Galectin-3 Ligand Binding: Mechanism and Driving Forces

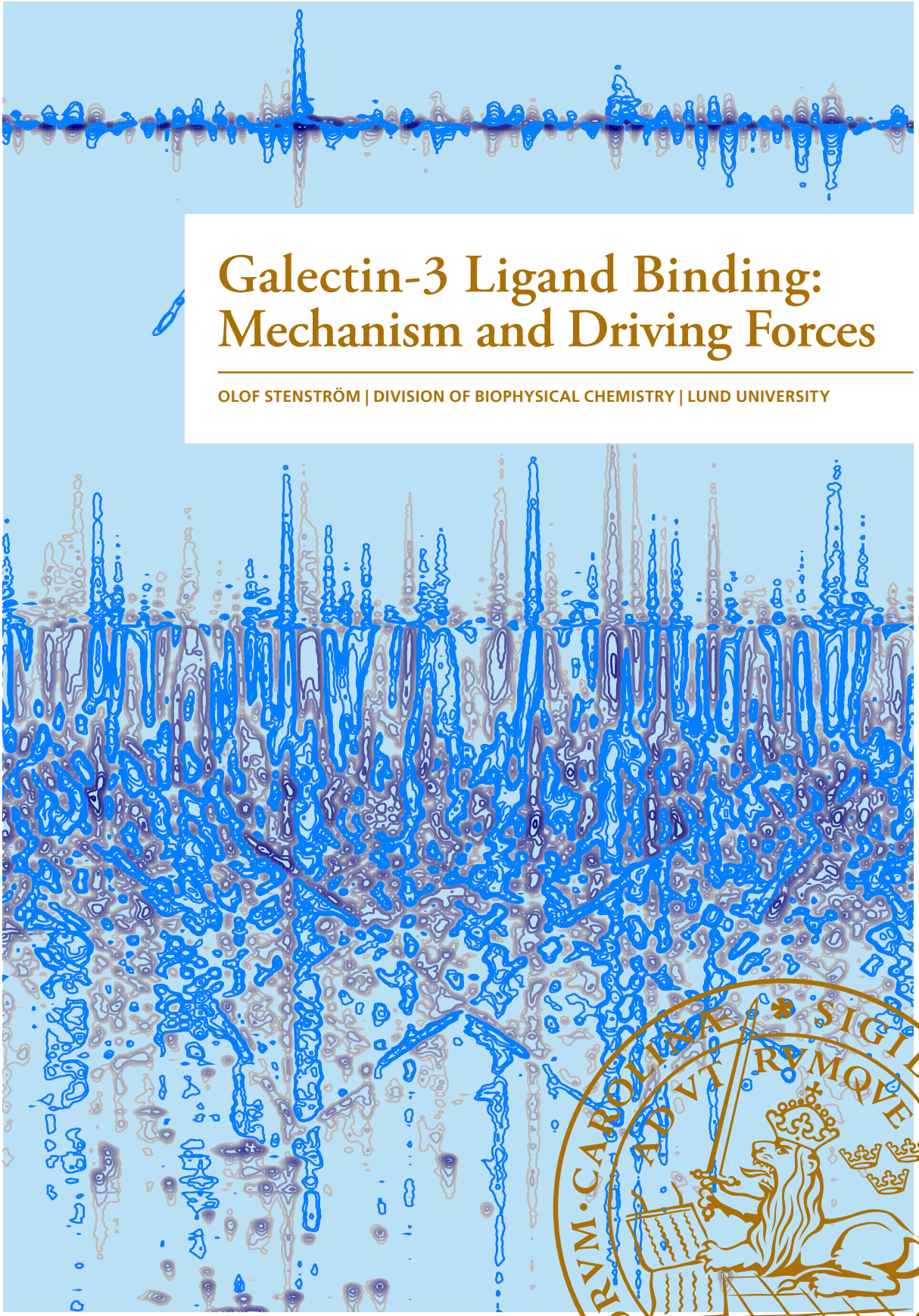
Stenström, Olof

2019

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

Galectin-3 Ligand Binding: Mechanism and Driving Forces

OLOF STENSTRÖM | DIVISION OF BIOPHYSICAL CHEMISTRY | LUND UNIVERSITY



Galectin-3 Ligand Binding: Mechanism and Driving Forces

Galectin-3 Ligand Binding: Mechanism and Driving Forces

by Olof Stenström

LUND
UNIVERSITY

Thesis for the degree of Doctor of Philosophy
Thesis advisors: Prof. Mikael Akke
Faculty opponent: Dr. Martin Blackledge

To be presented, with the permission of the Faculty of Engineering of Lund University, for public criticism in lecture hall A at the Center for Chemistry and Chemical Engineering on Friday, the 13th of December 2019 at 09:00.

Organization
LUND UNIVERSITY
Division of Biophysical Chemistry
Box 124
SE–221 00 LUND
Sweden

Document name
DOCTORAL DISSERTATION

Date of disputation
2019-12-13

Sponsoring organization

Author(s)
Olof Stenström

Title and subtitle
Galectin-3 Ligand Binding: Mechanism and Driving Forces:

Abstract
Knowledge of the molecular driving forces and mechanisms of ligand binding is central to the understanding cellular functions as well as in rational drug design. This thesis investigates the driving forces and mechanisms of ligand binding to the carbohydrate recognition domain of galectin-3 (galectin-3C). Galectin-3 is a beta-galaectosid binding protein which is involved in a myriad of bodily functions and have been indicated to play roll in for example different kinds of cancer and lung fibrosis. With the aid of Isothermal titration calorimetry (ITC), x-ray crystallography, nuclear magnetic resonance (NMR) relaxation experiments and molecular dynamics (MD) simulations we have dissected the different contributions to the entropy of ligand binding, found ligands with low nano molar affinity and characterised the dynamics of two ligand bound to galectin-3C. We Found that conformational entropy contributes significantly to the total entropy difference, solvation entropy entributes aswell but to a lesser extent. Additionally, we have found that lactose binds galectin-3C primarily through the induced fit mechanism and linear free energy relations show that the transition state is closer to the free state compared to the ligand bound.

Key words
Dynamics, NMR, ITC, Relaxation Dispersion, NMR Relaxation, Ligand Binding, Thermodynamics, Galectin-3, Entropy, Binding Mechanism, Induced Fit, Conformational Selection

Classification system and/or index terms (if any)

Supplementary bibliographical information

Language
English

ISSN and key title

ISBN
978-91-7422-706-2 (print)
978-91-7422-707-9 (pdf)

Recipient's notes

Number of pages
168

Price

Security classification

Galectin-3 Ligand Binding: Mechanism and Driving Forces

by Olof Stenström

LUND
UNIVERSITY

A doctoral thesis at a university in Sweden takes either the form of a single, cohesive research study (monograph) or a summary of research papers (compilation thesis), which the doctoral student has written alone or together with one or several other author(s). In the latter case the thesis consists of two parts. An introductory text puts the research work into context and summarizes the main points of the papers. Then, the research publications themselves are reproduced, together with a description of the individual contributions of the authors. The research papers may either have been already published or are manuscripts at various stages (in press, submitted, or in draft).

Cover illustration front: Water by Olof Stenström.

Faculty of Engineering, Division of Biophysical Chemistry

ISBN: 978-91-7422-706-2 (print)
ISBN: 978-91-7422-707-9 (pdf)

Printed in Sweden by Media-Tryck, Lund University, Lund 2019

To my Family

Contents

List of publications

This thesis is based on the following publications, referred to by their Roman numerals:

I **Systematic Tuning of Fluoro-galectin-3 Interactions Provides Thiodigalactoside Derivatives with Single-Digit nM Affinity and High Selectivity**

K. Peterson, R. Kumar, **O. Stenström**, P. Verma, P. R. Verma, B. Kahl-Knutsson, F. Zetterberg, H. Leffler, M. Akke, D. T. Logan, U. J. Nilsson,
J. Med. Chem. 2018, 61, 1164-1175

II **Interplay between Conformational Entropy and Solvation Entropy in ProteinLigand Binding**

M. L. Verteramo, **O. Stenström**, M. Misini Ignjatović, O. Caldararu, M. A. Olsson, F. Manzoni, H. Leffler, E. Oksanen, D. T. Logan, U. J. Nilsson, U. Ryde, M. Akke,
J. Am. Chem. Soc. 2019, 141, 5, 2012-2026

III **Resolving Conformational Selection and Induced Fit Pathways of Protein-Ligand Binding Using NMR Relaxation Dispersion**

O. Stenström, C. Diehl, K. Modig, M. Akke,
Manuscript

IV **Linear Free Energy Relationships of Ligand Binding Kinetics and Affinity Determined by Protein NMR Relaxation Dispersion**

O. Stenström, C. Diehl, K. Modig, M. Akke,
Manuscript

List of publications not included in thesis

I **A De Novo Design Coiled-Coil Peptide with a Reversible pH-Induced Oligomerization Switch**

R. Lizatović, O. Aurelius, **O. Stenström**, T. Drakenberg, M. Akke, D. T. Logan, I. André
Structure, 2016, 24, 946-955

II **Designing interactions by control of protein-ligand complex conformation: tuning argenine-arene interaction geometry for enhanced electrostatic protein-ligand interactions**

A.-L Noresson, O. Aurelius, C. T. Öberg, O. Engström, A. P. Sundin, M. Håkansson, **O. Stenström**, M. Akke, D. T. Logan, H. Leffler and U. J. Nilsson

Acknowledgements

Acknowledgements

I would like to start by thanking my supervisor **Mikael** for giving me this opportunity. When I started my master thesis in your group more than six years ago I did not know what I got myself into. You really challenged me in sometimes, and yes I struggled with relaxation theory. But I think i really needed that and you always took time to talk things through when I had questions.

Pär, my deputy supervisor, thank you for all the support and discussions. **Uli** I have so many things I would like to thank you for and here are a few: you taught me more or less everything I know about the practical part of NMR, for bringing me to coffee, all the beers and many other things. You, **Michalski** and **Johan** also made me feel very welcome in the group. **Kristofer** it has been a lot of fun teaching with you and thank you for all the discussions and support. **Kristine** for always getting my spirits up and sharing your knowledge, it has been great having you in the group. **Sven** and **Simon** for sharing a room with me and all the running. **Göran** for all the NMR help and for always exposing the holes in my understanding and knowledge. I always learn something when I talk to you.

I would also like to thank the Swedish NMR center in Gothenburg and especially **Cecilia** and **Ulrika** for all the help with running experiments. I would also like to thank all the people involved in the DecRec project for all the fun and collaborations. **Ulf R**, **Majda**, **Martin** and **Octav**, it has been great getting to know you and work with you. All the people at CAS and especially **Ulf N**, **Kristoffer** and **Maria** for all the great ligands. **Derek** and **Rohit** for all the structure(s) in my life. At last **Francesco**, I miss you a lot. You always knew how to cheer me up and I am happy to have such good memories together with you and sad that we can't make any new.

Many thanks to the rest of BPC, former and current members. I would especially like to thank **Filip** for all the good times, with and without kids and putting up with all my bs. **Backster** for always cheering me up with your smile and teasing. Just remember, in tennis, half the game is in your mind. **Zhiwei** my tennis partner, still undefeated. **Angus** you should probably be included here too since you were part of the department some time. Thank you for all the laugh and football and I will never forget your rap.

Thanks to all the PhD-students, post-docs, professors and all the other

people at CMPS. **Stefan**, my partner in crime. for all the beers, oysters, life at CMPS would have been a lot more boring if you were not there. **Sambutt**, **Sammi**, **Buttis**, **Sam**, **Butto**, the list goes on, it has been a pleasure hearing you complain and having you listen to me complain, lets keep doing that. For all the running together with **Thom**, **Karin**, **Björn**, **Janina** and the rest who ever joined. **Veronika** for starting BC with me and Stefan, there has been a few. **Rebecca** for just the tip and all the coffee talks. **Ginger lightning** for all the monologues. **Eva** I could always count on you when it came to well all kinds of stuff related to tv-series, It's been great having someone like minded to talk to. **Indian Mafia**, you know who you are, stay classy. **Magnus** for always being there when needed, often coffee or BC, ad for taking all my pr/lanks. **Risto** for all the singing, It's OK if it's from the Thai Way. **Tanja** we will always have Mexico. **Mattias** Vegetarian Cannibal. **Skyscraper** Bourbon, dark chocolate and cake, you know how to make a grown man purr.

There is also a few people outside of LU I would like to thank. First of all **Micke**, thank you for all the help with Vilda and everything else. Thanks to **Emil N**, **Axel**, **Victor**, **Jacob** and **Emil W** for all the distractions, wine and good times.

At last I would like to thank my whole family and especially my sister **Louise** for all the love and making me not think about work for a while. My **mother** and **father** for all the love support and help with Vilda and Freja. **Freja** you are the sunshine of my life! Finally **Emmi** I am sure this would not have been possible without you. Thank you for everything!

Popular summary in English

All bodily functions involve proteins in one way or another, for example movement, digestion and defence against infections. By studying how proteins interact with other proteins and molecules it is possible to get a deeper understanding of the bodily functions. Through this knowledge it is possible to design pharmaceuticals for when something goes wrong. One example of this is cancer where the body's own cells starts to grow uncontrollably. The creation of a new pharmaceutical is a long and expensive process which fail in a lot of stages during the development process. The reasons for failure can be for example a lack of uptake in the body, the drug can turn out to be toxic or it might not bind the target with sufficient strength.

