}$ experiment that complements the previously presented accordion-NUS R_1 experiment (Carlström et al., 2019). The present accordion-NUS $R_{1\rho}$ experiment allows for accurate and precise measurement of the transverse relaxation rate constant R_2 while reducing sampling of the indirect dimension by at least 50 %. The combination of accordion relaxation rate measurements with NUS achieves a time saving of an order of magnitude compared to conventional experiments, in keeping with previous results presented for the corresponding accordion-NUS R_1 experiment (Carlström et al., 2019). In addition to on-resonance R_2 measurements, demonstrated herein, we anticipate that this experiment will be useful for on- and off-resonance $R_{1\rho}$ experiments to characterize chemical exchange processes. The accordion-NUS approach has broad applications in heteronuclear relaxation studies; with suitable modifications, the pulse sequence reported here for backbone ^{15}N spins should be applicable to many other sites, e.g., ^{13}C spins.

Appendix A

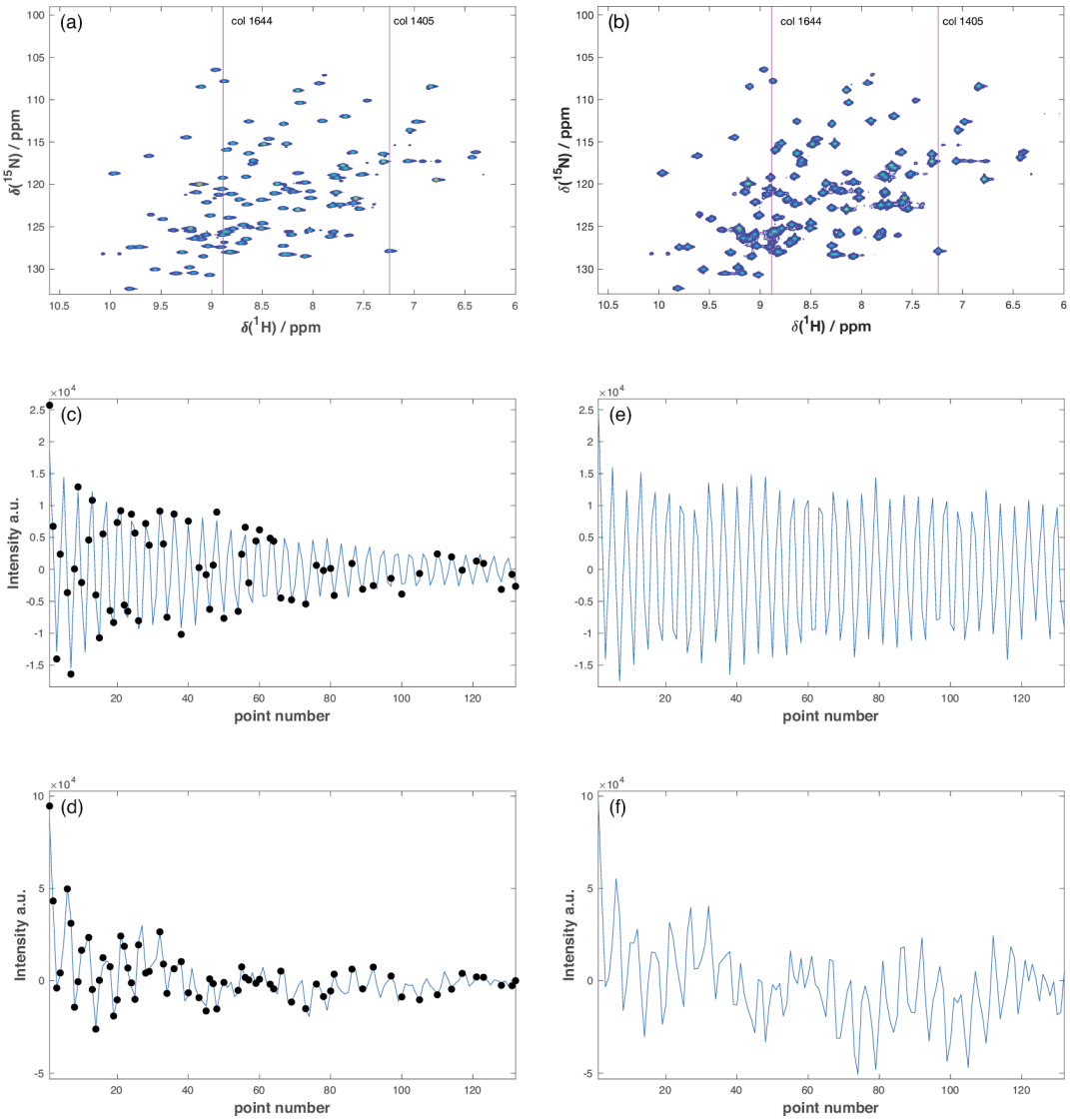


Figure A1. Representative accordion $R_{1\rho}$ datasets. **(a, b)** Full 2D ^{1}H-^{15}N HSQC spectra from the **(a)** US reference experiment and **(b)** 50 % NUS forward accordion experiment following DSURE reconstruction. The red lines identify column 1405 containing a single peak and column 1644 containing eight peaks. **(c, d)** DSURE models of the interferograms of **(c)** column 1405 and **(d)** column 1644 extracted from the 50 % NUS accordion experiment. The black dots indicate the sampled data points, while the blue curve shows the interferogram reconstructed by DSURE. **(e, f)** Interferograms of **(e)** column 1405 and **(f)** column 1644 extracted from the US reference experiment. Due to the constant time evolution period in t_1, the interferograms in the reference experiment **(e, f)** show essentially no decay, whereas the interferograms in the accordion experiment **(c, d)** show significant decays due to the encoded $R_{1\rho}$ relaxation rate(s).

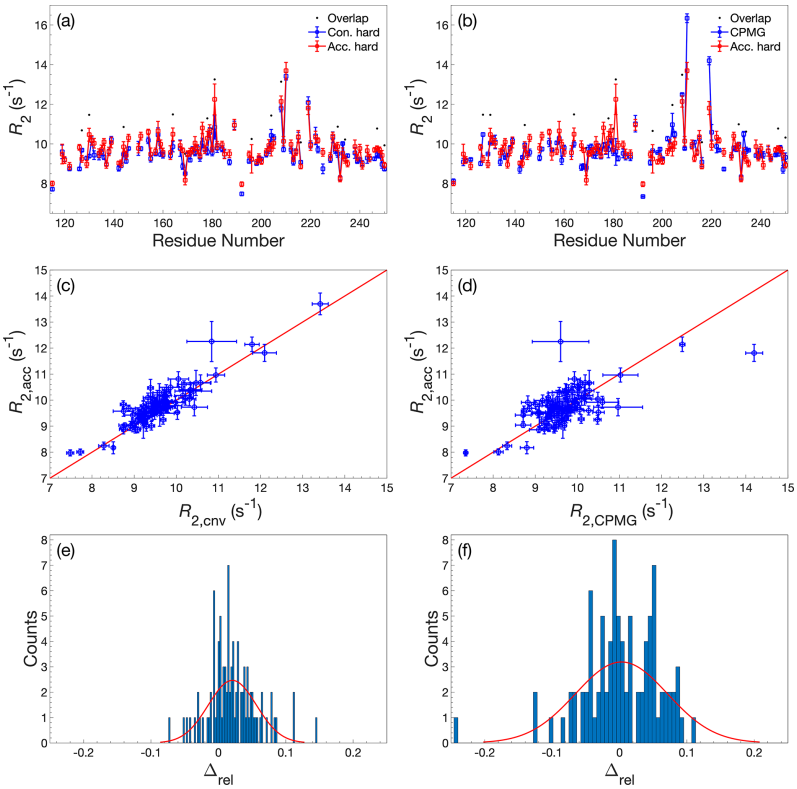


Figure A2. Comparison of R_2 determined by accordion $R_{1\rho}$ or conventional relaxation experiments. R_2 values determined by accordion $R_{1\rho}$ (red) spin-lock experiments compared with **(a, c, e)** R_2 determined by conventional $R_{1\rho}$ (blue) and **(b, d, f)** R_2 determined by conventional R_2 CPMG (blue). Both $R_{1\rho}$ experiments were acquired with hard-pulse alignment. **(a, b)** R_2 plotted versus residue number. Black dots indicate residues showing significant overlap in the ^{1}H-^{15}N HSQC spectrum. **(c, d)** Covariance plot of R_2 datasets. **(e, f)** Histogram of the relative differences between datasets. The red curve describes the normal distribution that best fits the data. In panels **(a)**–**(d)**, error bars indicate ± 1 SD.

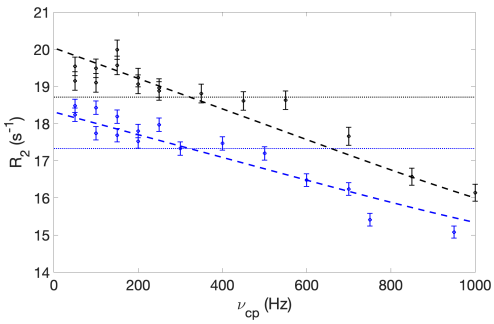


Figure A3. CPMG relaxation dispersion profiles of residue L219, acquired at static magnetic field strengths of 11.7 T (blue) and 14.1 T (black). Dashed lines show the results of fitting a two-state exchange model globally to both datasets, which yielded an exchange rate of $7150 \pm 4100\,s^{-1}$, major state population of 0.99 ± 0.01, and chemical shift difference of 7 ± 4 ppm. The dotted horizontal lines show the average R_2 value at each field strength.