All proteins in humans consists of 20 different amino acids which are linked into chains of different compositions and lengths. These chains in turn take on different folds depending on amino acid composition. It is the fold and the amino acid composition which gives the protein a certain function. One very important function of a protein is the ability to bind other molecules, so called ligands, which can be other proteins, DNA, small molecules or something else. One example of where ligand binding is important is the binding of oxygen to haemoglobin in the lungs to be transported in the blood to the rest of the body. Pharmaceuticals can also bind proteins to provide their healing powers. For example the antibiotic tetracyclin which binds the bacterial ribosome, which is the cells protein factory. Tetracyclin acts by binding and there by blocking the site where new amino acids enter to be added to the growing peptide chain. This leads to that the creation of the new protein is halted.

During my PhD studies I have investigated different aspects of ligand binding to a protein called Galectin-3. This is a protein involved in a multitude of cell functions in and around the cell, such as gene regulation and programmed cell death. Galectin-3 have also been indicated in a large number of diseases such as cancer and lunch fibrosis. This makes Galectin-3 an interesting target for drug interventions.

Isothermal titration calorimetry (ITC) and nuclear magnetic resonance (NMR) spectroscopy have been the two main techniques I have used to study the binding process. ITC is a techniques where small aliquoted of one of the interacting partners are repeatedly added to a of the other. During these additions the heat which is needed to keep the temperature constant

is measured. From the resulting titration curve the binding strength can be calculated alongside the thermodynamic profile of the binding. The Thermodynamics can be separated into two components, enthalpy and entropy. Enthalpy is the heat which is absorbed or released as the two partners bind and the entropy is a measure of the change in disorder of the system as the components bind.

In the second technique, NMR spectroscopy, the quantum mechanical phenomenon of nuclear magnetic spin is used. This phenomenon can be likened to that some atomic nuclei having a small bar magnet. When these are placed in a stong magnetic field the nuclear spins will align with the magnetic field. By manipulating the nuclear spin trough electromagnetic pulses the spins can get out of alignment with the magnetic field. Measuring the rate at which the spins realign with the magnetic field and through a mathematical model it is possible to calculate the amount of dynamics at individual sites in the protein which, in turn can be used to calculate the entropy. Today NMR is the only technique which can get an atomic resolution of the dynamics and entropy.

During my PhD time I have been involved in measuring ligand affinities with ITC for a number of ligands some of whom binds very strongly to Galectin-3. All of these ligands bind with a very favourable enthalpy but with a penalty in entropy. With the help of NMR I have also characterised the difference is dynamics and conformational entropy for a pair of ligands. Collaborators made simulations of the two complexes which indicates that the difference in entropy have a larger contribution from conformational entropy compared to solvation entropy. In another project I have tried to, in the same way as for the protein, calculate the dynamics and entropy for the ligand when it is bound to the protein. I have also looked at the on and off rates of different ligands to the Galectin-3 and correlated these rates to the binding strength. The results show that the ligand on rate is more or less independent of the binding strength but the off rate gets smaller when the binding affinity increases, so the ligand will stay bound for longer when the affinity increases. The last project have been to investigate the binding mechanisms of ligand binding to lactose. When a protein binds a ligand it is common that the protein changes the structure to accommodate the ligand. This leads to the question of which comes first the binding of the ligand or the conformational change? We have shown that for Galectin-3 binding lactose the dominant binding pathway is to first bind lactose and then undergo conformational change.

The hope is that the knowledge gathered in this thesis eventually will lead to a deeper understanding of the underlying driving forces and mechanisms in ligand binding to proteins. This in turn will hopefully lead to a efficient way of designing and develop pharmaceuticals with a high binding affinity, specificity and the desired effect.

Populärvetenskaplig sammanfattning på svenska

Kroppens funktioner utförs till stor del med hjälp av proteiner, exempel på dessa funktioner är allt från att röra på sig till matsmältning och försvara oss mot infektioner. Genom att studera proteiner och hur de interagerar med varandra och andra molekyler i kroppen kan vi få en djupare förståelse för hur kroppen fungerar på en molekylär nivå. Genom denne förståelse kan vi också påverka kroppen när något går fel, t.ex. vid cancer när kroppens egna celler börjar växa okontrollerat. Att skapa ett nytt läkemedel är en lång och kostsam process som kan misslyckad vid många tillfällen under utvecklingsprocessen. Läkemedelskandidaterna kan till exempel sakna förmågan att tas upp i kroppen, de kan visa sig att de är giftiga eller inte binder tillräckligt hårt till sitt mål.

Proteiner består av kedjor av aminosyror med varierande längd, deras funktion bestäms delvis av sekvensen av aminosyror men även hur kedjan veckas. Kroppens alla proteiner byggs upp av 20 olika aminosyror som kombineras på olika sätt för att skapa den unika sekvensen av ett protein. En av de viktigaste funktionerna ett protein har i kroppen är förmågan att kunna binda andra proteiner eller molekyler, så kallade ligander. Som ett exempel på bindning har vi Hemoglobinet i våra röda blodkroppar som binder syre i lungorna för att sedan transportera syret till alla delar av kroppen. Läkemedel binder ofta också till olika proteiner för att delge sin läkande kraft. Ett exempel är antibiotikan tetracyklin som binder till ribosomen i bakterier. Där blockerar antibiotikan nya aminosyror att nå peptidkedjan som skapas där och ribosomen kan därigenom inte skapa nya proteiner.

Under mina doktorandstudier har jag undersökt olika aspekter av ligand bindning till ett protein vid namn Galectin-3. Detta är ett protein som är inblandat i en uppsjö funktioner i och runt cellen som t.ex. reglering av genuttryck och celldöd. Detta gör Galectin-3 till ett intressant mål för läkemedels intervention eftersom forskning har visat att Galectin-3 är inblandat i flertalet sjukdomsförlopp så som olika typer av cancer och lungfibros för att nämna några.

Isoterm titreringskalorimetri (ITC) och kärnmagnetisk resonans spektroskopi (NMR) har varit de två tekniker jag har använt för att studera bindnings-processen. ITC har jag använt för att bestämma hur starkt Galectin-3 binder till de olika ligander som ingått i arbetet. Utöver det får man information om den termodynamiska profilen för bindningen. Termodynamiken för bindning kan separeras i två komponenter, entalpi och entropi.

Entalpin säger hur mycket värme som tas upp eller avges vid bindningen. Entropin är ett mått på hur mycket ordningen ändras i systemet (protein, ligand och lösningsmedel). I den andra metoden, NMR, utnyttjar man de kvantmekaniska fenomenet spinn som kan liknas vid att det finns en liten magnet i atomkärnan. När man utsätter spinnen för ett starkt magnetfält kommer de att rikta sig med magnetfältet. Genom att manipulera dessa spinn med elektromagnetiska pulser kan man få dem att ändra riktning. Genom att sedan mäta med vilken hastighet spinnen riktar tillbaka sig till magnetfältet kan man få information om rörelser i proteinet som sedan kan relateras till entropin. Detta gör att man kan undersöka entropin med atomär upplösning tillskillnad från ITC där man endast får den totala för hela systemet. I dagsläget är NMR den enda metoden man kan bestämma dynamiken och entropin experimentellt på en molekylär nivå.

Mitt doktorand-arbete har gått ut på att undersöka olika aspekter av ligand-bindning till Galectin-3. Jag har delvis gjort ITC för flertalet ligander och undersökt hur bindningsstyrkan och den termodynamiska profilen ändras när man ändrar utseendet på liganden. I samband med detta arbetet har vi hittat ligander som binder väldigt starkt. Dessa Ligander binder med en fördelaktig entalpi men en entropi som bidrar negativt till bindningen. Jag har med hjälp av NMR undersökt entropi skillnader mellan olika galectin-3 ligand komplex. och med hjälp av samarbetspartners som har simulerat bindningen undersökt samspelet mellan skillnaden i konformationell entropi och lösnings entropin. I detta fallet väger den konformationella entropin tyngre för att bestämma den totala entropiskillnaden. Jag har även arbetat med att på samma sätt som för proteinet försöka karaktärisera dynamiken och entropin för liganden. En del av mitt arbete har gått ut på att undersökt med vilken hastighet liganden binder till och släpper proteinet och korrelerat det till bindningsstyrkan för liganden. Där ser vi att hastigheten som liganden binder in med ändras inte mycket med bindningsstyrkan utan det är framförallt hastigheten som liganden lämnar proteinet som minskar när bindningsstyrkan ökar. Sist har jag undersökt mekanismerna för ligand bindning. Många proteiner genom går en ändring i struktur när de binder en ligand. För vissa proteiner kan man se att proteinet ibland antar den ligand bundna strukturen även när liganden inte är närvarande. Detta leder till frågan vad som sker först, ligand bindningen eller den strukturella ändringen? Vi har visat att för Galectin-3 som binder till laktos sker den strukturella ändringen huvudsakligen efter att laktosmolekylen har hittat till bindningssätet.

Förhoppningen är att denna detta arbetet ska kunna användas som en pusselbit för att få en djupare förståelse för de underliggande drivkrafterna och mekanismerna i ligandbindning till proteiner. Detta i sin tur kan leda till att man på ett effektivare sätt ska kunna designa ligander som specifikt binder ett visst mål i kroppen och som dessutom ger den önskade effekten.

Chapter 1

Introduction

Basically all processes in and around a cell involves proteins, these proteins are essential for the survival of the cell and hence all living organisms. Even though proteins are only made up of 20 different types of amino acids they take on a plethora of shapes, sizes and perform a huge range of tasks, from catalysis [1, 2] to signaling [3–5] and transport [6, 7]. A protein is a linear polymer which is built up of said amino acids connected with a peptide bond, see figure 1.1 A. In many cases the chain of amino acids take on distinct secondary structures like α-helices and β-sheets, panel B in figure 1.1, which in turn are folded in to a well defined 3-dimensional structure, panel C figure 1.1 giving the protein its unique function [8]. In some cases two or more proteins come together to form a complex in order to perform a function. One example is the ribosome [9] which is a large protein complex that produces new proteins. In other cases the protein lacks a well defined structure, these are called intrinsically disordered proteins (IDPs). Despite the lack of structure these proteins have proven to play an important role in biology [10].

One very important function of a protein is the ability to recognise and interact with other molecules, so called ligands. These molecules can be for example other proteins, DNA or small molecules. The binding ligands is central to the function of proteins for example to perform catalysis or signaling. The study of protein-ligand interactions is key in understanding the activity of a protein and therefore also in in drug development. The aim of this thesis is to try to get a deeper understanding for the underlying driving forces and mechanisms of ligand binding to proteins. To do this I have used a combination of nuclear magnetic resonance spectroscopy

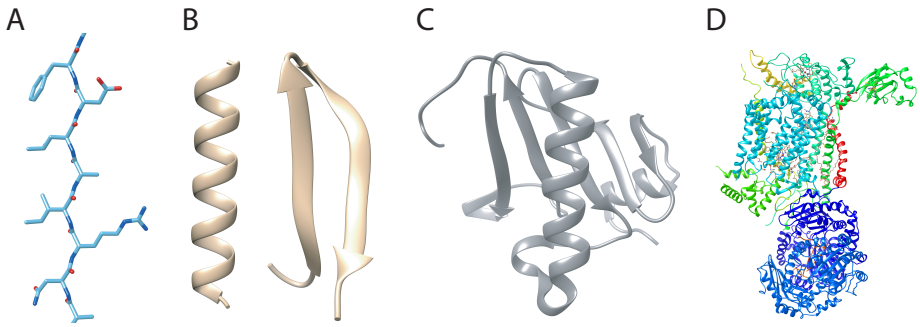


Figure 1.1: Showing a short chain of amino acids (A). Secondary structures commonly found in proteins (B), to the left an α-helix and to the right an anti-parallel β-Sheet with two strands. The protein ribonuclease from Streptomyces aureofaciens (1RGE) [11], with β-sheets and α-helices connected with loop regions (C). The Cytochrome bc_1 complex from chicken (1BCC) where each protein in the assembly is colored in different colors (D) [12].

(NMR) and isothermal titration calorimetry (ITC) and studied the binding of Galectin–3C to a wide range of ligands.

1.1 Protein-Ligand binding

The binding of a protein P to ligand L, see figure 1.2, to form the complex PL, can be described by the reaction formula:

$$P + L \underset{k_{off}}{\overset{k_{on}}{\rightleftharpoons}} PL \tag{1.1}$$

Where $\rightharpoonup$ represents the binding of the protein and ligand with a reaction rate constant k_{on} and $\leftharpoondown$ represents the dissociation of the complex with a reaction rate constant k_{off}. For a system at equilibrium the rates of association and dissociation are equal and can write as:

$$[L][P]k_{on} = [PL]k_{off} \tag{1.2}$$

In which the brackets indicates equilibrium concentrations. This means that when the system is at equilibrium the number of binding events is equal to the number of dissociation events meaning that the concentrations of the different species are unchanged over time as long as all parameters are kept constant. The association constant (K_a) and the dissociation constants (K_d) can be written as:

$$K_a = \frac{k_{on}}{k_{off}} = \frac{[PL]}{[P][L]} = \frac{1}{K_d} \tag{1.3}$$

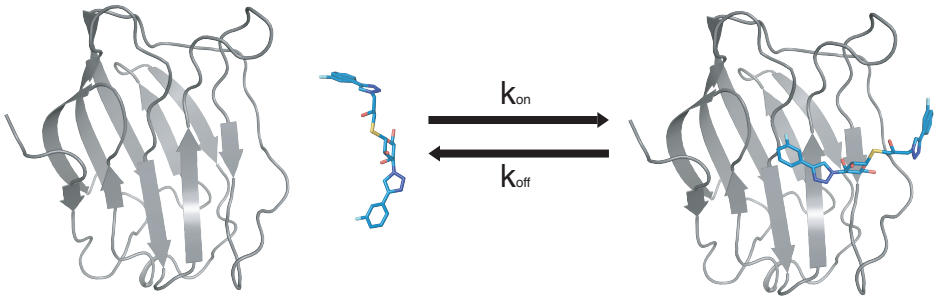


Figure 1.2: Binding of Galectin-3C with the Ligand S (see paper II), the backbone of the proteins is in cartoon representation in gray and the ligand is in blue.

The association and dissociation constant are related to the gibbs free energy of binding (ΔG_b) as:

$$\Delta G_b = RTln(K_d) = -RTln(K_a) \tag{1.4}$$

where T is the temperature in Kelvin and R is the gas constant. The binding is more favorable the more negative the free energy of binding is, which is equivalent to smaller K_d or a large K_a [13, 14]. The free energy of binding can be seperated into an enthalpic (ΔH_b) and entropic (ΔS_b) components as:

$$\Delta G_b = \Delta H_b - T\Delta S_b \tag{1.5}$$

More about entropy in section 1.1.1. The enthalpy of binding is the energy change of going from the free form to the bound. In protein ligand interactions this energy comes from the difference in energy of noncovalent interactions of the protein, ligand and solvent when dissociated compared to the interactions of the the protein-ligand complex and the solvent. These interactions can be for example hydrogen-bonds, van der Waals interactions and electrostatic interactions [14–17].

1.1.1 Entropy

The entropy is commonly referred to at the number of available states or the amount or disorder in a system. For a system with n energy levels with energy ϵ_i the population of an energy level i can be written according to

the Boltzmann distribution:

$$p_i = \frac{e^{\epsilon_i / k_b T}}{Q} \tag{1.6}$$

Where T is the temperature in Kelvin, k_B is Boltzmann's constant and Q is the partition function which is written as:

$$Q = \sum_n g_n e^{-\epsilon_n / k_b T} \tag{1.7}$$

where g_n is the degeneracy of energy level n. The energy of a system ϵ_i is the sum of the energy from all classes of energy levels $\epsilon_i = \epsilon^{trans} + \epsilon^{rot} + \epsilon^{vib} 0 \epsilon^{el}$ where trans represents translational energy levels, rot represents rotational, vib represents vibrational and el represents the energy levels of electronic excitations [18]. We can see from equation 1.7 that when T $\rightarrow$ 0 that q $\rightarrow$ g_0 and when T $\rightarrow \infty$ q $\rightarrow \infty$. This gives us an indication that the partition function gives us a measure of the number of states available at a certain temperature. From the partition function we can write Helmholtz free energy.

$$A = -k_B T ln Q \tag{1.8}$$

and from which we can get the entropy by taking the partial derivative of Helmholtz free energy with respect to temperature.

$$S = k_B ln Q \tag{1.9}$$

The total entropy of binding can be seperated the contributions:

$$\Delta S_b = \Delta S_{solv} + \Delta S_{conf} + \Delta S_{rot/trans} \tag{1.10}$$

in which ΔS_{solv} represent the change in the solvent entropy upon ligand binding to the protein. Examples of processes which contribute to the changes in solvent entropy can be rearrangement or release of waters on the surface of protein or ligand [18–20]. ΔS_{conf} is the result of changes in the conformational freedom of the protein and ligand, which might either increase of decrease in ligand binding [21–24]. The last component is $\Delta S_{rot/trans}$ which is the entropic contribution from the loss of translational and rotational degrees of freedom of the protein and ligand [25, 26].

1.1.2 Entropy-Enthalpy compensation

The phenomena of a small change in Gibbs free energy, as a result from often large and counteracting changes in both entropy and enthalpy, is

commonly known as entropy-enthalpy compensation. This effect is especially apparent in aqueous solutions [27,28]. Experiments offer evidence for a compensatory behaviour when varying thermodynamic parameters for example temperature or pressure in micelle formation experiments, denaturation, [29–32]. The binding of different ligands to a protein target have also shown a tendency for entropy-enthalpy compensation [33–37]. Different physical explanations for entropy-enthalpy compensation have been suggested, some of which are:

Solvent rearrangement when the ligand binds a protein have been suggested as an explanation for the compensation behaviour. Where the number of displaced waters determines Gibbs free energy and that the changes in hydrogen-bond network of the waters determines the entropy and enthalpy of binding [38]. Grunewald and Steel have even suggested a statistical mechanical formalism for solvent reorganization and attempt to demonstrate how entropy-enthalpy compensation is a feature of nearly all reactions in solvent [39]. Water release from the binding site of a model system have been shown to reduce or even nullify the reduction in entropy due to loss of conformational flexibility in a binding event thus circumventing entropy-enthalpy compensation [40].

A simple physical explanation of the compensatory behaviour proposed is stronger attractive forces pull ligand and receptor closer together thus restricting the mobility and reducing entropy [37, 41]. Dunitz [42] uses a thermodynamic method to argue for entropy-enthalpy compensation as a general property of weak interactions that is almost unavoidable.

1.1.3 Binding mechanisms

In 1894 Emil Fischer [43] introduced the Lock-and-Key model for ligand binding to a rigid receptor, in which the protein conformation of free and ligand bound are identical. This model have for a long time been considered too primitive to describe ligand binding since it neglects the plasticity in the receptor. X-ray crystallography, NMR spectroscopy and single molecular fluorescence detection have all shown that binding of ligands to proteins can be associated with both large and small conformational changes [44–48]. Theses conformational changes can be present in absence of ligand or when the ligand is bound to the protein [49–51]. This leads to the important question: How are the conformational change coupled to the ligand binding? Two scenarios named induced-fit and conformational selection can be imagined if the lifetimes of the individual states are significantly longer

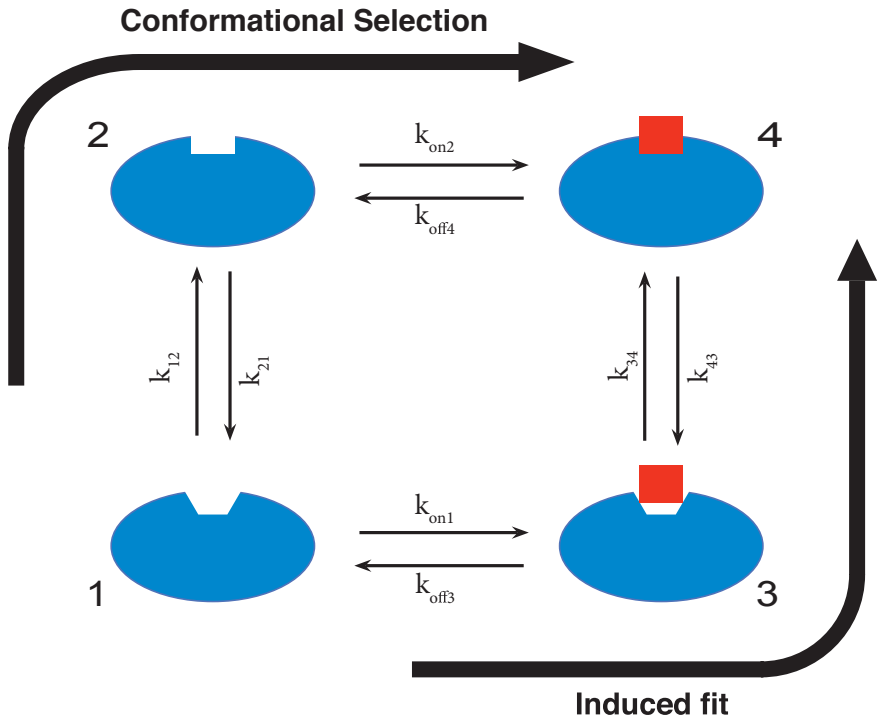


Figure 1.3: Molecular recognition mechanism for ligand binding coupled to conformational change. The protein is represented by the blue spheres and the ligand by the red square. States 1 and 2 represents the protein in the open, low affinity and closed high affinity state, respectively. The thin arrows represents the transitions available for each state of the protein, with corresponding rate constants denoted k. The thick arrows represent the two possible binding mechanisms, conformational selection and induced fit.

compared to the time of the conformational transitions and ligand binding. In the induced-fit model the ligand binds a low affinity ground state of the protein and induces a conformational change to a high affinity ligand bound state [52], see figure 1.3 . In the case of conformational selection, the protein samples a high energy state which has a high affinity for the ligand. The ligand selectively binds to this state of the protein [53]. Both NMR and single molecule fluorescence detection experiments have shown evidence of conformational change in the free state of the protein sampling a state similar to the ligand bound and the ligand bound protein sampling a state similar to the free [44, 49, 54–56]. These conformational changes in presence or absence of ligand have often been seen as evidence of indication of one or the other of the two binding mechanisms. These mechanisms are the extremes of binding but for protein-protein or protein peptide interactions for example this might not be the case due to of for example

larger conformational changes or folding [57–60]. In order to distinguish which pathway is dominant under given conditions the flux through each pathway have to be compared [61].

$$F_{IF} = \left(\frac{1}{k_{on1}[P_1][L]} + \frac{1}{k_{34}[P_3]} \right)^{-1} \tag{1.11}$$

$$F_{CS} = \left(\frac{1}{k_{12}[P_1]} + \frac{1}{k_{on2}[P_2][L]} \right)^{-1} \tag{1.12}$$

In which F is the flux though either the IF or CS pathway, k represents reaction rate constants, see figure 1.3 and square brackets denotes concentrations of the different components. In paper III we investigate the binding mechanism of lactose binding to galectin-3C and use the flux equations to destinguish between the two binding pathways.

1.2 Drug discovery

The drug discovery process starts with the identification of a potential target. This process is essential, since the two main reasons that a drug fails in clinical trials are that it is not safe or lacks the desired effect. After the target has been identified and validated to be involved in the diseases process the search for hit compounds can begin. There are many ways of going about finding a hit compound, some common are: high throughput screening in which a large amount of compounds are screened for activity. Fragment screening where crystals of protein is soaked with small compounds to find weakly binding molecules to be combined into larger scaffold for further development. Structure aided drug design in which the crystal structure of the target is used for designing molecules, to mention a few methods [62–65].

Next, the hit compounds are developed to increase the affinity for the target by introducing hydrogen bonds and increasing hydrophobicity for example, and at the same time follow chemical parameters such as Lipinskies rule of five [66]. Using structure based experimental and computational methods the hit compounds are developed into lead compounds which have the desired activity on the target [62, 67]. In order to optimize the ligand affinity for a target there is a need to overcome the enthalpy-entropy compensation.

A favourable enthalpic interaction, for example a hydrogen bond, is often penalized with a unfavourable entropy, see section 1.1.2. One reason for the entropy penalty can be the structuring of flexible regions in the protein. It has been suggested that this can be overcome by targeting already structured parts of the protein or directing more than one hydrogen bond at the same flexible region [68].

The lead compounds are modified in a way that retains favourable properties and improve deficiencies for example the ability to cross the blood brain barrier if the target is located within the central nervous system [62]. Last but not least one or two compounds are selected for clinical trials.

Chapter 2

Galectin-3

2.1 Galectins

Lectins make up a family of carbohydrate binding proteins in which galectins are a members. Galectins are characterized by an affinity for -galactosides and with a sequence similarity in the carbohydrate recognition domain (CRD) [69, 70]. This far, all known human galectins have a single polypeptide chain with one or two CRDs and a flexible peptide of varying length. Galectins are divided into three subgroups based on how their domains are organized [71], see figure 2.1. The proto-type have a single CRD followed by a short flexible peptide and includes galectins-1, -2, -5, -7, -10, -13, -14. The chimera type have a CRD and an other type of domain or a long flexible peptide, this group has only one member, galectin-3. The last subgroup is the tandem repeat which are galectins consisting of two CRDs and includes galectin-4, -6, -8, -9, -12 [72]. The galectin CRD is a slightly bent antiparalelle beta-sandwich of about 135 amino acids. The carbohydrate binding side of the sandwhich is concave and consists of six β-sheets and the opposite side is convex and consists of five. Galectins are synthesised by cytosolic ribosomes, they can however be found both extra- and intra-cellulary.

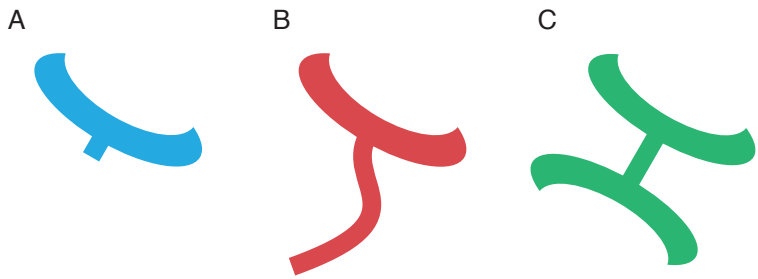


Figure 2.1: Galectin subclass types. The CRD are shown as crescents. A depicts the proto type group (blue), B the chimera type (red) and C the Tandem repeat type (green).

2.2 Galectin-3

Galectin-3 is this far the only known member of the chimera type group of galectins because of its, compared to the prototype group, long N-terminal tail. The human Galectin-3 has 250 residues with a C-terminal CRD and a N-terminal flexible peptide consisting of 18 conserved amino acids followed by 7-14 repeats, each having 8-11 amino acids that include one aromatic and multiple Pro and Gly, sometimes know as the collagen-like N-terminal domain [73, 74]. Galectin-3 can multimerise and the N-terminal domain have been shown to be essential for this function [75]. Some Matrix metalloproteinases have the ability to cleave the N-terminal of Galectin-3 which decreases the propencity for self association but increases the affinity for glucoconjugates [76].