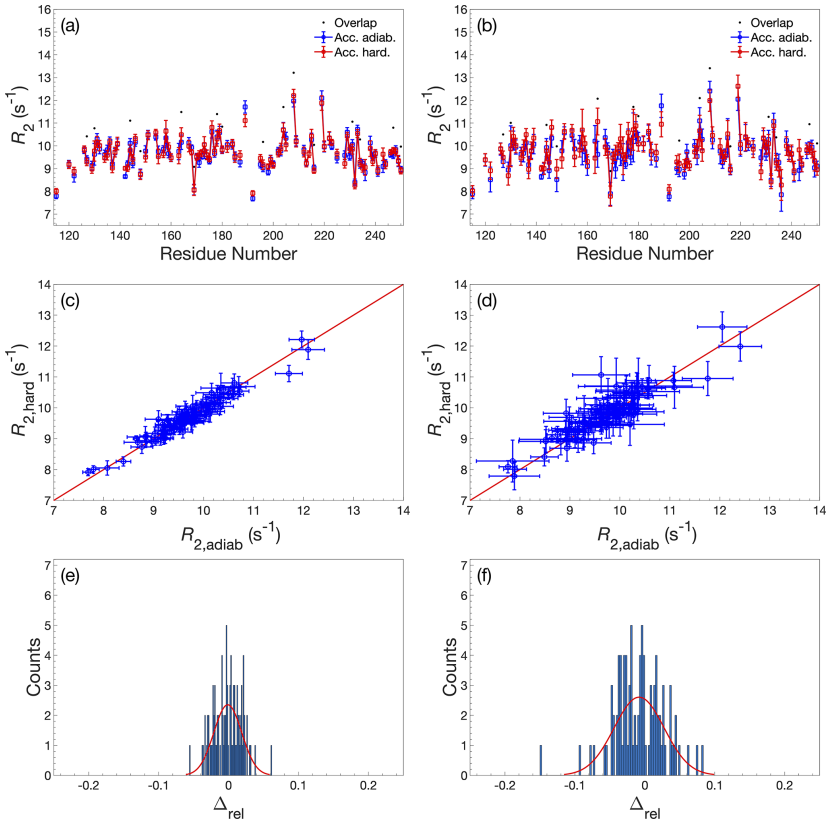

(a)
Overlap
Acc. adiab.
Acc. hard.
R_2 (s^{-1})
Residue Number
(b)
Overlap
Acc. adiab.
Acc. hard.
R_2 (s^{-1})
Residue Number
(c)
$R_{2,hard}$ (s^{-1})
$R_{2,adiab}$ (s^{-1})
(d)
$R_{2,hard}$ (s^{-1})
$R_{2,adiab}$ (s^{-1})
(e)
Counts
Δ_{rel}
(f)
Counts
Δ_{rel}

Figure A4.

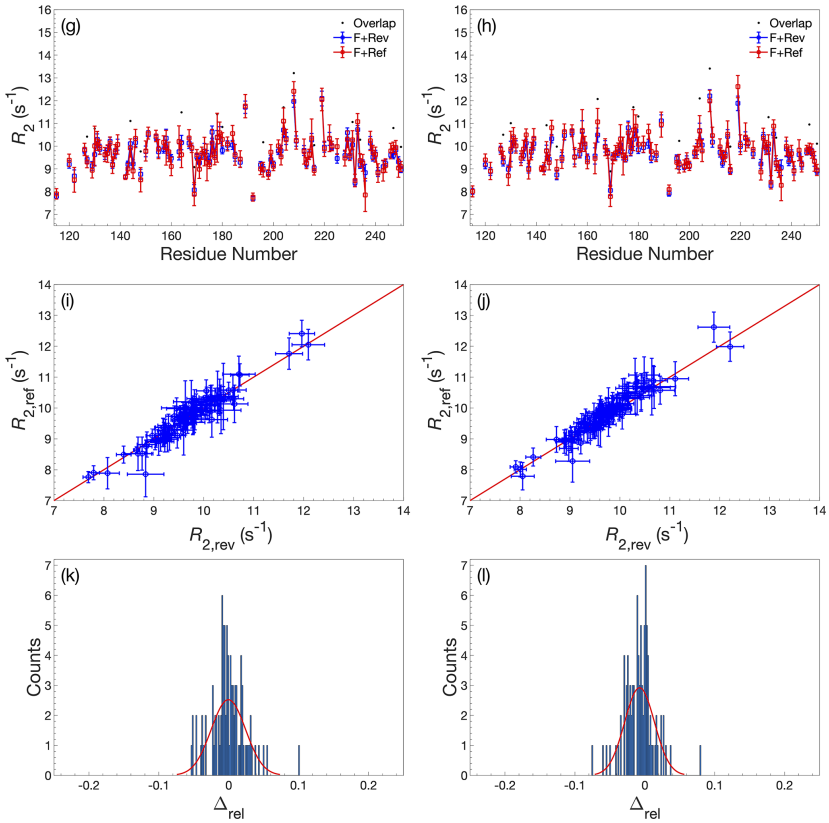


Figure A4. Comparison of R_2 determined by accordion $R_{1\rho}$ using various acquisition schemes. **(a–f)** Comparison of R_2 values determined using either adiabatic alignment (blue) or hard-pulse alignment (red). **(g–l)** Comparison of R_2 values determined using either forward–reverse (blue, F+Rev) accordion modes or forward–reference (red, F+Ref). **(a, c, e)** R_2 determined by forward–reverse accordion modes; **(b, d, f)** R_2 determined by forward–reference accordion modes. **(a, b, g, h)** R_2 plotted versus residue number. Black dots indicate residues showing significant overlap in the ^{1}H-^{15}N HSQC spectrum. **(c, d, i, j)** Covariance plot of R_2 datasets. **(e, f, k, l)** Histogram of the relative differences between datasets. The red curve describes the normal distribution that best fits the data. In panels **(a)–(d)** and **(g)–(j)**, error bars indicate ± 1 SD.

Code and data availability. Backbone chemical shift assignments have been deposited at the Biological Magnetic Resonance Bank with the accession code 50283 (https://doi.org/10.13018/BMR50283, Wallerstein et al., 2020). The NMR pulse sequences, relaxation datasets, and extracted relaxation rate constants are available at Mendeley Data together with MATLAB scripts implementing the DSURE algorithm (https://doi.org/10.17632/zyryxrgkc3.1; Wernersson et al., 2021).

Author contributions. SW, GC, and MA conceived the overall experimental approach and designed pulse sequences. AJ contributed the DSURE algorithm and designed the analysis protocol together with GC. SW and GC performed experiments. SW, GC, and MA analyzed data. All authors contributed to writing the article.

Competing interests. Mikael Akke is an associate editor of *Magnetic Resonance*.

Disclaimer. Publisher's note: Copernicus Publications remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Special issue statement. This article is part of the special issue "Geoffrey Bodenhausen Festschrift". It is not associated with a conference.

Financial support. This research has been supported by the Swedish Research Council (grant no. 2018-4995) and the Swedish SRA eSSENCE (grant no. 2020 6:2).

Review statement. This paper was edited by Paul Schanda and reviewed by Dominique Frueh and two anonymous referees.

References

Ahlner, A., Carlsson, M., Jonsson, B. H., and Lundström, P.: PINT: A software for integration of peak volumes and extraction of relaxation rates, J. Biomol. NMR, 56, 191–202, https://doi.org/10.1007/s10858-013-9737-7, 2013.

Akke, M. and Palmer, A. G.: Monitoring Macromolecular Motions on Microsecond–Millisecond Time Scales by R1r–R1 Constant-Relaxation-Time NMR Spectroscopy, J. Am. Chem. Soc., 118, 911–912, 1996.

Alderson, T. R. and Kay, L. E.: Unveiling invisible protein states with NMR spectroscopy, Curr. Opin. Struct. Biol., 60, 39–49, https://doi.org/10.1016/j.sbi.2019.10.008, 2020.

Bodenhausen, G. and Ernst, R. R.: The accordion experiment, a simple approach to three-dimensional NMR spectroscopy, J. Magn. Reson., 45, 367–373, https://doi.org/10.1016/0022-2364(81)90137-2, 1981.

Bodenhausen, G. and Ernst, R. R.: Direct determination of rate constants of slow dynamic processes by two-dimensional "accordion" spectroscopy in nuclear magnetic resonance, J. Am. Chem. Soc., 104, 1304–1309, 1982.

Bodenhausen, G. and Ruben, D. J.: Natural abundance nitrogen-15 NMR by enhanced heteronuclear spectroscopy, Chem. Phys. Lett., 69, 185–189, 1980.

Carlström, G., Elvander, F., Swärd, J., Jakobsson, A., and Akke, M.: Rapid NMR relaxation measurements using optimal nonuniform sampling of multidimensional accordion data analyzed by a sparse reconstruction method, J. Phys. Chem. A, 123, 5718–5723, https://doi.org/10.1021/acs.jpca.9b04152, 2019.

Carr, P. A., Fearing, D. A., and Palmer, A. G.: 3D Accordion Spectroscopy for Measuring 15N and 13CO Relaxation Rates in Poorly Resolved NMR Spectra, J. Magn. Reson., 132, 25–33, https://doi.org/10.1006/jmre.1998.1374, 1998.

Chen, K. and Tjandra, N.: Direct measurements of protein backbone 15N spin relaxation rates from peak line-width using a fully-relaxed Accordion 3D HNCO experiment, J. Magn. Reson., 197, 71–76, https://doi.org/10.1016/j.jmr.2008.12.001, 2009.

Davis, D. G., Perlman, M. E., and London, R. E.: Direct measurements of the dissociation-rate constant for inhibitor-enzyme complexes via the T1r and T2 (CPMG) methods, J. Magn. Reson. Ser. B, 104, 266–275, 1994.

Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A.: NMRPipe: A multidimensional spectral processing system based on UNIX pipes, J. Biomol. NMR, 6, 277–293, 1995.

Diehl, C., Genheden, S., Modig, K., Ryde, U., and Akke, M.: Conformational entropy changes upon lactose binding to the carbohydrate recognition domain of galectin-3, J. Biomol. NMR, 45, 157–169, https://doi.org/10.1007/s10858-009-9356-5, 2009.

Diehl, C., Engström, O., Delaine, T., Håkansson, M., Genheden, S., Modig, K., Leffler, H., Ryde, U., Nilsson, U. J., and Akke, M.: Protein flexibility and conformational entropy in ligand design targeting the carbohydrate recognition domain of galectin-3, J. Am. Chem. Soc., 132, 14577–14589, 2010.

East, K. W., Delaglio, F., and Lisi, G. P.: A simple approach for reconstruction of non-uniformly sampled pseudo-3D NMR data for accurate measurement of spin relaxation parameters, J. Biomol. NMR, https://doi.org/10.1007/s10858-021-00369-7, online first, 2021.

Farrow, N. A., Zhang, O., Forman-Kay, J. D., and Kay, L. E.: Comparison of the backbone dynamics of a folded and an unfolded SH3 domain existing in equilibrium in aqueous buffer, Biochemistry, 34, 868–878, 1995.

Gołowicz, D., Kasprzak, P., Orekhov, V., and Kazimierczuk, K.: Fast time-resolved NMR with non-uniform sampling, Prog. Nucl. Magn. Reson. Spectrosc., 116, 40–55, https://doi.org/10.1016/j.pnmrs.2019.09.003, 2020.

Grzesiek, S. and Bax, A.: The importance of not saturating H_2O in protein NMR. Application to sensitivity enhancement and NOE measurements, J. Am. Chem. Soc., 115, 12593–12594, 1993.