Galectin-3 is expresses in a range of tissue type but is mainly related to epithelial cells [77–85] and can be found both in the cytocol, nucleus and extra cellular matrix [86–88]. Many interaction partners have been identified for galectin-3 in and outside the cell indicating an involvement in a great range of biological processes among others inflammation [89] and wound healing [90]. In the cytocol galectin-3 interacts with for example Blc-2 [91] and CD95 [92] which are involved in reguation of apoptosis. In the nucleus Galectin-3 is involved in mRNA splicing [93, 94] and have also been shown to have single stranded DNA and RNA binding ability [87] and have additionally been implied as a regulator of gene transcription [89, 95, 96]. Extracellulary Galectin-3 have also shown to be involved large extent in processes such as interactions with the extracellular matrix, angiogenesis, endocytosis and neuronal functions [97]. With this large repertoire of biological functions it is not surprising that Galectin-3 have been found to

play a roll in many disease processes. Examples of diseases are different forms of cancers like colon [78], breast [84], protstate [98] and the expression levels can be an indicator of progression [99–101]. In addition, galectin-3 have also been implicated in other deceases like pulmonary fibrosis [82] and heart failure [102]. The combination of biological function and decease involvement makes galectin-3 highly interesting for drug intervention.

Chapter 3

Isothermal Titration Calorimetry

Ligand binding to proteins can be studied by a multitude of techniques, where most rely on either measuring population changes or signals from a reaction [103]. Calorimetry, on the other hand, is the only technique which measures the binding energetics directly where no reporter molecule is needed [104]. A drawback of measuring the heat is that the signal of the reaction is proportional to the binding enthalpy. Since protein ligand interactions are non-covalent and these interactions usually exhibits rather small binding enthalpy the resulting signal from an ITC experiment is weak. this results in the need for, compared to other techniques, large amounts of material.

3.1 Instrumentation

The ITC instrument, see figure 3.1, consists of a syringe with stirrer blades on the needle and matched reference and sample cells which are insulated from the environment. Both cells are coin shaped and sandwiched together with a thermocouple device which measures the difference in temperature between the two cells. On the flat outside surfaces of the cells are attached heaters to regulate the temperature. Let's imagine we are running an ITC experiment in which we have a protein in the cell and a ligand with affinity for the protein in the syringe. During an experiment a small amount of power is continuously fed into the reference cell which activates the ther-

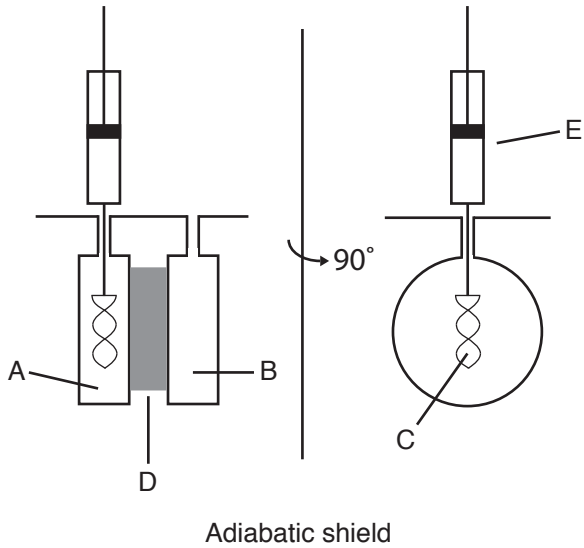


Figure 3.1: A Schematic figure of an ITC instrument. A is the sample cell, B is the reference cell, C is the tip of the syringe with the stirrer blades, D is the thermocouple device, E is the syringe containing the titrant.

mocouple device that regulates the power input to the sample cell to a baseline level [105]. After the baseline has been reached the syringe injects and mixes small aliquots of the ligand solution into the sample cell where the protein and ligand will bind. In case of an exothermic or endothermic reaction the feedback power will either decrease or increase, respectively, to counteract the heat change from the reaction. This temporary divergence from baseline is recorded in the form of a thermogram, left panel of figure 3.2. In the case of no reaction the feedback power will have a minor change due to other effects for example mixing.

3.2 Analysis

As described above, in the experiment we titrate the ligand into the cell and the protein bind the ligand resulting in a change of heat, until there is no more protein molecule not bound to the ligand. The result is a thermogram, see figure 3.2, which is the power input as a function of time. Each injection results in a peak in the thermogram and the area under the peak corresponds to the heat absorbed or released from that injection. By integrating all the peaks in the thermogram and fitting the results to the

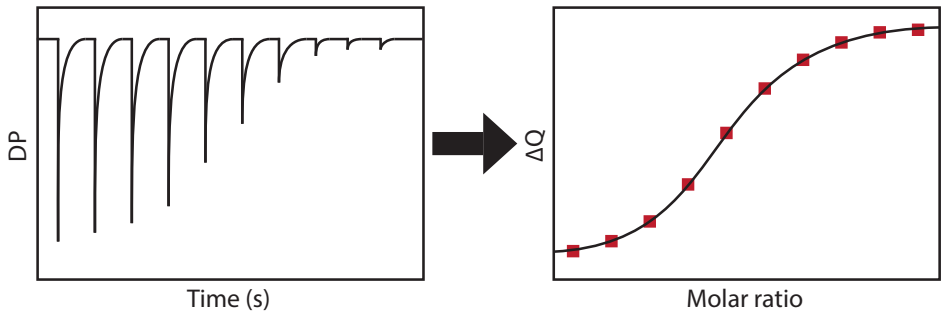


Figure 3.2: Left panel is a schematic overview of a thermogram with time in seconds on the x-axis and differential power in $\mu J/s$ on the y-axis. The right panel depict the integrated heats of the thermogram to the left where the red dots represents the area under each peak with molar ratio on the x-axis and the heat in kJ/mol in the y-axis.

equation 3.1 it is possible to determine the enthalpy of binding, dissociation constant and stoichiometry, through which the entropy and Gibbs free energy can be calculated. The heat developed as a result of each injection can be calculated as [106]:

$$\Delta Q(i) = Q(i) + \frac{V_i}{V_c}\left[\frac{Q(i) + Q(i-1)}{2}\right] - Q(i-1) + Q_0 \tag{3.1}$$

In which V_i is the injection volume, V_c is the cell volume, Q_0 is the offset heat accounting for the heat of mixing and $\Delta Q(i)$ is the heat function following the ith injection. For a reaction where one protein binds one ligand from the example above the heat function can be written as:

$$Q(i) = (\Delta H * V_c/2) * \left[\alpha - \sqrt{\alpha^2 - 4nM_iX_i}\right] \tag{3.2}$$

Where $\alpha = nM_i + X_i + K_d$ in which M_i and X_i is the cell and syringe component, respectively. K_d is the dissociation constant, n is the stoichiometry and ΔH is the enthalpy of binding. Gibbs free energy (ΔG) can then be calculated as: $\Delta G = RTln(K_d))$ where R is the gas constant, T is the temperature in Kelvin. The binding entropy in turn can be calculated according to $\Delta G = \Delta H - T\Delta S$.

In the heat function above there are three unknown parameters to be fitted, namely the binding constant, stoichiometry and the binding enthalpy. The effect on variations of the different parameters can be seen in figure 3.3. The effect of changing the stoichiometry on the isotherm (figure 3.3 panel A) is displacement in the X-direction. The dissociation constant K_d effects

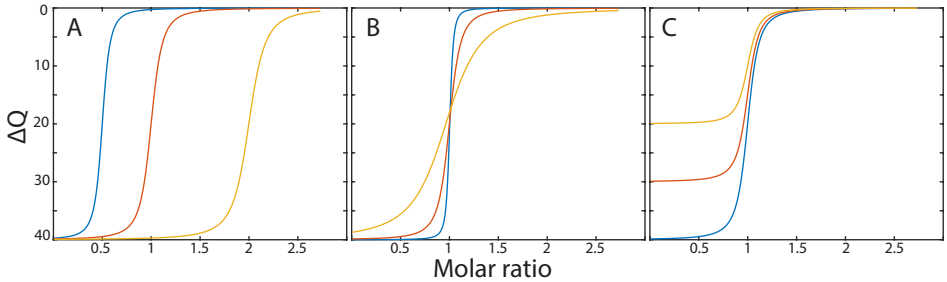


Figure 3.3: ITC isotherms simulated under different conditions. Panel A depict differences in stoichiometry, 0.5, 1 and 2 for blue, red and yellow respectively. Panel B shows different heat curves for binding constants K_d of 10, 100 and 1000 nM for blue, red and yellow, respectively, which corresponds to c-values of 3000, 300 and 30. Panel C depict the resulting heat profiles of different enthalpys (ΔH), -40, -30 and -20 kJ/mol for blue, red and yellow respectively.

the isotherm by changing the slope (panel B), a lower dissociation constant results in a steeper slope. The size of the isotherm step is effected by the binding enthalpy (panel C), where a large binding enthalpy results in a large step and a small enthalpy gives a small step.

The sources of error in ITC are many but a large portion of them can be minimized and/or circumvented by careful design of the experiments. The c-value, a unitless parameter, is a useful tool to design experiments in a way that all parameters are properly determined:

$$c = K_a[M_t]n \tag{3.3}$$

In which , K_a is the association constant ($K_d = 1/K_a$) and M_t is the total cell concentration. For c-values between 1 and 1000 the enthalpy and binding constant can be determined accurately [105, 107]. The enthalpy is best determined for c-values over 50. Under these conditions the first injections goes to near completion and ΔH is almost independent of K_d. To determine the dissociation constant accurately c-value below 500 is optimal. The resulting isotherm has enough data points in the transition for a good determination of the K_d [108]. In order to get around the restrictions described above experimental parameters can be changes, for example temperature or pH [109, 110]. The displacement method have also been successfully applied for both low and high affinity reactions, see section 3.2.1 below [12, 35]. The recommendation of c-values between one and a thousand have been questioned and Tellinghuisen [111] have shown that determination of the dissociation constant is reliable for c-values down to 10^{-4} and enthalpy through its temperature dependence.

3.2.1 Displacement ITC

A problem with ITC is to accurately determine the binding constant of high affinity ligands (dissociation constant in low nanomolar range and lower), since the measurement of a very strong binding ligand would require concentrations in the range where the heat signal would be undetectable. Different can be employed to counter this problem for example by changing the temperature or pH. These methods rely equations to extrapolate back to the original experimental conditions [112]. We have instead employed a method called displacement or competitive ITC [35].

In this approach the protein (M) is first saturated with a weaker ligand (L) with known affinity and enthalpy. The high affinity ligand (X) is then titrated into the mixture of protein and weak ligand and out competes the weak ligand in the protein binding site. The drawbacks of this method is that you need a second ligand that binds with a factor at least 10 or weaker to the protein and has a significantly different enthalpy [113]. The heat function associated with the formation and dissociation of the complexes now becomes:

$$Q(i) = V_c * (\Delta H_L * [ML]_i + \Delta H_X * [MX]_i) \tag{3.4}$$

In which ΔH_L and ΔH_X is the binding enthalpys and [ML] and [MX] is the complex with the protein with the lignads L and X, respectively. The concentrations of the complexes in turn can be expressed in terms of dissociation constants and stoichiometry [112, 113].

3.3 Sources of Error

The disadvantage of measuring the energetics is that everything releasing or absorbing heat during the titration will contribute to the signal in the thermogram. This problem can be countered by matching all components in the solutions being mixed except for the components of interest [15]. This can be done by dialysing the solutions against the same buffer for example. The need to know the exact concentrations of the reagents is also critical to reliably determine the enthalpy and dissociation constant [114].

Systematic errors can be detected by plotting the residuals of the fitted titration curve. A nonrandom distribution around zero can indicated that

the wrong model have been chosen or that there is an error in one of the concentrations or volumes. Systematic errors that will not be detected by the residuals are errors in the voltage in the instrument or for some model selections [108]. The systematic error in the voltage can be detected by systematic use of standard reactions to validate the accuracy of ITC instruments [115]. Baranauskiene et. al. [116] highlights systematic errors between instruments and suggests possible reactions to use for calibration. The errors in model selection can be detected by repeating the experiments at different concentrations.

Common practice in ITC is also to have a small first injection which in the analysis is discarded on the grounds that reactant from the syringe diffuses out of the tip of the syringe. It has been reported that the primary cause of the first injection anomaly is a result of a mechanical error. If the last action of the syringe before starting the experiment is to drive the plunger up the change of direction when doing the first injection can cause a backlash error from changing direction. This error can easily be overcome by executing a small down action of the plunger before starting the experiment [117].

Chapter 4

NMR

The nuclear magnetic resonance (NMR) phenomena was discovered almost simultaneously and independently by the groups of Purcell (December 1945) and Block (January 1946) and published their findings in the same issue of Physical review in 1946 [118,119]. Since then the number of applications for NMR have exploded and the range of uses range from imaging of the brain or other parts of the body [120, 121] to Structure determination [122, 123] of proteins and study of chemical exchange [124, 125]. This chapter will give a brief introduction to NMR and the majority of the chapter is based on the textbooks [126–128].

4.1 Spin and Magnetic Moment

Many nuclei has an intrinsic angular momentum called spin, denoted **I**. Some examples of nuclei with spin $\neq 0$ are the ones I have used during my PhD studies which are ^1H, ^2H, ^{15}N, ^{13}C and ^{19}F. The angular momentum of particles with spin is not due to rotation of the particle but an intrinsic property of the nuclei which can take values of:

$$|\mathbf{I}| = [\mathbf{I} \cdot \mathbf{I}]^{1/2} = \hbar[I(I+1)]^{1/2} \tag{4.1}$$

where $\hbar$ is Plancks constant divided by 2π. A particle with spin **I** have 2I+1 sublevels, where I is the nuclear spin angular momentum quantum number. The sublevels takes on different energies in the presence of a magnetic field but are otherwise degenerate. Nuclei with a spin, additionally posseses a nuclear magnetic moment, μ. The nuclear spin angular momentum is

proportional to, and colinear to the nuclear magnetic moment:

$$\mu = \gamma \mathbf{I} \tag{4.2}$$

in which γ is the gyromagnetic ratio, unique for each type of nucleus, for example ^{1}H and ^{15}N has gyromagnetic ratios of 2.675×10^{8} and -2.713×10^{7} $(Ts)^{-1}$, respectively. By convention the z-component of the nuclear spin angular momentum is $I_z = \hbar m$ where m = (-I,-I+1,...I-1,I) is the magnetic quantum number. The magnetic moment μ interacts with a magnetic field B_0 as:

$$E = -\mu_z B_0 = \gamma \mathbf{I}_z B_0 = \gamma \hbar m B_0 \tag{4.3}$$

Positioning the spins in a magnetic field induces a precession of the magnetic moment around the direction of the magnetic field (z-axis). The frequency of the precession is called the Larmor frequency ω_0, which is equal to the gyromagnetic ratio multiplied by the magnetic field strength. The introduction of a magnetic field splits the population between 2I+1 different energy levels called Zeeman levels, with an energy difference between levels equal to $\Delta E = \hbar \gamma B$. The application of radio frequency pulses can induce transitions between the energy levels. The Frequency which is needed to transfer a spin between energy levels are $\omega = \Delta E / \hbar = \gamma B$. This is central in NMR spectroscopy as we will see. At equilibrium the populations differences between the Zeeman levels follow the Boltzmann distribution, see equation 4.4.

$$\frac{N_m}{N} = e^{\frac{-E_m}{k_B T}} \Big/ \sum_{m=-I}^{I} e^{\frac{-E_n}{k_B T}} \approx \frac{1}{2I+1} \left(1 + \frac{m \hbar \gamma B_0}{k_B T} \right) \tag{4.4}$$

Where N is the total number or spins, N_m is the number of spins in the mth Zeeman levels, T is the temperature and k_B is Boltzmann constant. In NMR all the microscopic magnetic moments of the spins add up to one macroscopic magnetisation vector which is what is detected in an NMR experiment, hence NMR is called a coherent spectroscopic method. Probably the greatest problem with NMR is the relative insensitivity. Using a magnetic field strength of 14.1 T, a common magnetic field strength in NMR, the difference in populations is about one in ten thousand. The sensitivity can, according to equation 4.4 be increased by enlargement of the population difference between the states, which can be achieved by decreasing the temperature or increasing the energy difference between the states, which

is accomplished by increasing the magnetic field strength of the static field. This explains the pursuit of ever stronger magnets.

The chemical environment around the nucleus will affect the local magnetic field experienced by the nuclei, by either shielding or deshielding the spin from the static magnetic field. This effect, called chemical shift, will slightly change the Larmor frequency by σ, the isotropic shielding constant. This makes it possible to distinguish nuclei of the same type in different position in for example a protein.

$$\omega = -(1 - \sigma)B_0 \tag{4.5}$$

The difference in Larmor frequency arise because of motions of electrons which generates secondary magnetic fields. The electron motion is induced by the static magnetic field. The chemical shift is central in NMR and is reported in ppm which is calculated as the difference to a reference resonance signal from a standard molecule:

$$\delta = \frac{\Omega - \Omega_{ref}}{\omega_0} \times 10^6 = (\sigma_{ref} - \sigma)10^6 \tag{4.6}$$

in which Ω and Ω_{ref} are the offset frequencies. This makes the shift difference independent of magnetic field. An example of chemical shift can be seen in figure 4.1 where a one dimensional spectra have been recorded for a ligand with two ^{19}F nuclei.

4.2 The Bloch Equations

The same year as the discovery of NMR, 1946, Bloch [129] formulated a semi-classical vector model which describe the behaviour of a spin-1/2 nuclei without interactions in a static magnetic field. The macroscopic magnetic moment $\mathbf{M}(\mathrm{t})$ is represented as a vector which can evolve over time. Bloch also introduced two processes to account for the loss of NMR signal. One in which the thermal equilibrium is reinstated by transitions between Zeeman levels back to Boltzmann equilibrium. This process is know as longitudinal relaxation, or spin-lattice relaxation (R_1). The other is the process of loss of magnetization in the transverse plane, the x-y plane, and it is called transverse relaxation, or spin-spin relaxation (R_2) and is a result of all spins experiencing slightly different local magnetic

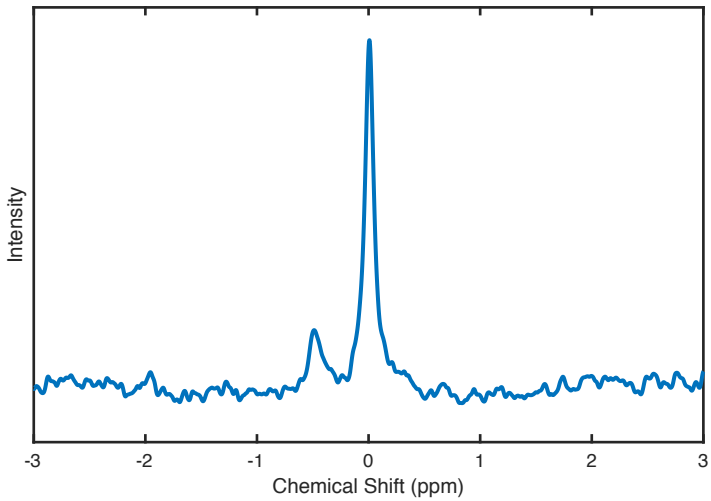


Figure 4.1: ^{19}F NMR spectra of the ligand R, see paper II, bound to Galectin-3C.

field, resulting in an uncertainty in the Zeeman-levels over the sample, meaning they will precess at slightly different frequency resulting in the loss of coherence and the spins will come out of sync. More on relaxation in the comming chapter. The Bloch equations are written as:

$$\begin{aligned}
\frac{dM_x(t)}{dt} &= -\Omega M_y(t) + \omega_1 sin(\phi) M_z(t) - R_2 M_x(t) \\
\frac{dM_y(t)}{dt} &= \Omega M_x(t) + \omega_1 sin(\phi) M_z(t) - R_2 M_y(t) \\
\frac{dM_z(t)}{dt} &= \omega_1 \left[-sin(\phi) M_x(t) + cos(\phi) M_y(t) \right] - R_1 [M_z(t) - M_0]
\end{aligned} \tag{4.7}$$

In which M is the macroscopic bulk magnetisation with component x, y and z. Ω is the offset frequency, ω_1 is the frequency of the applied radio-frequency (rf) field ϕ is the angle in the transverse plane of the applies rf-field, M_0 represents equilibrium magnetization.

4.2.1 Chemical Exchange

The Bloch equations was expanded by McConnell [130] to include chemical exchange, and goes under the name Bloch-McConnell equations, or simply McConnell equations. Chemical exchange can be monitored by NMR if the states have different magnetic environment, hence different chemical shift. The exchange process can be a result of for example a chemical reaction

or a conformational change [54, 131]. The most simple form of chemical exchange is that of a two state process in which the spins exchange between two distinct states, i and j, described as:

$$\mathrm{A_i} \underset{k_{ji}}{\overset{k_{ij}}{\rightleftharpoons}} \mathrm{A_j} \tag{4.8}$$

Where A_i and A_j denotes the different states i and j where k_{ij} is the forward rate going from state i to j and k_{ji} is the reverse rate going from state j to i. The chemical kinetic rate laws can be written as in matrix form as:

$$\frac{d\mathbf{A}(t)}{dt} = \mathbf{K}\mathbf{A}(t) \tag{4.9}$$

For which the matrix elements of $\mathbf{K}$ is given by

$$\begin{aligned} K_{ij} &= k_{ji} \qquad (i \neq j) \\ K_{ii} &= -\sum_{j=1, j\neq i}^{N} k_{ij} \end{aligned} \tag{4.10}$$

where for a two state exchange N is equal to 2, N can however be expanded to include all states needed to describe the exchange process. The expanded Bloch equations in the absence of a rf field becomes:

$$\begin{aligned} \frac{dM_x(t)}{dt} &= -\Omega M_y(t) - R_2 M_x(t) + \sum_{k=1}^{N} K_{jk} M_{kx} \\ \frac{dM_y(t)}{dt} &= \Omega M_x(t) - R_2 M_y(t) + \sum_{k=1}^{N} K_{jk} M_{ky} \\ \frac{dM_z(t)}{dt} &= -R_1[M_z(t) - M_0] + \sum_{k=1}^{N} K_{jk} M_{kz} \end{aligned} \tag{4.11}$$

In paper III we use the Bloch-McConnell equations to characterize the binding of lactose to Galectin-3C using CPMG relaxation dispersions experiments at different lactose concentrations. In this way we can distinguish in what proportions the binding goes through the conformational selection and induced fit pathway of ligand binding.

4.3 The NMR Experiment

Modern NMR experiments are performed using radiofrequency pulses oscillating at or close to the larmor frequency of the nuclei of interest. These kinds of pulses are said to be on resonance and have the ability to rotate the magnetization vector out of the direction of the B_0 field. The most simple NMR experiment consists of just one pulse followed by acquisition of the NMR signal called the free induction decay or FID for short, see figure 4.2. The magnetization starts with the equilibrium state along the z-direction (B_0 field), this is ensured by leaving a delay before the start of the experiment.

$$\mathbf{M}_0 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \tag{4.12}$$

Next a rf-pulse along the y-axis rotates the magnetization 90° into the transverse plane:

$$\mathbf{M}(\tau_p) = \mathbf{R}_y(\alpha)\mathbf{M}_0 = \begin{bmatrix} cos(\alpha) & 0 & sin(\alpha) \\ 0 & 1 & 0 \\ -sin(\alpha) & 0 & cos(\alpha) \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \tag{4.13}$$

Followed by free precession, which is the same as rotation around the z-axis. During the free precession the NMR signal is recorded as the rotating magnetization induces a current in the detection coil.

$$\mathbf{R}_z(t\Omega)\mathbf{M}_0 = \begin{bmatrix} cos(\Omega t) & -sin(\Omega t) & 0 \\ sin(\Omega t) & cos(\Omega t) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} cos(\Omega t) \\ sin(\Omega t) \\ 0 \end{bmatrix} \tag{4.14}$$

During the free precession relaxation of the magnetization back to equilibrium takes place so that the magnetization evolves as:

$$\begin{aligned} M_x(t) &= M_0\, cos(\Omega t)\, exp(-R_2 t) \\ M_y(t) &= M_0\, sin(\Omega t)\, exp(-R_2 t) \\ M_z(t) &= M_0 - M_0\, exp(-R_1 t) \end{aligned} \tag{4.15}$$

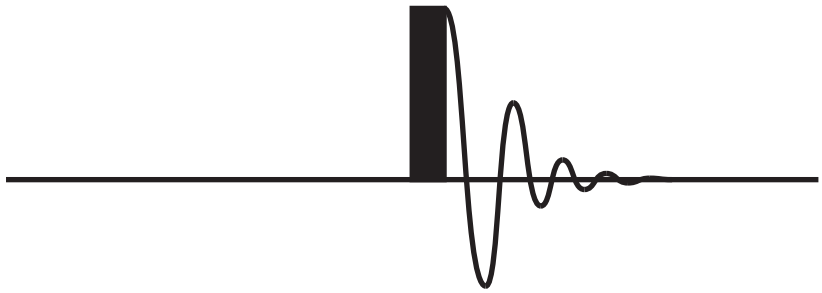


Figure 4.2: A visual representation of the one-dimensional experiment described in section 4.3. The black box represents a 90° electromagnetic pulse followed by acquisition of the FID.

4.4 The Density Operator

In quantum mechanics all knowable information about a system is contained in the wave-function or state vector $|\Psi\rangle$. The state vector can be written as the sum of orthonormal basis functions:

$$|\Psi\rangle = \sum_{n=1}^{N} c_n |n\rangle \tag{4.16}$$

in which $|n\rangle$ are the basis kets, c_n are complex numbers and the vector space have N dimensions. If an operator A, which represents a measurable, acts on the state vector, the expectation value can be written as:

$$\langle A\rangle = \langle\Psi|A|\Psi\rangle = \sum_{nm} c_m^* c_n \langle m| A |n\rangle \tag{4.17}$$

For a basis set $\langle m| A |n\rangle$ are constants and the observable A is given by the product of the coefficients $c_m^* c_n$. The coefficients c can be written in matrix form known as the density matrix (σ) which describes the state of a system and from which the outcome of an experiment can be calculated.

$$\sigma = \sum_{nm} c_m^* c_n \tag{4.18}$$

Diagonal elements of the density matrix are referred to as populations and the off diagonal elements as coherences. The evolution of the density matrix over time is described by the Liouville-von Neumann equation:

$$\frac{d\sigma(t)}{dt} = -i[\mathcal{H}, \sigma(t)] \tag{4.19}$$

Where $\mathcal{H}$ is the Hamiltonian operator which represents the energy of the system and may be time dependent or independent.