Hansen, D. F. and Kay, L. E.: Improved magnetization alignment schemes for spin-lock relaxation experiments, J. Biomol. NMR, 37, 245–255, https://doi.org/10.1007/s10858-006-9126-6, 2007.

Harden, B. J. and Frueh, D. P.: SARA: A software environment for the analysis of relaxation data acquired with accordion spectroscopy, J. Biomol. NMR, 58, 83–99, https://doi.org/10.1007/s10858-013-9807-x, 2014.

Hyberts, S. G., Takeuchi, K., and Wagner, G.: Poisson-gap sampling and forward maximum entropy reconstruction for enhancing the resolution and sensitivity of protein NMR data, J. Am. Chem. Soc., 132, 2145–2147, https://doi.org/10.1021/ja908004w, 2010.

Hyberts, S. G., Arthanari, H., Robson, S. A., and Wagner, G.: Perspectives in magnetic resonance: NMR in the post-FFT era, J. Magn. Reson., 241, 60–73, https://doi.org/10.1016/j.jmr.2013.11.014, 2014.

Jameson, G., Hansen, A. L., Li, D., Bruschweiler-Li, L., and Brüschweiler, R.: Extreme Nonuniform Sampling for Protein NMR Dynamics Studies in Minimal Time, J. Am. Chem. Soc., 141, 16829–16838, https://doi.org/10.1021/jacs.9b08032, 2019.

Juhlin, M., Elvander, F., Sward, J., and Jakobsson, A.: Fast Gridless Estimation of Damped Modes, in: ISPACS 2018–2018 International Symposium on Intelligent Signal Processing and Communication Systems, 27–30 November 2018, Okinawa, Japan, 346–351, 2018.

Kay, L. E., Nicholson, L. K., Delagio, F., Bax, A., and Torchia, D. A.: Pulse sequences for removal of the effects of cross correlation between dipolar and chemical-shift anisotropy relaxation mechanisms on the measurement of heteronuclear T1 and T2 values in proteins, J. Magn. Reson., 97, 359–375, 1992a.

Kay, L. E., Keifer, P., and Saarinen, T.: Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity, J. Am. Chem. Soc., 114, 10663–10665, 1992b.

Korzhnev, D. M., Tischenko, E. V., and Arseniev, A. S.: Off-resonance effects in 15N T2 CPMG measurements, J. Biomol. NMR, 17, 231–237, 2000.

Linnet, T. E. and Teilum, K.: Non-uniform sampling of NMR relaxation data, J. Biomol. NMR, 64, 165–173, https://doi.org/10.1007/s10858-016-0020-6, 2016.

Loria, J. P., Rance, M., and Palmer, A. G.: A relaxation-compensated Carr-Purcell-Meiboom-Gill sequence for characterizing chemical exchange by NMR spectroscopy, J. Am. Chem. Soc., 121, 2331–2332, 1999.

Mandel, A. M. and Palmer, A. G.: Measurement of relaxation rate constants using constant-time accordion heteronuclear NMR spectroscopy, J. Magn. Reson. Ser. A, 110, 62–72, 1994.

Månsson, A., Jakobsson, A., and Akke, M.: Multidimensional Cramér-Rao lower bound for non-uniformly sampled NMR signals, in: European Signal Processing Conference, 1–5 September 2014, Lisbon, Portugal, 974–978, 2014.

Massi, F., Johnson, E., Wang, C., Rance, M., and Palmer, A. G.: NMR R1ρ Rotating-Frame Relaxation with Weak Radio Frequency Fields, J. Am. Chem. Soc., 126, 2247–2256, https://doi.org/10.1021/ja038721w, 2004.

Mayzel, M., Ahlner, A., Lundström, P., and Orekhov, V. Y.: Measurement of protein backbone 13 CO and 15 N relaxation dispersion at high resolution, J. Biomol. NMR, 69, 1–12, https://doi.org/10.1007/s10858-017-0127-4, 2017.

Mittermaier, A. K. and Kay, L. E.: Observing biological dynamics at atomic resolution using NMR, Trends Biochem. Sci., 34, 601–611, https://doi.org/10.1016/j.tibs.2009.07.004, 2009.

Mobli, M. and Hoch, J. C.: Nonuniform sampling and non-Fourier signal processing methods in multidimensional NMR, Prog. Nucl. Magn. Reson. Spectrosc., 83, 21–41, https://doi.org/10.1016/j.pnmrs.2014.09.002, 2014.

Morris, G. A. and Freeman, R.: Enhancement of nuclear magnetic signals by polarization transfer, J. Am. Chem. Soc., 101, 760–762, 1979.

Mosteller, F. and Tukey, J. W.: Data Analysis and Regression. A Second Course in Statistics, Addison-Wesley, Reading, MA, USA, 1977.

Mulder, F. A. A., de Graaf, R. A., Kaptein, R., and Boelens, R.: An off-resonance rotating frame relaxation experiment for the investigation of macromolecular dynamics using adiabatic rotations, J. Magn. Reson., 131, 351–357, 1998.

Mulder, F. A. A., Skrynnikov, N. R., Hon, B., Dahlquist, F. W., and Kay, L. E.: Measurement of slow (microseconds-milliseconds) time scale dynamics in protein side chains by 15N relaxation dispersion NMR spectroscopy: Application to Asn and Gln residues in a cavity mutant of T4 lysozyme, J. Am. Chem. Soc., 123, 967–975, 2001.

Niklasson, M., Otten, R., Ahlner, A., Andresen, C., Schlagnitweit, J., Petzold, K., and Lundström, P.: Comprehensive analysis of NMR data using advanced line shape fitting, J. Biomol. NMR, 69, 93–99, https://doi.org/10.1007/s10858-017-0141-6, 2017.

Palmer, A. G.: NMR characterization of the dynamics of biomacromolecules, Chem. Rev., 104, 3623–3640, https://doi.org/10.1021/cr030413t, 2004.

Palmer, A. G., Cavanagh, J., Wright, P. E., and Rance, M.: Sensitivity improvement in proton-detected two-dimensional heteronuclear correlation NMR spectroscopy, J. Magn. Reson., 93, 151–170, 1991.

Palmer, A. G., Skelton, N. J., Chazin, W. J., Wright, P. E., and Rance, M.: Suppression of the effects of cross-correlation between dipolar and anisotropic chemical shift relaxation mechanisms in the measurement of spin-spin relaxation rates, Mol. Phys., 75, 699–711, 1992.

Press, W. H., Flannery, B. P., Teukolsky, S. A., and Vetterling, W. T.: Numerical Recipes. The Art of Scientific Computing, Cambridge University Press, Cambridge, UK, 1986.

Rabier, P., Kieffer, B., Koehl, P., and Lefevre, J.-F.: Fast measurement of heteronuclear relaxation: frequency-domain analysis of NMR accordion spectroscopy, Magn. Reson. Chem., 39, 447–456, 2001.

Skelton, N. J., Palmer, A. G., Akke, M., Kördel, J., Rance, M., and Chazin, W. J.: Practical aspects of two-dimensional proton-detected 15N spin relaxation measurements, J. Magn. Reson. Ser. B, 102, 253–264, 1993.

Stetz, M. A. and Wand, A. J.: Accurate determination of rates from non-uniformly sampled relaxation data, J. Biomol. NMR, 65, 157–170, https://doi.org/10.1007/s10858-016-0046-9, 2016.

Swärd, J., Adalbjörnsson, S. I., and Jakobsson, A.: High resolution sparse estimation of exponentially decaying N-dimensional signals, Signal Process., 128, 309–317, https://doi.org/10.1016/j.sigpro.2016.04.002, 2016.

Swärd, J., Elvander, F., and Jakobsson, A.: Designing sampling schemes for multi-dimensional data, Signal Process., 150, 1–10, 2018.

Urbańczyk, M., Nowakowski, M., Koźmiński, W., and Kazimierczuk, K.: Joint non-uniform sampling of all incremented time delays for quicker acquisition in protein relaxation studies, J. Biomol. NMR, 68, 155–161, https://doi.org/10.1007/s10858-017-0115-8, 2017.

Verteramo, M. L., Ignjatović, M. M., Kumar, R., Wernersson, S., Wallerstein, J., Carlström, G., Chadimova, V., Ekberg, V., Leffler, H., Zetterberg, F., Logan, D. T., Ryde, U., Akke, M., and Nilsson, U. J.: in preparation, 2021.

Wallerstein, J., Misini Ignjatovic, M., Kumar, R., Caldararu, O., Peterson, K., Leffler, H., Logan, D., Nilsson, U., Ryde, U., and Akke, M.: Entropy-Entropy Compensation Between the Conformational and Solvent Degrees of Freedom Fine-tunes Affinity in Ligand Binding to Galectin-3C, Biological Magnetic Resonance Data Bank, https://doi.org/10.13018/BMR50283, 2020.

Wallerstein, J., Ekberg, V., Misini Ignjatović, M., Kumar, R., Caldararu, O., Peterson, K., Leffler, H., Logan, D. T., Nilsson, U. J., Ryde, U., and Akke, M.: Entropy–Entropy Compensation Between the Conformational and Solvent Degrees of Freedom Fine-tunes Affinity in Ligand Binding to Galectin-3C, JACS Au, 1, 484–500, https://doi.org/10.1021/jacsau.0c00094, 2021.

Wang, A. C. and Bax, A.: Minimizing the effects of radio-frequency heating in multidimensional NMR experiments, J. Biomol. NMR, 3, 715–720, 1993.

Waudby, C. A., Burridge, C., and Christodoulou, J.: Optimal design of adaptively sampled NMR experiments for measurement of methyl group dynamics with application to a ribosome-nascent chain complex, J. Magn. Reson., 326, 106937, https://doi.org/10.1016/j.jmr.2021.106937, 2021.

Wernersson, S., Carlström, G., Jakobsson, A., and Akke, M.: Data for: Rapid measurement of heteronuclear transverse relaxation rates using non-uniformly sampled R1rho accordion experiments, Mendeley Data, V1 [data set], https://doi.org/10.17632/zyryxrgkc3.1, 2021.

Yip, G. N. B. and Zuiderweg, E. R. P.: A phase cycle scheme that significantly suppresses offset-dependent artifacts in the R-2-CPMG N-15 relaxation experiment, J. Magn. Reson., 171, 25–36, 2004.