Instrumentation

High-resolution NMR spectroscopy requires magnets with high field strengths and a very homogenous magnetic field where the sample is situated in the magnet. This is accomplished using superconducting magnets which are cooled using liquid helium which has a temperature of about 4 K (-269 $^{\circ}C$). The liquid helium is insulated from the surrounding using a vacuum space which is submerged in a bath of liquid nitrogen (77 K or -196 $^{\circ}C$) Inside the coil which makes up the superconducting magnet, around the sample is a set of shim coils. Each of these produces a weak magnetic field and has the job of making the magnetic field in the sample homogenous. In the center of the magnetic coils is a vertical tube called the bore. The probe inserted into the bottom of the bore and it contains coils for applying radio frequency pulses and detecting the signal, and it also serves as a sample holder.

Lets once again use the experiment in figure 4.2 as an example. A radio frequency pulse is created and applied through the coil in the probe on to the sample. The resulting signal produced by the spins in the sample is detected by the detection coil and amplified. The signal is now split into two where one half is mixed with the reference frequency, and the other with the same reference frequency but phase shifted with 90°. By choosing a reference frequency in the middle of the spectral range, and subtracting that from the signal, we reduce the frequency of the signal to a few kHz. The two signals are interpreted as the real and imaginary part of the complex signal. This is called quadrature detection and lets us determine whether the signal oscillates faster or slower compared to the reference frequency hence, determine where in the spectra the peak will appear. At this point the analog signal is converted to a digital signal and saved on a computer. This signal can now be Fourier transformed to produce a spectra.

Pulse Sequences

Pulse sequences were already introduced in its most simple form in section 4.3, where a one pulse experiment can be seen in figure 4.2 and will result in

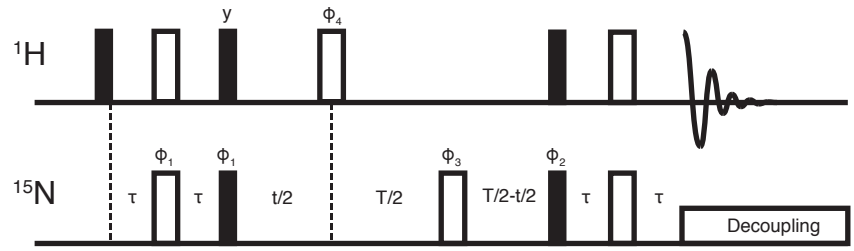


Figure 4.3: A pulse sequence for measuring two dimensional constant time heteronuclear single quantum coherence (HSQC) spectra. The filled black boxes represents 90° pulses and the open boxes represents 180° pulses. All pulses are applied with x-phase if nothing else is indicated. $2\tau = 1/(2J_{HN})$ where J_{HN} is the scalar coupling between the ^{15}N and the ^{1}H. The ϕ indicates phase cycling of the pulse which is presented in [132, 133].

a one dimensional spectra as in figure 4.1. The pulse program is a sequence of pulses and delays which manipulate the spins in a way to get the desired information. They can be as simple as in figure 4.2, but in protein NMR pulse sequences usually includes a large number of pulses and delays and often involves more than one nuclei, for example the pulse sequence in figure 4.3, where we have pulses on both hydrogen and nitrogen. In this experiment we start off on hydrogen and than transfer the magnetization to nitrogen which is possible due to a coupling between the nuclei known as the J-coupling or scalar coupling. The J-coupling make it possible to transfer coherence between different nuclei and increase the dimensionality do get better resultion of the peaks. Before transferring the magnetization back to hydrogen for detection the second dimension is created by incrementing the delay t.

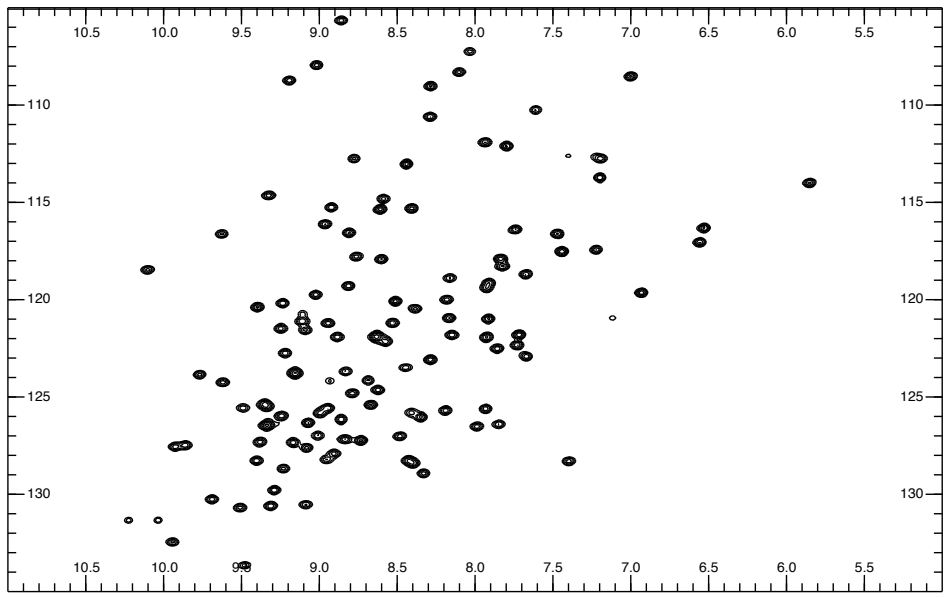


Figure 4.4: ^{15}N HSQC spectra of apo Galectin-3C backbone amides. The x-axis represents the HN dimension in ppm and the y-axis the ^{15}N dimension in ppm.

Chapter 5

NMR Relaxation

Relaxation in NMR is the phenomenon of loss of magnetization coherence in the transverse plane, and the drive of magnetization back to equilibrium along the static magnetic field. NMR relaxation takes place on the time scale of seconds which is slow compared to other techniques utilizing molecular energy levels, such emission spectroscopy which has relaxation times on the low micro second time scale. The relatively slow relaxation in NMR is both a blessing and a curse, it means among other things that the spectroscopist have to wait relatively long between experiments for the magnetisation to reach thermal equilibrium and the next experiment can be started. On the other hand, it provides the possibility to manipulate the magnetization to provide information which would otherwise be unobtainable. By measuring the relaxation rate a wealth of knowledge is accessible, for example internal dynamics of the protein and structural information.

I briefly touched the topic of relaxation in the previous sections on the Bloch equations (section 4.2) and again in the example of the one pulse experiment where the resulting NMR signal is lost by transverse relaxation and the equilibrium magnetization is reinstated through longitudinal relaxation. In this section I will go deeper on he origin of these phenomena. In solution a protein for example, will tumble due to Brownian motion. In this protein the spin experiences a local magnetic field due to other nuclei and electrons. The Brownian motion will cause reorientation of the local field in relation to the static magnetic field, which will cause the spin to experience a time dependent local field fluctuation, which causes relaxation. For a spin = 1/2 the dominant contributions to the fluctuating internal magnetic field is the dipole-dipole interactions and chemical shielding anisotropy (CSA) and for

a spin $> 1/2$ the quadropolar coupling is dominating.

5.1 Spin-Lattice Relaxation

The main source of relaxation in NMR is the coupling to the surroundings also known as the lattice. The lattice modifies the magnetic fields locally so that the nucleus experiences a local magnetic field, B_{local} in addition to the static magnetic field B_0. Following the treatment in Goldman 2001 [134] the local field will have a time dependence due to thermal motion and the Hamiltonian can be written as:

$$\mathcal{H}(t) = \mathcal{H}_0 + \mathcal{H}_1(t) \tag{5.1}$$

Where $\mathcal{H}_0$ is the time independent Hamiltonian and $\mathcal{H}_1(t)$ is the time dependent stochastic Hamiltonian accounting for the coupling to the lattice. This means that the density matrix will evolve over time as:

$$\frac{d\sigma(t)}{dt} = -i[\mathcal{H}_0 + \mathcal{H}_1(t), \sigma(t)] \tag{5.2}$$

By a transformation into the interaction frame which, if the time independent Hamiltonian is solely due to the Zeeman interaction, is synonymous with a rotating frame transformation namely:

$$\tilde{\sigma}(t) = exp(i\mathcal{H}_0 t)\sigma(0)exp(-i\mathcal{H}_0 t) \tag{5.3}$$

This, after skipping a few steps, leads to the master equation for spin-lattice relaxation:

$$\frac{d\tilde{\sigma}}{dt} = -i[\tilde{\mathcal{H}_1}(t), \tilde{\sigma}(0)] - \int_0^t [\tilde{\mathcal{H}_1}(t), [\tilde{\mathcal{H}_1}(t'), \tilde{\sigma}(t')]]dt' \tag{5.4}$$

Next we replace $\tilde{\sigma}(t)$ with $\tilde{\sigma}(t) - \tilde{\sigma}_{eq}$ and expand the Hamiltonian:

$$\mathcal{H}_1(t) = \sum_\alpha V_\alpha F_\alpha(t) = \sum_\alpha V_\alpha^\dagger F_\alpha^*(t) \tag{5.5}$$

where V_α contains spin operators and $F_\alpha(t)$ is a random function or time. In the next step, first taken the ensemble average and since $\tilde{\mathcal{H}_1}(t)$ is zero when averaged, the first term vanishes. The random fluctuation is faster compared to the physical property studied in this case relaxation, which normally takes place on the second time scale. Now let the time scale of the random fluctuation be the overall tumbling of a protein, which is normally

on the nanoseconds time scale (for Galectin-3C 7-8 nanoseconds). Giving $t - t' = \tau_c$, over which a product $F_\alpha(t)F *_\beta (t')$ decay by a substantial amount. Now $t >> \tau_c$, which results in that $\tilde{\sigma}(t')$ can be written as $\tilde{\sigma}(t)$. This also results in that each member of the ensemble have experienced the random process many times τ_c the effect averages out so that we can replace $\overline{\tilde{\sigma}(t)}$ by $\tilde{\sigma}(t)$. Now we can rewrite equation 5.4.

$$\frac{d\tilde{\sigma}(t)}{dt} = -\sum_{\alpha,\beta}\int_0^t [V_\alpha(t), [V_\beta^\dagger(t'), (\tilde{\sigma}(t) - \tilde{\sigma}_{eq})]]\overline{F_\alpha(t)F_\beta^*(t')}dt' \tag{5.6}$$

From this equation we can identify the time correlation function, $C(\tau) = \overline{F_\alpha(t)F_\beta^*(t')}$ in which $\tau = t - t'$. The time correlation function, also known as a memory function, can be described as how long a property of a system persists until it have been averaged out by thermal motion.

5.2 Dipole-Dipole coupling

The dipole-dipole coupling is the interaction between two nuclei. Any nucleus with a spin will generate a magnetic field that will fluctuate as the molecule tumbles. The energy of the interaction between two magnetic dipoles are given by:

$$E = r^{-3}\left[\vec{\mu}_1 \bullet \vec{\mu}_2 - 3(\vec{\mu}_1 \bullet n_{1,2})(\vec{\mu}_2 \bullet n_{1,2})\right] \tag{5.7}$$

In which μ is the magnetic moment for the spins 1 and 2, r is the distance between the nuclei and $n_{1,2}$ is the unit vector connecting the two nuclei. This equation shows that the dipole interaction depends on the distance between the two spins and the orientation relative to the magnetic moments. Both of which can fluctuate over time due to to overall rotation and internal dynamics, which will induce relaxation. Often, the two spins that interacts belong to the same molecule and is covalently bonded so that r is fixed, seen on the NMR time-scale. The magnetic moment is proportional to $\gamma\hbar$ so that the strength of the interaction can be written as:

$$\omega_{DD} = -\sqrt{6}\left(\frac{\mu_0}{4\pi}\right)\frac{\hbar|\gamma^I\gamma^S|}{r^3} \tag{5.8}$$

where μ_0 is the permeability of free space [126, 128].

5.3 Chemical Shielding Anisotropy

The chemical shielding anisotropy (CSA) is the effect of the static magnetic field inducing local magnetic fields through the electrons. These local magnetic field are not uniform in the molecule fixed coordinate system and depending on how the molecule is oriented in relation to B_0, the nuclei will experience different local fields. Hence as the molecule reorients due to Brownian motion the nuclei will experience different local field which causes relaxation. The chemical shielding anisotropy frequency is defined as:

$$\omega_{SA} = \frac{\gamma B_0 \Delta\sigma}{\sqrt{3}} \tag{5.9}$$

In which $\Delta\sigma$ is the chemical shielding anisotropy [126, 128].

5.4 Quadropolar Coupling

In the case where the nuclear spin has a quantum number greater than 1/2 the nuclei possess an electric quadruple moment. It is the divergence from spherical symmetry of the nuclear charge distribution which gives rise to the electric quadruple moment and it is unique to the particular nucleus. The quadrupolar coupling is given by:

$$\omega_Q = \frac{e^2 qQ}{4\hbar I(2I-1)} \tag{5.10}$$

In which e is the elementary charge, eq is the principle value of the electric field gradient tensor, Q is the nuclear quadropolar moment and I is the spin angular momentum quantum number [126, 128].

5.5 The Spectral Density Functions

The Fourier transform of the time correlation function is called the spectral density function, which gives the frequency distribution of the fluctuating magnetic field.

$$j(\omega) = \int_0^\infty \overline{F_\alpha(t)F_\beta(t-\tau)} exp(-i\omega\tau)d\tau = \int_0^\infty C(\tau)exp(-i\omega\tau)d\tau \tag{5.11}$$

in which $F(t) = c_0(t)Y_2^0[\Omega(t)]$ where $c_0(t)$ is a function of physical constants and spatial variables given in sections 5.10, 5.9 and 5.8 and $Y_2^0[\Omega(t)]$ is

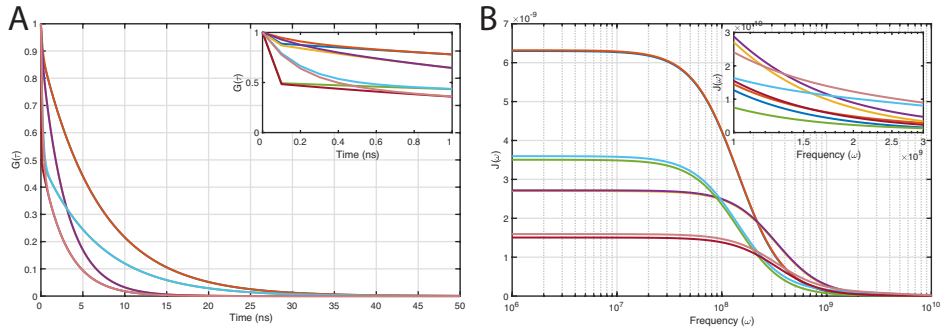


Figure 5.1: Panel A depicts the correlation function and panel B the corresponding spectral density function using the model-free formalism. Blue: $S^2 = 0.9$, $\tau_c = 7\ ns$ and $\tau_e = 10\ ps$. Orange: $S^2 = 0.9$, $\tau_c = 7\ ns$ and $\tau_e = 200\ ps$. Yellow: $S^2 = 0.9$, $\tau_c = 3\ ns$ and $\tau_e = 10\ ps$. Purple: $S^2 = 0.9$, $\tau_c = 3\ ns$ and $\tau_e = 200\ ps$. Green: $S^2 = 0.5$, $\tau_c = 7\ ns$ and $\tau_e = 10\ ps$. Light Blue: $S^2 = 0.5$, $\tau_c = 7\ ns$ and $\tau_e = 200\ ps$. Red: $S^2 = 0.5$, $\tau_c = 3\ ns$ and $\tau_e = 10\ ps$. Pink: $S^2 = 0.5$, $\tau_c = 3\ ns$ and $\tau_e = 200\ ps$.

spherical harmonics and $\Omega(t)$ is the polar angles. So that the stochastic correlation function can be written as

$$C(\tau) = \overline{c_0(t)c_0(t+\tau)Y_2^0[\Omega(t)]Y_2^0[\Omega(t+\tau)]} \tag{5.12}$$

A correlation function for a rigid molecule tumbling in solution $c_0(t) = c_0$ can be written as the orientational correlation function:

$$C_{00}^2(\tau) = \overline{Y_2^0[\Omega(t)]Y_2^0[\Omega(t+\tau)]} = 1/5\ exp[-\tau/\tau_c] \tag{5.13}$$

in which τ_c is the global correlation time, which reports on the time it takes the molecule on average to rotate 1 radian. Now we can write the spectral density function as $j(\omega) = c_0^2 J(\omega)$. where $J(\omega)$ is the orientational spectral density function [126, 135].

$$J(\omega) = Re\left\{\int_0^\infty C_{00}^2(\tau)exp[-i\omega\tau]\right\} = \frac{2}{5}\frac{\tau_c}{(1+\omega^2\tau_c^2)} \tag{5.14}$$

5.5.1 Model-Free Formalism

The model-free formalism takes in addition to the global correlation time, into account the possibility and internal dynamics in terms of motion of individual bonds [136–138]. For isotropic tumbling only one correlation time is needed to describe the Brownian motion. By assuming that the overall and internal motion are independent the total correlation function can be written as a function of the internal and overall:

$$C(t) = C_O(t)C_I(t) \tag{5.15}$$

In which the overall correlation function can be written as in equation 5.13 and the internal correlation function is written as:

$$C_I(t) = S^2 + (1 - S^2)\ exp[-t/\tau_e] \tag{5.16}$$

Where S^2 is the generalized order parameter and τ_e is the internal correlation time. The generalized order parameter reports on the amplitude of motion and makes no assumption on what type of motion that is present. A completely restricted bond vector gives an order parameter of 1 while a bond vector which have the freedom to take what ever angle relative to the rest of the protein gives an order parameter of 0. By combining equations 5.13 and 5.16 we can write the total correlation function as:

$$C(t) = 1/5\left(S^2\ exp(-t/\tau_c) + (1 - S^2)\ exp(-t/\tau)\right. \tag{5.17}$$

in which $\tau^{-1} = \tau_c^{-1} + \tau_e^{-1}$ and the resulting spectral density function is:

$$J(\omega) = \frac{2}{5}\left[\frac{S^2\tau_c}{1 + (\omega\tau_c)^2} + \frac{(1 - S^2)\tau}{1 + (\omega\tau)^2}\right] \tag{5.18}$$

simulated correlation functions and corresponding spectral density functions for the original model-free formalism can be seen in figure 5.1 where the effect changes in the parameters are apparent.

Clore and coworkers expanded the original model-free equations in 1990 after failing to account for the NOE for some residues when fitting relaxation data from SNase and IL-1β using the original model-free equation. The expanded model-free equation include two order parameters (S_f^2 and S_s^2) as well as internal correlation times on two time scales (τ_f and τ_s) which are time scale separated, where f and s denotes fast and slow, respectively [139].

$$J(\omega) = \frac{2}{5}\left[\frac{S^2\tau_c}{1 + (\omega\tau_c)^2} + \frac{(1 - S_f^2)\tau_f'}{1 + (\omega\tau_f')^2} + \frac{(S_f^2 - S^2)\tau_s'}{1 + (\omega\tau_s')^2}\right] \tag{5.19}$$

In which $\tau_i' = (\tau_i\tau_c)/(\tau_c + \tau_i), i = f, s$ and $S^2 = S_f^2 S_s^2$. If the protein diffuses like a sphere meaning all the principle axes in the diffusion tensor are equal, $D_x = D_y = D_z = D$ the spectral density will be as in equation 5.18 and $D = 1/6\tau_c$. If one component differs from the other two, the diffusion is said to be axially symmetric, where $D_\parallel$ is the unique axis and $D_\perp$ represent the other two. In this case the spectral density will be the sum of three components which depends on the vectors angle to the unique axis of the diffusion tensor. If all components in the diffusion tensor are different then the diffusion tensor is called anisotropic, $D_x \neq D_y \neq D_z$. The spectral density will now be the sum of 5 components [135, 140, 141].

Chapter 6

Protein Dynamics

Proteins functionality often relies on the ability to change conformation or to have internal dynamics [142–144]. The conformations of a protein are separated by energy barriers which have to be overcome for the proteins to go from one conformation to an other. The height of the barriers determines how frequent a protein goes between the conformations or on what times scale the conformational changes takes place. The height of the energy barriers can vary over a large range so that conformational changes can take place on time scales from picoseconds for librations and rotations to seconds for unfolding and larger conformational rearrangements [135, 145–147].

NMR is unique in the way that it can probe dynamics on an atomic level and on a wide range of timescales, from bond vibrations and molecular rotations on the picosecond timescale to protein folding and catalysis which can take place in seconds [21, 148–151]. This section is not meant as an exhaustive examination of all NMR methods for characterisation of molecular dynamics but a review of the methods I have used in the articles presented in this thesis.

6.1 Picosecond to Nanosecond Dynamics

As shown in the previous chapter the reason for spin-relaxation is the stochastic fluctuations in the Hamiltonian on a timescale equal to or faster then the overall tumbling of the molecule. The primary interactions in the Hamiltonian that is modulated to cause relaxation is the Dipole-Dipole coupling and the Chemical Shielding Anisotropy for spins = 1/2 and the

quadropolar coupling for spins with a quantum number higher then 1/2 [148].

6.1.1 Backbone Dynamics

In this thesis I have used the N-H bond vector to characterise the backbone dynamics by measuring ^{15}N relaxation rates. It should also be mentioned that the alpha- or carbonyl-carbon can be used to characterise the backbone dynamics [152–155]. The three relaxation rates measured are the longitudinal relaxation rates R_1, the transverse relaxation rate R_2 and the $\{^1H\}$–^{15}N heteronuclear NOE [145, 156]. Since ^{15}N has a spin quantum number of 1/2 the dominant relaxation mechanisms are dipole-dipole and chemical shielding anisotropy. The dipole-dipole interaction scales with the gyromagnetic ratio of the nuclei involved in the interaction and the inverse of the distance to the power of six, which results in that the only nuclei which will contribute significantly to relaxation of the nitrogen of the backbone amide is the protein directly attached to that it, see equation 5.8 [157].

$$R_1 = \frac{1}{4}d^2\left[J(\omega_H - \omega_N) + +3J(\omega_N) + 6J(\omega_H + \omega_N)\right] + c^2 J(\omega_N) \tag{6.1}$$

$$R_2 = \frac{1}{8}d^2\left[4J(0) + J(\omega_H - \omega_N) + 3J(\omega_N) + 6J(\omega_H) + 6J(\omega_H + \omega_N)\right] + \frac{1}{6}c^2\left[4J(0) + 3J(\omega_N)\right] + R_{ex} \tag{6.2}$$

$$NOE = 1 + \frac{d^2}{4R_1}\frac{\gamma_N}{\gamma_H}\left[6J(\omega_H + \omega_N) - J(\omega_H - \omega_N)\right] \tag{6.3}$$

The Relaxation rates are sums of spectral density functions, J(ω), for different combinations of the nitrogen and hydrogen frequencies, ω. The prefactors are given by $d = (\mu_0 h\gamma_N\gamma_H/8\pi^2)\langle 1/r_{NH}^3\rangle$ and $c = \omega_N\Delta\sigma/\sqrt{3}$ in which μ_0 is the permeability of free space, h is Planck's constant, γ is the gyromagnetic ratio of N and H, r_NH is the distance of the bond vector between H and N (1.02 Å), $\Delta\sigma$ is the chemical shielding anisotropy of the nitrogen (-172 ppm) and R_{ex} is the contribution to R_2 from chemical exchange on a micro- to millisecond timescale. For ^{15}N relaxation The CSA tensor is often assumed to be co-linear with the covalent hydrogen nitrogen bond

vector. However, this is not the case but it is only a problem if the diffusion tensor of the molecule significantly deviates from a sphere [158, 159].

In order to interpret the relaxation data we have applied the model-free approach, see section 5.5.1, using the program suite relax [160–162]. Three different global diffusion tensors are often used to account for the global Brownian motion: sphere, axially symmetric and anisotropic. In the sphere all principal axes of the diffusion tensor are equal, for the axially symmetric one axes is unique and in the anisotropic all three axis are different. Ten different models for the internal dynamics is commonly used to explain the relaxation data with the following residue specific parameters:

$$
\begin{aligned}
m0 &= \{\} \\
m1 &= \{S^2\} \\
m2 &= \{S^2,\ \tau_e\} \\
m3 &= \{S^2,\ R_{ex}\} \\
m4 &= \{S^2,\ \tau_e,\ R_{ex}\} \\
m5 &= \{S^2,\ S_f^2,\ \tau_s\} \\
m6 &= \{S^2,\ \tau_f, /S^2,\ \tau_s\} \\
m7 &= \{S^2\ \ S_f^2,\ \tau_s\ \ R_{ex}\} \\
m8 &= \{S^2,\ \tau_f, /S^2,\ \tau_s\ \ R_{ex}\} \\
m9 &= \{R_{ex}\}
\end{aligned}
\tag{6.4}
$$

Relax works in an iterative process where the diffusion tensor is first estimated and fixed followed by fitting of all selected internal models. The internal models are selected for each residue using Akaike's Information Criterion (AIC) [163] and the last step is a global minimization of all parameters. If the model converges the program goes to the next diffusion model, if not the the process goes back to fixing the diffusion tensor and fitting the internal parameters. In the end AIC is used to select diffusion model.