Paper V

An accordion NUS experiment measuring transverse cross-correlated dipole/CSA relaxation

Sven Wernersson[1], Göran Carlström[2], Andreas Jakobsson[3], David Fushman[4], Mikael Akke[1*]

[1] Division of Biophysical Chemistry, Center for Molecular Protein Science, Department of Chemistry, Lund University, Sweden; [2] Centre for Analysis and Synthesis, Department of Chemistry, Lund University, Sweden. [3]Department of Mathematical Statistics, Lund University, P.O. Box 118, 22100 Lund, Sweden

ABSTRACT: This article describes a novel heteronuclear NMR relaxation experiment designed for time-efficient measurements of the transverse dipole/CSA cross-relaxation rate η_{xy}. This is done by combining non-uniform sampling (NUS) with the use of the accordion method to reduce the dimensionality of the experiment. The accordion method is itself combined with spin state selective methods, as the difference in relaxation between spin states naturally mimic the different accordion experiments typically used for accordion NMR relaxation measurements. The new experiment is used to measure η_{xy} for the ^{15}N nuclei in a fully protonated Ubiquitin sample. The estimated η_{xy} rates are favorably compared with η_{xy} rates estimated from conventional NMR relaxation experiments, which validates the approach taken with new experiment.

NMR relaxation experiments are a powerful tool to study protein dynamics, and a multitude of experiments exists that explores different relaxation phenomena (1). Cross-correlated dipole/CSA relaxation in spin ½ nuclei is active in the absence of decoupling (2, 3). This effect is commonly utilized for TROSY detection (4) and CRINEPT type magnetization transfer (5), typically for larger molecules of study. Less common is the direct measurement of the relaxation rates for this mechanism, which is the topic at hand. The relevant transverse and longitudinal rate constants (η_{xy}, η_z) can be described for an ^{15}N-^{1}H spin pair as (6):

$$\eta_{xy} = \frac{\sqrt{3}}{6} cdP_2(\cos\theta)[4J(0) + J(\omega_N)]$$

$$\eta_z = \sqrt{3}cdP_2(\cos\theta)J(\omega_N)$$

Where $c = \gamma_N B_0 \Delta\sigma/\sqrt{3}$ is the CSA coupling constant, $d = (\mu_0 h \gamma_H \gamma_N)/(8\pi^2 r_{NH}^3)$ is the dipolar coupling constant, γ_H and γ_N are the gyromagnetic ratios of hydrogen and nitrogen, $\Delta\sigma$ is the chemical shift anisotropy, μ_0 is the permeability of free space, h is planks constant, r_{NH} is the ^{15}N-^{1}H bond length, $P_2 = (3x^2 - 1)/2$ is the second degree Legendre polynomial, and θ is the angle between the principal axes of the dipolar and CSA tensors. An interesting property of the transverse relaxation rate η_{xy} is the absence of any chemical exchange (R_{ex}) contributions. This makes measured η_{xy}, η_z rates useful for unambiguous estimates of the spectral density at zero frequency (J(0)), the chemical shift anisotropy ($\Delta\sigma$), and the rotational diffusion tensor (6). In addition, the θ angle, seemingly a nuisance parameter, also appears in the R_1, R_2 auto-relaxation rates (7). For these relaxation rates, the effect of θ scales with B_0 and anisotropy in the rotational diffusion tensor. This effect is typically ignored as it is negligible when measured at historically available B_0 strengths, but it can be significant at the strongest commercially available field strengths of around 21 T and beyond.

Two different methods are typically used for designing experiments measuring η_{xy}, η_z. The first method is to use the fact that cross-correlated dipolar/CSA relaxation causes interconversion between in-phase and anti-phase magnetization. This interconversion rate provides an estimate of the relaxation rate (6, 8–10). The origin of this method is an ηxy experiment designed by Tjandra and Co-workers (11). This experiment was later improved and expanded upon to allow for η_z measurements by Kroenke and co-workers (6). The basis of both experiments is the measurement of the auto-relaxation of in-phase magnetization and the cross transfer to antiphase relaxation via cross-correlated relaxation. Later, Pelupessy and co-workers published the symmetrical reconversion approach (9, 10), which gains improved accuracy via the additional measurements of the auto-relaxation of antiphase magnetization and the cross transfer of antiphase to in-phase magnetization via cross-correlated relaxation. The basis of the other method is to directly measure the difference in auto-relaxation between the peak multiplets in coupled spectra, as the sign of the cross-correlated dipolar/CSA relaxation contribution follows the spin state of the measured magnetization (±½) (12–16). The simplest experiment following this spin state selective approach is to measure the difference in peak height between the coupled peaks in a CT HSQC spectrum (3). This has the clear disadvantage of significant spectral overlap, and possible errors from measuring a single relaxation time point. Hall and co-workers designed an experiment that measures pair of spectra as in-phase/antiphase (IP/AP) (12). This allows for selective addition and subtraction to achieve spin state

selection with high sensitivity, but notable spectral artefacts. The same result with reduced artifacts was later achieved via the use of S3CT filtering (14).

A challenge to any measurement of η_{xy}, η_z in protonated samples is the presence of spin diffusion. This refers to random spin flips, which occur spontaneously but are mainly induced by interactions between different spin (17). For measurements on ^{15}N-^{1}H spin pairs, this affects the ^{1}H-spin as it interacts with other remote ^{1}H spin. This increases the auto-relaxation rate of anti-phase magnetization by an amount equivalent to the ^{1}H longitudinal auto-relaxation rate ($R_{1,H}$) in a manner similar to chemical exchange, as the spin states experience different chemical shifts (2). The increased accuracy of the symmetrical reconversion approach stems in part from reduced sensitivity to this difference in the auto-relaxation rates between in- and antiphase magnetization (9). For spin state selective experiments this effect averages the auto-relaxation rate of the spin states, thus reducing the apparent cross-correlated relaxation rate. For transverse magnetization this averaging is reduced as $R_{1,H}$ is smaller than the auto-relaxation rate R_2, causing the magnetization to decay faster than it interconverts (18, 19). In general the effect of spin diffusion can be refocused for transverse magnetization by implementing spin echoes for every $1/4J_{NH}$ time period of the relaxation period (9).

In this article we present a novel spin state selective accordion NUS experiment measuring the transverse cross-correlated dipole/CSA relaxation rate η_{xy}. The experiment, outlined in figure 1, is structured in a similar fashion to our earlier accordion NUS experiments measuring R_1 and $R_{1\rho}$ respectively (20, 21).

MATERIAL AND METHODS

NMR sample preparation. Lyophilized uniformly ^{15}N-enriched human ubiquitin was purchased from Sigma Aldrich. The protein was dissolved in a sample buffer consisting of 20mM sodium phosphate at pH 6.8 in 90% H_2O and 10% D_2O (v/v). The final protein concentration in the sample was estimated via UV absorbance spectroscopy to approximately 200 μM.

NMR relaxation experiments. Conventional symmetrical reconversion (9) and accordion spin state selective η_{xy} experiments were acquired with uniform sampling (US) on a Bruker NEO 600 MHz instrument equipped with a 5 mm HPCN QCI cryo-probe. The ^{1}H dimension was acquired with a spectral width of 9615 Hz, sampled over 2306 complex data points, in all experiments. The ^{15}N dimension was acquired with a spectral width of 2129 Hz, sampled over 64 increments in the symmetrical reconversion experiments and with 64 or 128 increments in the accordion experiments. The accordion experiments were performed with the magnetization at the start of the relaxation period prepared as both $H^{\alpha}N_x$ ('forward'), and $H^{\beta}N_x$ ('reverse'), acquired in an interleaved fashion. Four symmetrical reconversion experiments (I-IV) were acquired interleaved with the t_1-increments for each conventional dataset. Three accordion datasets using an accordion scaling factor κ = 2 were acquired with a number of scans (ns), increments (ni) equal to (64, 64), (32, 128) and (64, 128). Two accordion datasets using an accordion scaling factor κ = 4 were acquired using (ns, ni) equal to (32, 128) and (64, 128). One accordion dataset using accordion scaling factors κ(forward, reverse) = (2, 4) were acquired using (ns, ni) = (64, 128). Three symmetrical reconversion datasets were acquired for relaxation periods of 40, 60, 80 ms. The symmetrical reconversion experiments utilized ^{1}H GARP decoupling during the relaxation period.

Nonuniform sampling schemes. The uniformly sampled accordion η_{xy} dataset acquired using κ = 2, 128 increments and 64 scans was used to generate NUS schemes. In-house MATLAB scripts were used to implement Single-column CRLB optimized (col-opt) (22, 23) and sine-weighted Poisson-gap (24) sampling schemes. Poisson-gap sampling schemes were generated by randomly varying the argument of the sinusoidal weighting function between 0 and $\pi/2$; see (24). The best sampling scheme was identified as the one having the lowest sum of the CRLB calculated over all columns containing peaks. In the case of the col-opt approach, the selection was made among 97 different single-column CRLB optimized schemes corresponding to each slice of the interferogram containing protein signals, whereas in the case of Poisson-gap sampling 1000 different schemes were compared. The sampling scheme was optimized individually for each of the forward and reverse accordion experiments. In each case, we generated individual datasets sampled with N = 16, 20, 24, 28, 32, 36, 40, 44, 48, 52, 56, 60, 64, 68, 72, 76, 80, 84, 88, 92, 96, 100, 104, 108, 112, 116, 120 and 124 increments in the indirect dimension. The NUS datasets resulting from the different sampling schemes were subsequently reconstructed using the DSURE algorithm.

Data reconstruction, processing, and analysis. Symmetrical reconversion datasets were processed using NMRPipe (25), with forward linear prediction to double the number of data points, cosine-squared window apodization, zero-filling to twice the size rounded to the nearest power of two, polynomial baseline correction, and solvent filtering. The η_{xy} relaxation constant was estimated from the symmetrical reconversion datasets by integrating the peak volumes using PINT (26, 27), followed by fitting the volumes of experiments ($V_I(T)$-$V_{IV}(T)$) using in-house MATLAB scripts to the hyperbolic function (9):

$$\left|\eta_{xy}\right| = \frac{\tanh^{-1}\left(\sqrt{\frac{V_{II}(T)\cdot V_{III}(T)}{V_I(T)\cdot V_{IV}(T)}}\right)}{T}$$

Where T is the length of the relaxation period. The uncertainty (E_{SR}) in the estimated η_{xy} rates from the symmetrical reconversion experiments were propagated using the expression:

$$E_{SR} = \frac{\sigma}{\alpha\sqrt{2}}\sqrt{1+\beta^2\left(\alpha^2+\alpha^{-2}\right)/2}$$

Where σ is the volume uncertainty, α is a scaling factor accounting for relaxation in the conversion periods flanking the relaxation period, and β is the fraction of magnetization transferred by η_{xy}. For the symmetrical reconversion experiments, we found the propagated uncertainties to be unreasonably small ($E_{SR}/\eta_{xy} \approx 0.2\%$). We relate this issue to PINT underestimating the volume uncertainty σ

for NMR experiments acquired using cryo-probes. For this reason, σ was scaled using the reduced χ^2 value of the volume integration before estimating the uncertainty in η_{xy}. The accordion datasets were processed in the direct (t_2) dimension using NMRpipe, with cosine-squared window apodization, zero-filling to twice the size rounded to the nearest power of two, polynomial baseline correction, and solvent filtering. The processed datasets were then analyzed in MATLAB using the DSURE algorithm (28). NUS reconstruction was performed on individual t_1 interferograms using DSURE, as described previously (23).

Statistical analysis. Different metrics were used to compare datasets. The root mean square deviation (RMSD) between datasets was calculated pairwise over all residues (N_{res}), RMSD = $[(\sum_i(x_i - y_i)^2/N_{res})]^{1/2}$. The mean relative uncertainty of a given dataset is the mean, calculated over all residues, of the individual uncertainty in x_i (σ_{xi}; one standard deviation, as estimated by the DSURE algorithm) divided by x_i, MRU = $(\sum_i \sigma_{xi}/x_i)/N_{res}$). The relative difference between datasets x and y are defined for a given residue i as $\Delta_{rel} = 2(x_i - y_i)/(x_i + y_i)$. The root mean square error (RMSE) between datapoints x and y were calculated as RMSE = |x - y|. The number of neighboring 1H-nuclei to the ^{1}H of each ^{1}H-^{15}N spin pair was calculated in Matlab using the NMR structure with PDB-ID: 2KOX (29).

RESULTS AND DISCUSSION

Pulse sequence design. We have previously published two CT Accordion (20, 21) NUS experiments for the measurement of the R_1 and $R_{1\rho}$ relaxation rates (23, 30). The key to this approach is the use of the DSURE (damped super-resolution estimator) algorithm (28) to reconstruct and estimate signal linewidths. The experiment described here is designed on the same format for the measurement of ηxy.

There are several things to consider for such an experiment. Although the DSURE algorithm is expected to resolve multi-exponential relaxation in principle, this is completely untested and could prove quite complicated to verify. In addition, accordion type experiments require arbitrary increments of the relaxation period. This makes it highly impractical to include CPMG pulse trains in the relaxation period, a problem we discuss in a previous article (30). It also prohibits refocusing of the t_1-evolution during the relaxation period, as it would require relaxation period increments of $1/4J_{NH}$ (10). Finally, for measurements of η_{xy} the use of spin locks during the relaxation period is complicated, as off-resonance magnetization can be expected to exhibit multi-exponential relaxation in protonated samples (31).

To fulfill these criteria, we designed the experiment in a spin state selective fashion, using the S3E filter of Meissner and coworkers (32). For these types of experiments, the build-up of multi-exponential relaxation is limited by the fact that the build-up rate ($\approx R_{1,H}$) is smaller than the auto-relaxation rate R_2 (18, 19). This type of experiment has the additional benefit that the difference in the relaxation rates of spin-up ($H^{\alpha}N_x$) and spin-down ($H^{\beta}N_x$) relaxation corresponds in a natural way to that of the forward and reverse experiments used for accordion NMR spectroscopy (21), as the $H^{\alpha}N_x$ magnetization experience additional relaxation equal to η_{xy}, and the $H^{\beta}N_x$ magnetization experience an equivalent reduction in relaxation. This allows the estimation of η_{xy} from the expression:

$$\eta_{xy} = \frac{R_2^{\alpha} - R_2^{\beta}}{2\kappa}$$

Where R_2^{α} is the auto-relaxation rate of magnetization state $H^{\alpha}N_x$, R_2^{β} is the auto-relaxation rate of magnetization state $H^{\beta}N_x$ and κ is the accordion scaling factor $\kappa = \Delta t_r/\Delta t_1$ describing the incrementation (Δt_r) of the relaxation period t_r with each increment of t_1. Given the unavoidable t_1 evolution during relaxation, these periods have been combined into the element $\tau_1 - 180(^{15}N)-\tau_2$, with $\tau_1 = t_r - t_1$ and $\tau_2 = t_r + t_1$. This can be contrasted with the experiment of Liu and Prestegard (19). For that η_{xy} experiment, as t_1-evolution occurs during both the relaxation and t_1-period, the chemical shift is different for each length of the relaxation period. This is not tenable for an accordion experiment, as it would lead to highly convoluted lineshapes. Our Accordion η_{xy} experiment is presented in figure 1. The preparation step consists of an INEPT transfer and a spin state selective filter. This is followed by the combined accordion relaxation and t_1 period. This period is followed by a PEP-PFG (33–35) back transfer step, and TROSY readout. This experiment is repeated for magnetization with the ^{1}H spin selected to generate either +1/2 ($H^{\alpha}N_x$) or -1/2 ($H^{\beta}N_x$) magnetization.

Comparison of DSURE-modeled accordion ηxy relaxation data and conventional relaxation data. We compared the new accordion pulse sequence against the conventional pulse sequence of Pelupessy and co-workers (9), by performing measurements on fully protonated human Ubiquitin. The small size of this protein enables conventional NMR relaxation experiments to be performed with 64 increments in the indirect (t_1) dimension. This is in contrast to our earlier accordion NMR experiments performed with 128 t_1-increments (23, 30). Therefore, panel A of figure 2 is a comparison of three different US accordion datasets all acquired using κ = 2 but using differing number of increments (ni) and scans (ns). The dataset in red uses (ni, ns) = (64, 64), the dataset in black (ni, ns) = (128, 64) and the dataset in green (ni, ns) = (128, 32). In general, the datasets with ni=128 are more similar, with an RMSD of 0.33. In contrast, the dataset with ni=64 has an RMSD of 0.65 to the dataset with the black dataset and an RMSD of 0.72 with the green dataset. Similarly, the red dataset had the highest relative mean uncertainty of 12.79%, followed by the green dataset with 8.88% and the black dataset with 5.28%. This indicates that the number of increments in general is of greater importance for accordion datasets than the number of scans. For the sake of better statistical analysis of NUS sampling in the next section, further comparisons will be made using accordion datasets using (ni, ns) = (128, 64).

Panel B of figure 2 compares accordion datasets measured using κ = 2 (black), κ = 4 (green) and κ_f = 2, κ_r = 4 (red), where κ_f, κ_r refers to the accordion scaling factors for the forward and reverse accordion experiment respectively. The possibility of using different κ scaling factors for

the forward and reverse accordion experiments is important for larger proteins, where the differential relaxation is expected to be highly significant. In the case of ubiquitin, the relative mean uncertainty of 5.17% for the dataset using (κ_f = 2, κ_r = 4) is only slightly lower than for the black dataset with κ = 2. In addition, the three datasets have rather large differences in η_{xy} values, with an RMSD of 0.56 between the κ = 2 and κ = 4 datasets, RMSD = 0.70 between the κ = 2 and κ = 2, 4 datasets, and RMSD = 0.98 between the κ = 4 and κ = 2, 4 datasets. Panel (C) compares the estimated η_{xy} rates of the accordion dataset measured using κ = 2 (black), and the three conventional datasets measured for relaxation period of 40 ms (red), 60 ms (green) and 80 ms (blue). The RMSD between the accordion dataset and the conventional 60 ms dataset is slightly smaller (0.37) than between the accordion dataset and the conventional 40, 80 ms datasets (0.43, 0.42). However, this is larger than the RMSD between the conventional datasets; the RMSD between the 40 and 60 ms datasets is 0.21, between the 40 and 80 ms datasets 0.24, and between the 60 and 80 ms datasets 0.18.

Panels (D-F) compares the accordion dataset with κ = 2 to the mean of the conventional datasets, where (D) plots each dataset against residue number, (E) is a plot of covariance between the estimated η_{xy} rates of each dataset, and (F) is a histogram of the differences between the datasets, plotted together with a fitted normal distribution. Covariance and histogram plots between the accordion dataset with κ = 2 and the individual conventional datasets are shown in SI figure 1, and the parameters of all fitted normal distributions are shown in SI table 1. For all fitted distributions, the means differ slightly from zero, but this difference is generally within the uncertainty of the datasets.

Comparison of non-uniformly sampled and uniformly sampled accordion η_{xy} relaxation data. We then performed NUS sampling of the accordion dataset with κ = 2. Figure 3 shows the mean RMSD against the US dataset and the mean relative uncertainty for different number of sampled points using sampling schemes optimized via Poisson-gap (PG) and column optimized (col-opt) sampling. In general, the down-sampled datasets for both sampling schemes show similar characteristics as our previously reported accordion NUS experiments down to about 50% sampling rates, below which the RMSD shows larger fluctuations. PG-sampling generally show somewhat better performance below 50% NUS, which is in line with our previous studies (30).

Spectral characteristics affecting accuracy and precision of accordion-NUS η_{xy} relaxation rate constants. Spectra measured with the accordion η_{xy} sequence contain small signal artifacts stemming from the unwanted spin multiplet (see example in SI figure 2). This is a known issue with TROSY sequences that relates to an reverse INEPT generated imbalance in the hypercomplex t_1 pairs due to differential relaxation (see (19) for further discussion). The presence of this artifact is not expected to increase cross-relaxation between the multiplets during the relaxation period, which is tested below. It does present a practical issue in estimating lineshapes. The obvious effect is an increase in spectral crowding, similar to any experiment employing TROSY pulse sequences. The other effect is an effective increase in the uncertainties of the estimated η_{xy} values for all peaks. This is due to that the uncertainties are related to the lineshape fit of the entire interferogram. SI figure 2 shows an example of the isolated peak from residue Ile13, fitted with and without the TROSY artifact. In both cases the estimated ηxy rate is 4.1 s-1. However, for the fit without the TROSY artifact the estimated uncertainty is 0.2 s-1, while for the fit with the TROSY artifact the estimated uncertainty is 0.1 s-1. This complicates the interpretation of the estimated uncertainties. Another potential issue that should be accounted for is spin diffusion. To investigate this, SI figure 3 shows the RSE between the three conventional datasets and the accordion dataset (κ = 2) used for downsampling plotted against the number of ^{1}H-neighbours within 5 Å of the ^{1}H-nuclei of the ^{1}H-^{15}N spin pair. No apparent relation can be found between ^{1}H-nuclei density and the RSE from this figure.

Given the presence of unwanted peak signals, we then investigated potential spectral crowding effects. Figure 4 shows RSE values calculated between different datasets, plotted alternatively against the number of signals in a given interferogram (left side) and nearby interferograms (right side). The top row plots the RSE between US accordion (κ = 2) data and the mean of the conventional datasets, the middle row plots the RSE between US and NUS (ni=88) accordion (κ = 2) data and the bottom row plots the RSE between US and NUS (ni=48) accordion (κ = 2) data. The NUS dataset with 88 sampled points was selected as a representative of the well behaved (<50% downsampling) regime, and the dataset with 48 sampled points and representative of the poorly behaved (>50% downsampling) regime. This overview shows no obvious link between spectral crowding and decreased precision for any of the compared datasets.

Finally, we discuss potential improvements to experimental design. A potentially very significant gain in sensitivity could be achieved if a mean to refocus chemical exchange effects could be implemented. As discussed above, for $R_{1\rho}$ based spin locks this would require deuteration or more advanced processing methods to account for multi-exponential relaxation. For CPMG pulse trains in the relaxation period, this would likely require some sort of non-uniform incrementation of the relaxation period. The other obvious avenue for improvement is the suppression of the unwanted signal components.

Conclusions. We have described a novel Accordion NUS experiment for measurement of the η_{xy} relaxation rate constant. The estimated η_{xy} rates shows a small but statististically insignificant difference to η_{xy} rates estimated from conventional NMR experiments. NUS sampling of this dataset could be reliably performed down to around 50% NUS sampling. Several ideas on potential improvements are presented.

ASSOCIATED CONTENT

Supporting Information

Supporting information is available as a separate file, containing; covariance and histogram plots of the differences between US sampled accordion and conventionally sampled data, normal distribution fitting parameters to said difference, fitted accordion interferogram and spectra of both spin states, and estimates of spin diffusion.

AUTHOR INFORMATION

Corresponding Author

* Mikael Akke (mikael.akke@bpc.lu.se)

REFERENCES

1. A. G. Palmer, NMR Characterization of the Dynamics of Biomacromolecules. *Chem. Rev.* **104**, 3623–3640 (2004).
2. M. Goldman, Interference effects in the relaxation of a pair of unlike spin- 1 2 nuclei. *J. Magn. Reson.* **60**, 437–452 (1984).
3. B. Brutscher, Principles and applications of cross-correlated relaxation in biomolecules. *Concepts Magn. Reson.* **12**, 207–229 (2000).
4. K. Pervushin, R. Riek, G. Wider, K. Wüthrich, Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. *Proc. Natl. Acad. Sci. U. S. A.* **94**, 12366–12371 (1997).
5. R. Riek, G. Wider, K. Pervushin, K. Wüthrich, Polarization transfer by cross-correlated relaxation in solution NMR with very large molecules. *Proc. Natl. Acad. Sci.* **96**, 4918–4923 (1999).
6. C. D. Kroenke, J. P. Loria, L. K. Lee, M. Rance, A. G. Palmer, Longitudinal and transverse 1H-15N dipolar/15N chemical shift anisotropy relaxation interference: Unambiguous determination of rotational diffusion tensors and chemical exchange effects in biological macromolecules. *J. Am. Chem. Soc.* **120**, 7905–7915 (1998).
7. D. Fushman, D. Cowburn, The effect of noncollinearity of 15N-1H dipolar and 15N CSA tensors and rotational anisotropy on 15N relaxation, CSA/dipolar cross correlation, and TROSY. *J. Biomol. NMR* **13**, 139–47 (1999).
8. D. Fushman, N. Tjandra, D. Cowburn, Direct measurement of 15N chemical shift anisotropy in solution. *J. Am. Chem. Soc.* **120**, 10947–10952 (1998).
9. P. Pelupessy, G. M. Espallargas, G. Bodenhausen, Symmetrical reconversion: Measuring cross-correlation rates with enhanced accuracy. *J. Magn. Reson.* **161**, 258–264 (2003).
10. P. Pelupessy, F. Ferrage, G. Bodenhausen, Accurate measurement of longitudinal cross-relaxation rates in nuclear magnetic resonance. *J. Chem. Phys.* **126** (2007).
11. N. Tjandra, A. Szabo, A. Bax, Protein backbone dynamics and 15N chemical shift anisotropy from quantitative measurement of relaxation interference effects. *J. Am. Chem. Soc.* **118**, 6986–6991 (1996).
12. J. B. Hall, K. T. Dayie, D. Fushman, Direct measurement of the 15N CSA/dipolar relaxation interference from coupled HSQC spectra. *J. Biomol. NMR* **26**, 181–186 (2003).
13. J. B. Hall, D. Fushman, Direct measurement of the transverse and longitudinal 15N chemical shift anisotropy-dipolar cross-correlation rate constants using 1H-coupled HSQC spectra. *Magn. Reson. Chem.* **41**, 837–842 (2003).
14. P. R. Vasos, J. B. Hall, D. Fushman, Spin-state selection for increased confidence in cross-correlation rates measurements. *J. Biomol. NMR* **31**, 149–154 (2005).
15. S. Bouguet-Bonnet, P. Mutzenhardt, D. Canet, Measurement of 15N csa/dipolar cross-correlation rates by means of Spin State Selective experiments. *J. Biomol. NMR* **30**, 133–142 (2004).
16. D. S. Weaver, E. R. P. Zuiderweg, Z κ: AA transverse relaxation optimized spectroscopy NMR experiment measuring longitudinal relaxation interference. *J. Chem. Phys.* **128** (2008).
17. D. S. Weaver, E. R. P. Zuiderweg, Protein proton-proton dynamics from amide proton spin flip rates. *J. Biomol. NMR* **45**, 99–119 (2009).
18. L. E. Kay, L. K. Nicholson, F. Delaglio, A. Bax, D. . Torchia, Pulse sequences for removal of the effects of cross correlation between dipolar and chemical-shift anisotropy relaxation mechanisms on the measurement of heteronuclear T1 and T2 values in proteins. *J. Magn. Reson.* **97**, 359–375 (1992).
19. Y. Liu, J. H. Prestegard, Direct measurement of dipole-dipole/CSA cross-correlated relaxation by a constant-time experiment. *J. Magn. Reson.* **193**, 23–31 (2008).
20. G. Bodenhausen, R. R. Ernst, The accordion experiment, a simple approach to three-dimensional NMR spectroscopy. *J. Magn. Reson.* **45**, 367–373 (1981).
21. A. M. Mandel, A. G. Palmer, Measurement of Relaxation-Rate Constants Using Constant-Time Accordion NMR Spectroscopy. *J. Magn. Reson. Ser. A* **110**, 62–72 (1994).
22. J. Swärd, F. Elvander, A. Jakobsson, Designing sampling schemes for multi-dimensional data. *Signal Processing* **150**, 1–10 (2018).
23. G. Carlström, F. Elvander, J. Swärd, A. Jakobsson, M. Akke, Rapid NMR Relaxation Measurements Using Optimal Nonuniform Sampling of Multidimensional Accordion Data Analyzed by a Sparse Reconstruction Method. *J. Phys. Chem. A* **123**, 5718–5723 (2019).
24. S. G. Hyberts, K. Takeuchi, G. Wagner, Poisson-gap sampling and forward maximum entropy reconstruction for enhancing the resolution and sensitivity of protein NMR data. *J. Am. Chem. Soc.* **132**, 2145–2147 (2010).
25. F. Delaglio, *et al.*, NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6**, 277–293 (1995).
26. A. Ahlner, M. Carlsson, B. H. Jonsson, P. Lundström, PINT: A software for integration of peak volumes and extraction of relaxation rates. *J. Biomol. NMR* **56**, 191–202 (2013).
27. M. Niklasson, *et al.*, Comprehensive analysis of NMR data using advanced line shape fitting. *J. Biomol. NMR* **69**, 93–99 (2017).
28. M. Juhlin, F. Elvander, J. Sward, A. Jakobsson, Fast Gridless Estimation of Damped Modes. *ISPACS 2018 - 2018 Int. Symp. Intell. Signal Process. Commun. Syst.*, 346–351 (2018).
29. R. B. Fenwick, *et al.*, Weak long-range correlated motions in a surface patch of ubiquitin involved in molecular recognition. *J. Am. Chem. Soc.* **133**, 10336–10339 (2011).
30. S. Wernersson, G. Carlström, A. Jakobsson, M. Akke, Rapid measurement of heteronuclear transverse relaxation rates using non-uniformly sampled <i>R</i><sub>1<i>ρ</i></sub > accordion experiments. *Magn. Reson.* **2**, 571–587 (2021).
31. P. Kadeřávek, *et al.*, Cross-correlated relaxation measurements under adiabatic sweeps: Determination of local order in proteins. *J. Biomol. NMR* **63**, 353–365 (2015).
32. A. Meissner, J. Duus, O. W. Sørensen, Spin-State-Selective Excitation. Application for E.COSY-Type Measurement of JHH Coupling Constants. *J. Magn. Reson.* **128**, 92–97 (1997).
33. A. G. Palmer, J. Cavanagh, P. E. Wright, M. Rance, Sensitivity improvement in proton-detected two-dimensional heteronuclear correlation NMR spectroscopy. *J. Magn. Reson.* **93**, 151–170 (1991).
34. M. Rance, J. P. Loria, A. G. Palmer, Sensitivity Improvement of Transverse Relaxation-Optimized Spectroscopy. *J. Magn. Reson.* **136**, 92–101 (1999).
35. L. E. Kay, P. Keifer, T. Saarinen, Pure Absorption Gradient Enhanced Heteronuclear Single Quantum Correlation Spectroscopy with Improved Sensitivity. *J. Am. Chem. Soc.* **114**, 10663–10665 (1992).

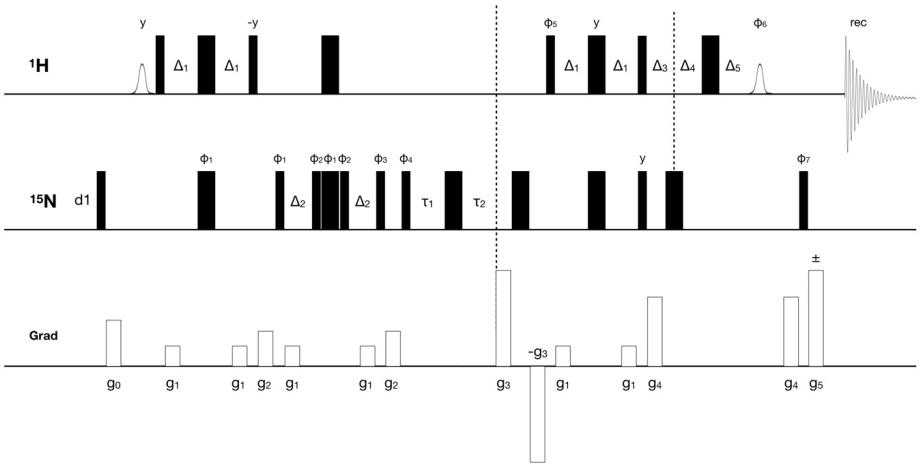


Figure 1: Pulse sequence for the ^{1}H-detected accordion-NUS ^{15}N spin state selective η_{xy} experiment. Thin (thick) black bars correspond to 90° (180°) non-selective pulses. All pulses have phase x, unless otherwise indicated. The open bell-shaped pulse at the beginning and end of the sequence is a selective 90° pulse on the water resonance (21). INEPT polarization transfer steps (22, 23) use $\Delta_1 = 1/(4J_{HN})$, where J_{HN} is the one-bond scalar coupling constant. The spin selection element is based on the S3E element of Meissner and coworkers (24), and uses $\Delta_2 = 1/(8J_{HN})$. For the TROSY back-transfer step, $\Delta_3 + \Delta_4 = \Delta_5 = 1/(4J_{HN})$. The combined relaxation and t_1-period contain delays $\tau_1 = n{\cdot}\Delta t_1(\kappa - 1)$ and $\tau_2 = n{\cdot}\Delta t_1(\kappa + 1)$, where n is the sampled point number in t_1, Δt_1 is the length of the t_1 incrementation step, and κ is the accordion scaling factor. The phase cycle is ϕ_1 = (x, -x), ϕ_2 = (-y, y), ϕ_4 = (2(-y), 2(y)), ϕ_7 = -x. For the forward ($H^{\alpha}N_x$) experiment ϕ_3 = (4(5π/4), 4(π/4)), ϕ_5 = y, ϕ_6 = -y and for the reverse ($H^{\beta}N_x$) experiment ϕ_3 = (4(7π/4), 4(3π/4)), ϕ_5 = -y, ϕ_6 = y. The receiver phase cycle is rec = (x, –x, –x, x, x, –x, –x, x). Since these phases correspond to rotations around a right-handed coordinate system, spectrometer specific changes might be required (25). Gradient enhanced PEP polarization transfer (26, 27) is achieved by acquiring a second data set with inversion of phases ϕ_5, ϕ_7 and gradient g_5. The gradients g_3 and g_5 are used for coherence selection. The phase φ_3 and the receiver phase are inverted for each t_1 increment. The gradient times and levels are g0: 1 ms, 8.9 G cm^{-1}; g1: 0.5 ms, 6 G cm^{-1}; g2: 1 ms, 10 G cm^{-1}; g3: 0.9 ms, 28 G cm^{-1}; g4: 1 ms, 20 G cm^{-1}; g5: 0.182 ms, 28 G cm^{-1}.

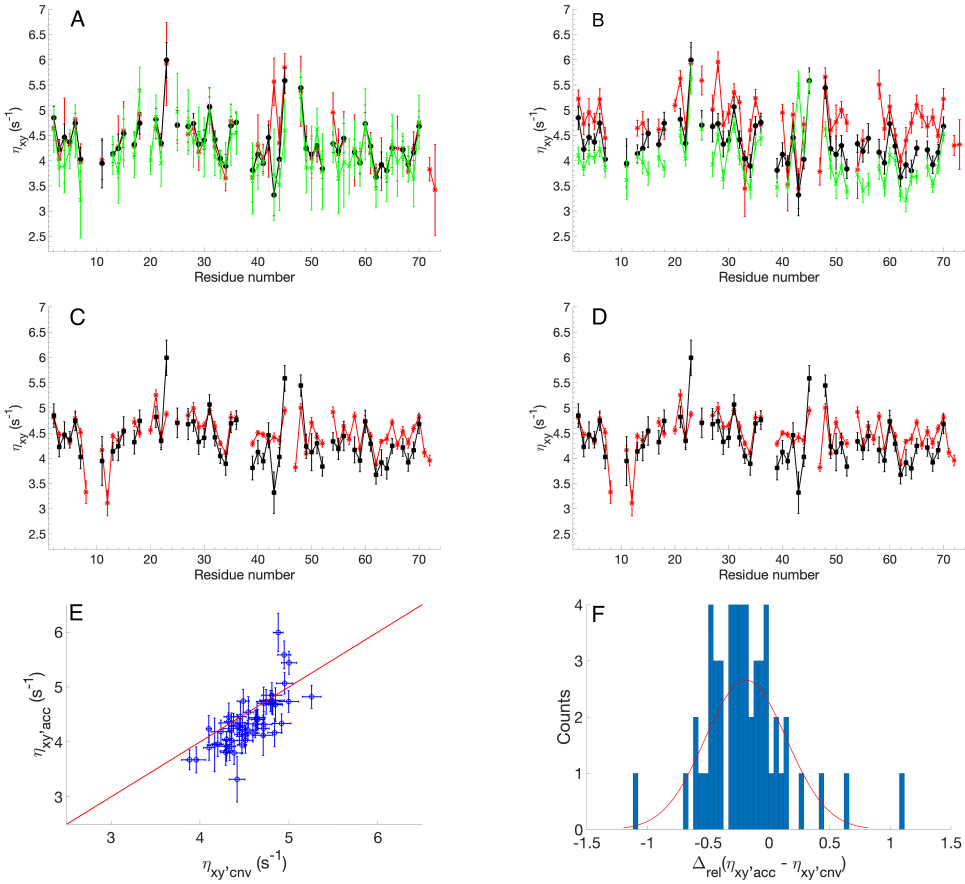


Figure 2: Comparison between US accordion and conventional ηxy data. (A) Plot of US accordion datasets measured using κ = 2, with number of increments (ni), number of scans (ns) equal to red: (64, 64), black: (128, 64), green: (128, 32). (B): Plot of US accordion datasets measured using κ = 2 (black), κ = 4 (green), κf = 2 and κr = 4 (red). (C) Comparison of US accordion dataset measured using κ = 2 (black) with conventionally measured datasets using a relaxation period of 40 ms (red), 60 ms (green), 80 ms (blue). (D) Comparison of US accordion dataset measured using κ = 2 (black) with the mean (red) of the conventional datasets plotted in (C). (E, F) The datasets in (D) compared in a covariance plot (E) and their difference in a histogram plot (F). The histogram plot is fitted with a normal distribution presented in SI table 1.

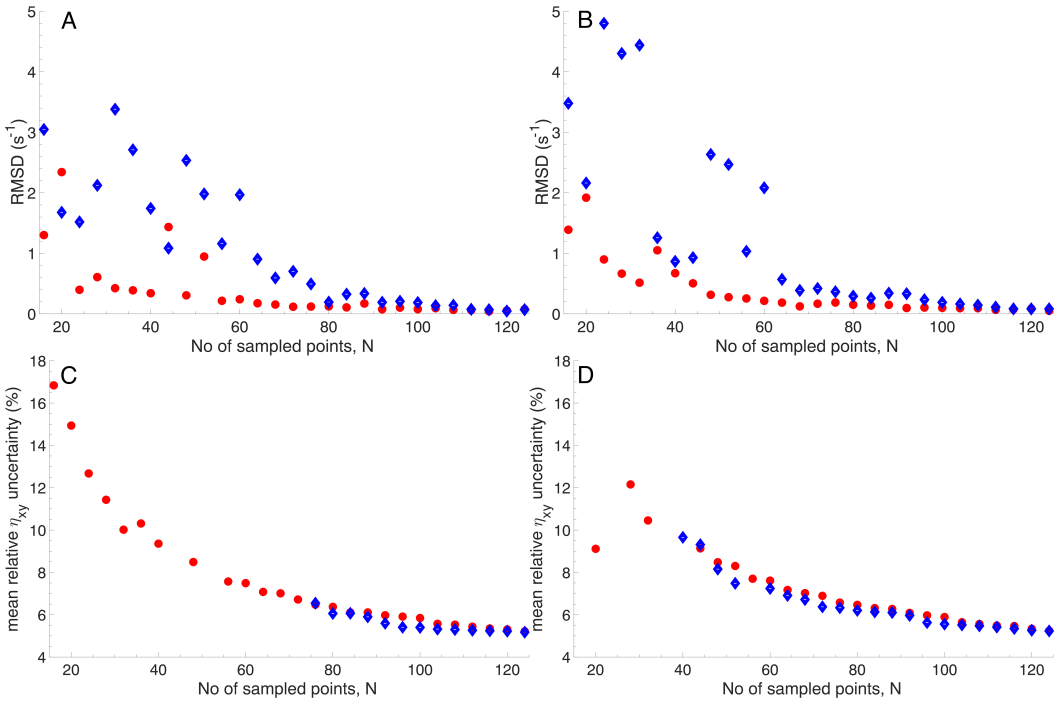


Figure 3: Analysis of NUS datasets created using Poisson-gap (PG) schemes (red) and column-optimized (col-opt) schemes (blue). (A, B) Mean RMSD of estimated η_{xy} rates against the US dataset. (C, D) Mean relative uncertainty for a given number of sampled points in the indirect dimension. Left-side panels were created using NUS schemes optimized against the forward experiment, Right-side panels were created using NUS schemes individually optimized for the forward, reverse experiment.

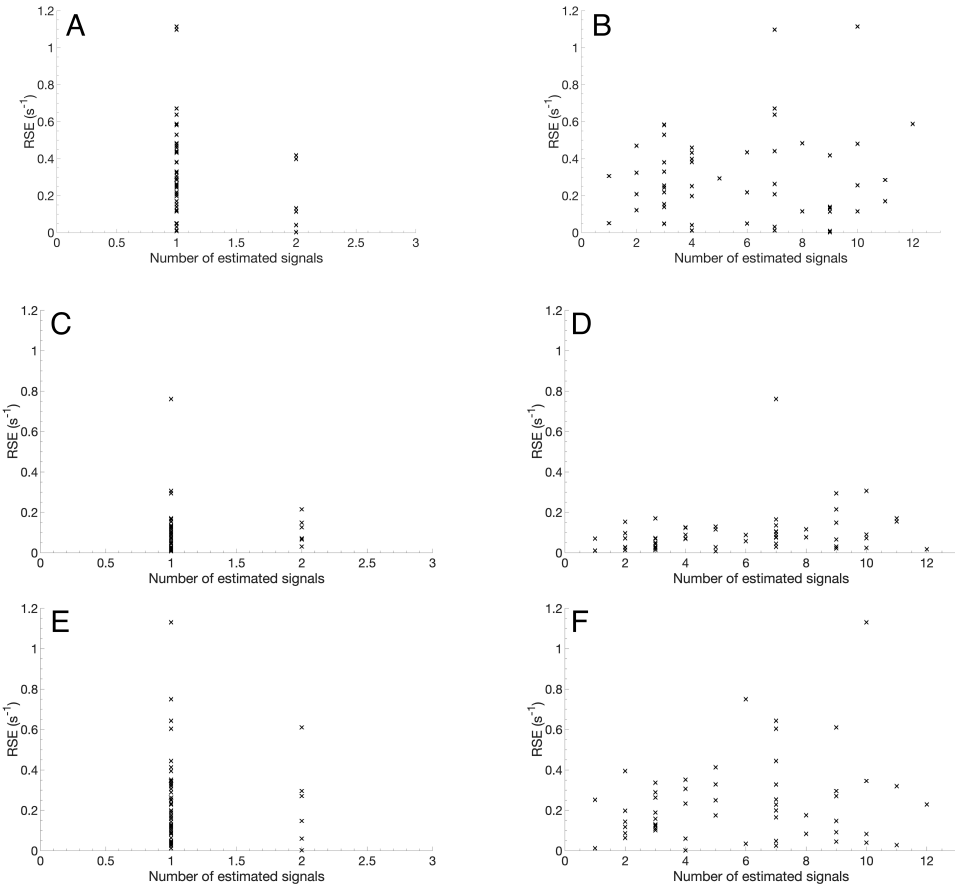


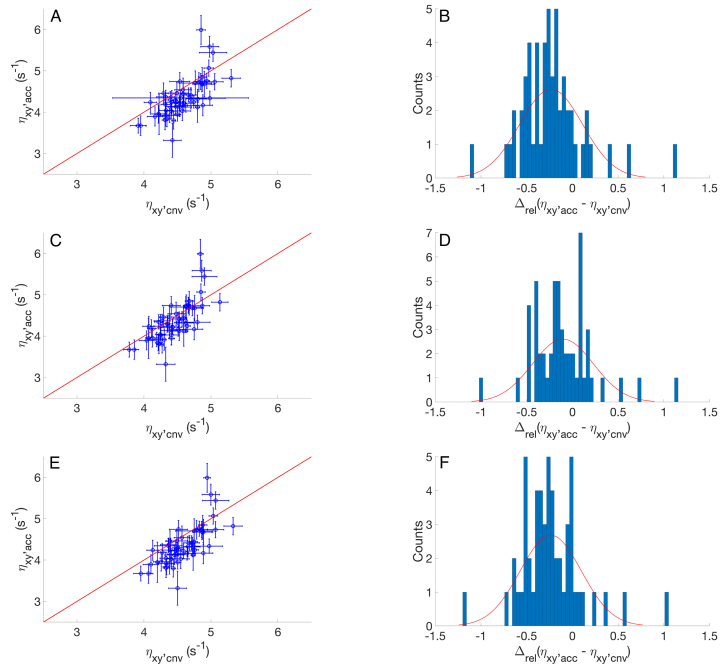
Figure 4: Comparing RSE values against number of estimated signals for the accordion dataset used in NUS sampling. The RSE values is plotted against (A, C, E) the number of estimated signals in the same interferogram, (B, D, F) the number of signals within ±0.1ppm (^{1}H) on either side of a given peak in the spectra. Panels (A, B) plots the RSE between the US accordion dataset and the mean of the conventional datasets, (C, D) plots the RSE between the US and PG(F+R) NUS (ni=88) accordion dataset, and (E, F) plots the RSE between the US and PG(F+R) NUS (ni=48) accordion dataset.

Supplementary Information

An accordion NUS experiment measuring transverse cross-correlated dipole/CSA relaxation

Sven Wernersson[1], Göran Carlström[2], Andreas Jakobsson[3], David Fushman[4], Mikael Akke[1*]

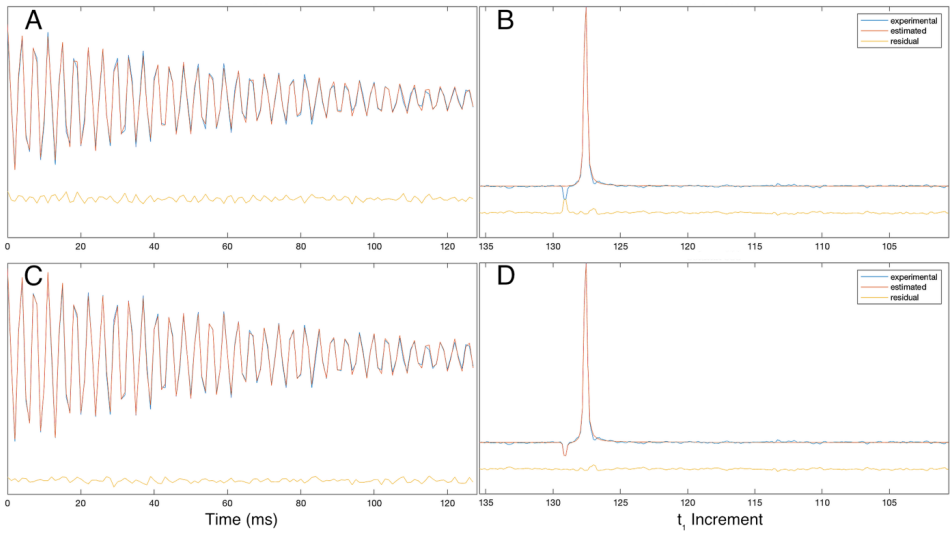
[1] Division of Biophysical Chemistry, Center for Molecular Protein Science, Department of Chemistry, Lund University, Sweden; [2] Centre for Analysis and Synthesis, Department of Chemistry, Lund University, Sweden. [3]Department of Mathematical Statistics, Lund University, P.O. Box 118, 22100 Lund, Sweden



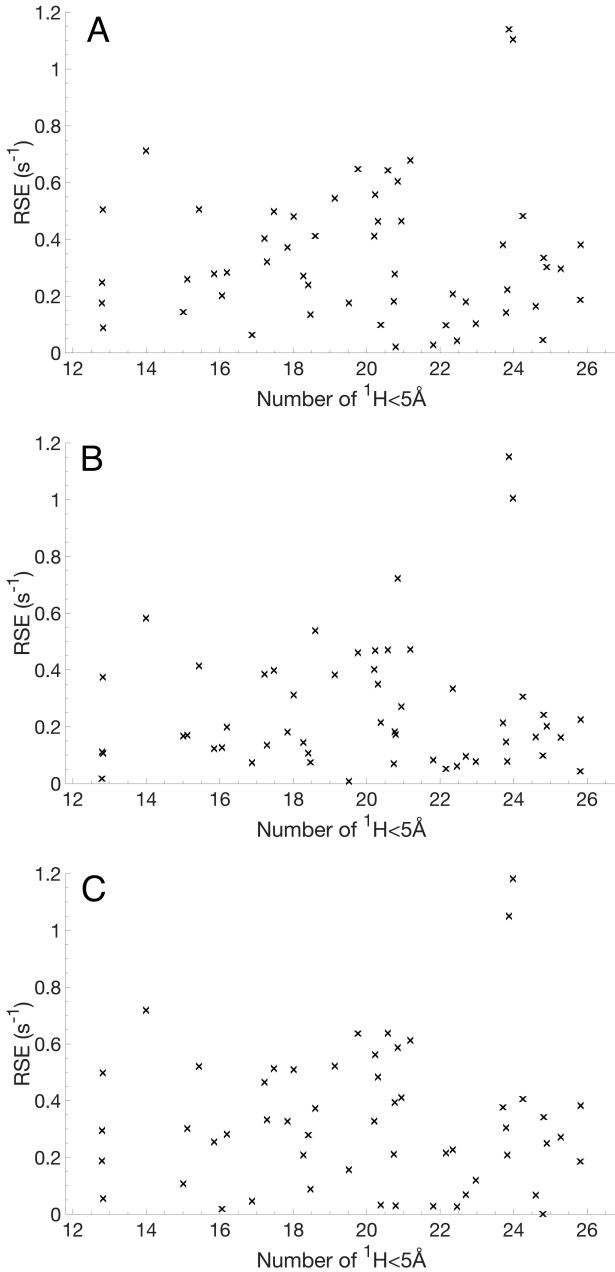
Supplementary Figure 1: Covariance plots and histograms showing the difference between an US sampled η_{xy} accordion dataset (κ=2) and different conventional datasets. Histograms are plotted with fitted normal distributions detailed in SI table 1. (A, B) Covariance plot, histogram comparing the accordion dataset with the conventional dataset measured with a relaxation time of 40ms. (C, D) Covariance plot, histogram comparing the accordion dataset with the conventional dataset measured with a relaxation time of 60ms. (E, F) Covariance plot, histogram comparing the accordion dataset with the conventional dataset measured with a relaxation time of 80ms.

Supplementary Table 1: Fitted normal distributions to the difference between US sampled η_{xy} accordion dataset (κ=2) and different conventional datasets. Distribution mean μ, standard deviation σ and 95% confidence intervals (Conf.) for both parameters.

Data set	μ (s^{-1})	σ (s^{-1})	μ, 95% Conf.	σ, 95% Conf.
Conv, 40ms	-0.2328	0.3437	-0.3266, -0.1390	0.2889, 0.4243
Conv, 60ms	-0.1032	0.3334	-0.1942, -0.0122	0.2803, 0.4116
Conv, 80ms	-0.2371	0.3350	-0.3286, -0.1457	0.2816, 0.4136
Conv, mean	-0.1910	0.3357	-0.2827, -0.0994	0.2822, 0.4145



Supplementary figure 2: Fitted accordion interferogram (A, C) and spectra (B, D) for residue 13 with the TROSY artifact excluded from the fit (A, B) and included in the fit (C, D). Dataset: Accordion η_{xy} acquired using κ=2, (ni, ns) = (128, 64).



Supplementary Figure 3: Difference between ηxy estimated using conventional and accordion experiments plotted against number of ^{1}H-nuclei present within 5 Å of the ^{1}H nuclei in the ^{1}H-^{15}N spin pair. The accordion experiment was acquired using κ=2, (ni, ns) = (128, 64). Conventional datasets are acquired using a relaxation period of length (A) 40 ms, (B) 60 ms, (C) 80 ms.