6.1.2 Side-Chain Dynamics

The dynamics for protein side-chains are generally not as simple to probe as the backbone. All side-chains have different composition and have different dynamic modes, j-couplings and interactions. This means there is not

a universal experiment to probe all side-chain dynamics, instead different experiments have been designed for different type of side-chains [164–170]. For certain isolated side-chain nitrogens, experiments designed for backbone nitrogen relaxation can be used to characterise the dynamics, for example the Arginine N_ϵ. Other side-chains I have characterised the dynamics of are methyl groups. This is done using protein which have been expressed in a solution of about 60% D_2O which will yield four different combinations of deuturation in the methyl groups: CH_3, CH_2D, CHD_2 and CD_3. The pulse sequence is designed to select for the CH_2D isotopomer and the relaxation is measured on the deuteron. Since the deuterium have a spin quantum number of one, the relaxation of the deuterium will be dominated by the quadropolar mechanism which simplifies the interpretation of the relaxation rate in terms of model-free parameters. In this kind of experiment five different relaxation rates can be measured:

$$R^Q(D_Z) = \frac{3}{40}\left(\frac{e^2qQ}{\hbar}\right)^2 (J(\omega_D) + 4J(2\omega_D)) \tag{6.5}$$

$$R^Q(3D_Z^2 - 2) = \frac{3}{40}\left(\frac{e^2qQ}{\hbar}\right)^2 (3J(\omega_D))) \tag{6.6}$$

$$R^Q(D_+) = \frac{1}{80}\left(\frac{e^2qQ}{\hbar}\right)^2 (9J(0) + 15J(\omega_D) + 6J(2\omega_D)) \tag{6.7}$$

$$R^Q(D_+D_Z + D_ZD_+) = \frac{1}{80}\left(\frac{e^2qQ}{\hbar}\right)^2 (9J(0) + 3J(\omega_D) + 6J(2\omega_D)) \tag{6.8}$$

$$R^Q(D_+^2) = \frac{3}{40}\left(\frac{e^2qQ}{\hbar}\right)^2 (J(\omega_D) + 3J(2\omega_D)) \tag{6.9}$$

In which $(e^2qQ)/\hbar$ is the quadropolar relaxation constant. Four different kinds of motions are the main contributors to the relaxation of the deuterium in the methyl group. These four processes are: fast methyl spinning which contributes with the factor 1/9 in equations 6.10 and 6.12. Fast side-chain motions such as torsion librations and bond fluctuation which contributes with S_f^2. Slow local dynamics like rotameric transistions, S_S^2 and at last contributor is global tumbling. The slow and fast internal mo-

tion additionally have correlation times, τ_s and τ_f.

$$\begin{aligned} J(\omega) =& \frac{1}{9} S_f^2 S_s^2 \frac{\tau_0}{1+\omega^2\tau_0^2} \ + \ \frac{1}{9} S_f^2 (1-S_s^2) \frac{\tau_1}{1+\omega^2\tau_1^2} \ + \\ & S_s^2 (1-S_f^2) \frac{\tau_2}{1+\omega^2\tau_2^2} \ + \ (1-S_s^2)(1-S_f^2) \frac{\tau_3}{1+\omega^2\tau_3^2} \end{aligned} \tag{6.10}$$

$$\begin{aligned} \tau_0 &= (1/\tau_c)^{-1} \\ \tau_1 &= (1/\tau_c \ + \ 1/\tau_s)^{-1} \\ \tau_2 &= (1/\tau_c \ + \ 1/\tau_f)^{-1} \\ \tau_3 &= (1/\tau_c \ + \ 1/\tau_f \ + \ 1/\tau_s)^{-1} \end{aligned} \tag{6.11}$$

In the cases that the time scale of the slow dynamics is similar to the overall tumbling time, τ_c, the two will be hard to separate. The correlation function is approximated with a single exponential decay with an effective correlation time, τ_c^{eff}, and the spectral density function becomes:

$$J(\omega) = \frac{1}{9} S_f^2 \frac{\tau_c^{eff}}{1+\omega^2(\tau_c^{eff})^2} \ + \ \left(1 - \frac{1}{9} S_f^2\right) \frac{\tau}{1+\omega^2\tau^2} \tag{6.12}$$

in which $1/\tau = 1/\tau_c^{eff} + 1/\tau_f$. If the methyl group lacks slow dynamics the effective correlation times in equation 6.12 becomes the global correlation time. This results in three models for motions in methyl groups with parameters $\{S_f^2, \ \tau_f, \ S_s^2, \ \tau_s\}$, $\{S_f^2, \ \tau_f, \ \tau_c^{eff}\}$ and $\{S_f^2, \ \tau_f\}$.

6.1.3 Entropy from NMR

The generalized order parameter reports on the equilibrium distribution of the orientation of a bond-vector, for example the N-H backbone bond vector [136].

$$S^2 = \frac{4\pi}{5} \sum_{m=-2}^{2} \mid \langle Y_2^m(\theta,\phi) \rangle \mid^2 \tag{6.13}$$

In which Y_2^m are spherical harmonics of the bond vector, where the orientations are defined by θ and ϕ.

$$\langle Y_2^m(\theta,\phi) \rangle = \int_0^{2\pi} \int_0^{\pi} p(\theta,\phi) Y_2^m(\theta,\phi) sin(\theta) d\theta d\phi \tag{6.14}$$

where $p(\theta, \phi)$ is the Boltzmann probability of finding the vector with a certain orientation, (θ, ϕ). The variable p is a function of the number of available states of the bond vector, this means the backbone order parameter can be related to entropy difference as [171, 172], see section 1.1.1

$$\Delta S_{BA} = -R \sum_{j=1}^{N} ln \left[\frac{(1 - S^2_{j,A})}{(1 - S^2_{j,B})} \right] \tag{6.15}$$

Where ΔS_{BA} is the entropy difference going from state B to A, R is the gas constant. For side chains the order parameters can be related to entropy as [172]:

$$S = k_B M \left[A + Bf(1 - S^2) \right] \tag{6.16}$$

Where M denotes the number of dihedral angles, A and B are fit parameters (see reference [172]) and f(x) is either x or ln(x). It should be mentioned that the order parameter only reports on motions on a time scale shorter than the global correlation time. So if the protein has internal motions on a time scale longer than the time scale of global tumbling these will not be included in the entropy calculated using the order parameter. Correlated motions are not accounted for in these equations either. however, through molecular dynamics simulations it has been shown that for the backbone amide correlated motions are limited [173].

6.1.4 NMR and MD simulations

Molecular Dynamics (MD) simulations is a powerful tool to get atomic level descriptions of molecular mechanisms, and in contrast to NMR, all bond vectors are available for study. Using NMR and MD in combination provides an elegant way of explaining molecular phenomenon, where for example NMR order parameters can be calculated from an MD trajectory and used to confirm the validity of the MD simulation [33, 174–177]. The method used for calculating the MD order parameters is iRED, which stands for isotropic reorientational eigenmode dynamics [178].

6.2 Chemical Exchange

Processes on the milli- to microsecond time scales is common in biological macromolecules such as ligand binding [33, 49], allostery [179, 180] and catalysis [181]. These kinds of processes can be characterized using the

Carr-Purcell-Meiboom-Gill (CPMG) pulse sequence if the process transfers the nuclei between different magnetic environments [182–184]. The chemical exchange on these time-scales contributes to an increase to R_2, which can be quenched by applying refocusing pulses, so that at low frequency of refocusing pulses the R_2 is increased due to exchange processes. As the frequency of pulsing increase the R_2 decreases until the effect on R_2 for exchange is negligible, see figure 6.1. The shape of the CPMG curve will in addition to the exchange rate depend on the chemical shift difference between, and the populations of the exchanging species [185,186]. Consider the exchange process of ligand binding as in:

$$P + L \underset{k_1}{\overset{k_{-1}}{\rightleftharpoons}} PL \tag{6.17}$$

Where P is a protein, L is a ligand and PL is the protein ligand complex, k_1 is the rate of association and k_{-1} is the dissociation rate. The association constant can be rewritten as $k_1' = k_1[L]$. The protein and ligand will due to interactions with each other experience different chemical shifts in the free form and bound. The transverse relaxation rate as a function of the time between the refocusing pulses in the CPMG train is described as:

$$R_2^{ex}(\tau_{cp}) = (R_2^f + R_2^b + k_1' + k_{-1}) - (1/\tau_{cp}) ln\lambda^+ \tag{6.18}$$

$$ln\lambda^+ = ln\left[(D_+ cosh^2\xi - D_- cos^2\eta)^{1/2} + (D_+ sinh^2\xi + sin^2\eta)^{1/2}\right] \tag{6.19}$$

$$D_\pm = \frac{1}{2}\left[\pm 1 + (\psi + 2\Delta\omega^2)/(\psi^2 + \zeta^2)^{1/2}\right] \tag{6.20}$$

$$\xi = \frac{\tau_{cp}}{\sqrt{8}}\left[\psi + (\psi^2 + \zeta^2)^{1/2}\right]^{1/2} \tag{6.21}$$

$$\eta = \frac{\tau_{cp}}{\sqrt{8}}\left[-\psi + (\psi^2 + \zeta^2)^{1/2}\right]^{1/2} \tag{6.22}$$

$$\psi = (R_2^f + R_2^b + k_1' + k_{-1})^2 - (\Delta\omega)^2 + 4k_1' k_{-1} \tag{6.23}$$

$$\zeta = 2\Delta\omega(R_2^f + R_2^b + k_1' + k_{-1}) \tag{6.24}$$

Where R_2^f and R_2^b is the transverse relaxation rate in absence of exchange of the exchanging species, in the case of ligand binding, free and bound protein and $\Delta\omega = 2\pi\Delta\nu$ is the chemical shift difference between the species. Using these equations CPMG relaxation dispersions have been simulated

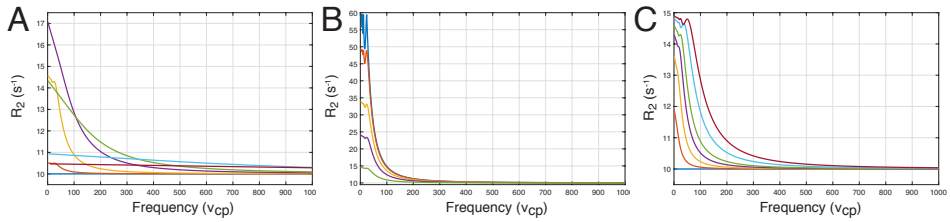


Figure 6.1: Showing simulated CPMG dispersion curves at a static magnetic field strength of 500 MHz when the exchange rate, populations and chemical shift difference is varied. Panel A shows relaxation dispersions using the following chemical exchange rates (s^{-1}): 0 (blue), 10 (red), 100 (yellow), 500 (purple), 1000 (green), 5000 (sky blue), 10000 (dark red). The populations was set to 0.95 and the chemical shift difference to 1 ppm. Panel B shows the effect of changes in population on the relaxation dispersion curve, populations used are 0.5 (blue), 0.6 (red), 0.75 (yellow), 0.85 (purple) an 0.95 (green). The exchange rate was fixed at 100 s^{-1} and the chemical shift difference to 1 ppm. Panel C shows the effect of changes in the chemical shift difference on the relaxation dispersion curve, 0 (blue), 0.25 (red), 0.5 (yellow), 0.75 (purple), 1 (green), 1.5 (sky blue), 2 ppm (dark red). The exchange rate was fixed at 100 s^{-1} and the population to 0.95. for all relaxation dispersion curves $R_{2,0}$, the transverse relaxation rate in absence of exchange, is fix at 10 s^{-1}

with different, exchange rates (panel A), populations (panel B), and chemical shifts (panel C), the resulting curves can be seen in figure 6.1. All the parameters significantly alters the shape of the curve however all parameters can be fitted by recording CPMG relaxation dispersion at multiple magnetic fields.

Chemical exchange can be divided into three regimes depending on the rate of exchange in relation to the chemical shift difference between the exchanging species. When the chemical exchange is much larger compared to the chemical exchange difference the exchange, $k_{ex} \gg \Delta\omega$ is said to be fast. If this i the case only one peak will be visible in the spectra. The position of the peak will be a population average of the two states. So if both states are equally populated the peak will appear half way distance from the chemical shift of to the other. If instead the population are skewed in favour of one the shift of the peak will be close to the dominant species. When the chemical shift difference is similar to the chemical exchange rate, $k_{ex} \sim \Delta\omega$, the exchange process is said to be intermediate. In this regime the single peak from the fast exchange regime is significantly broadened, not unusually beyond detection. If the chemical shift difference is larger compared to the exchange rate, $k_{ex} \ll \Delta\omega$, the exchange is in the slow exchange regime. In this regime the two peaks are visible in the spectra with intensities relative the populations of the species [186].

Chapter 7

Guide to the Papers

7.1 Paper I

In this paper a number of Galectin-3 inhibitors have been synthesised and the affinity to Galectins -1 and -3 have been investigated. Through tuning of the interactions between flourines on the ligands and galectin-3, low nano molar affinity ligands with high selectivity to galectin-3 over galectin-1 was discovered. Because of high binding affinity of the developed ligand a new probe for measuring competitive fluorescence polarization was developed, as well as a ligand to be used in displacement ITC measurements. The thermodynamic profile was determined for ligands with the highest affinity using ITC. All measured ligands have a strong enthalpic term driving the binding of the ligands. Additionally, all ligands display an entropic penalty counteracting the enthalpy. The series of ligands display a general trend of entropy-enthalpy compensation. For two of the asymmetric tiodigalectosides x-ray structures were determined. The difference between the two ligands are the fluorination pattern of a phenyl group, the first ligand has a 3-fluorophenyl while the other has a 3,4,5-trifluorophenyl group. The structures reveal a split occupancy for the mono fluorinated ligand with equal populations of the 3-fluorophenyl close to R186 and R144, while the trifluorophenylated ligand binds with the fluorinatedphenyl in proximety to R144. This leads us to conclude that an increase in the fluorination pattern drives the ligand to one binding pose.

7.2 Paper II

Here we have used a combination of ITC, X-ray crystallography, NMR relaxation and MD simulations in order to try to pinpoint the underlying driving forces of molecular recognition in Galectin-3C binding two diastereomeric ligands. By using these ligands we were able to attribute the difference in binding energetics to the complexes since the ligands have essentially the same chemical potential free in solution. ITC reveal similar binding affinities, 1.0 and 2.1 μM for R and S, respectively, and hence similar Gibbs free energy. Greater differences was observed in entropy and enthalpy which exhibited compensatory behaviour, $-T\Delta\Delta S(R-S) = 3 \pm 1\ kJ/mol$ and $\Delta\Delta H(R-S) = -5 \pm 1\ kJ/mol$.

The x-ray structures show similar binding modes for the two ligands but with small differences in the part close to Arg186, see figure 7.1. The difference in binding pose depends on the hydroxyl group in the stereocenter of the ligands making a hydrogen bond with Glu184. Ensemble refinment of the crystal structures show greater flexibility of the S ligand compered to R. Trying to quantify the entropy from the ensamble refinement resulted in a qualitative result agreeing with other results but the standard errors were greater than the difference between the complexes. NMR relaxation experiments were performed for the ^{15}N amide backbone and for ^{2}H of methyl side-chains and interpreted using the model-free formalism. The backbone order parameters are very similar for the two complexes with significant differences for residues which are not directly situated in the binding pocket. Which indicated that the different stereo chemistry of the ligands affect the mobility of the protein at remotes sites. The differences in methyl order parameters are also small with a few significant differences of which one, Val172, is situated directly in the binding site next to the stereocenter of the ligand. The order parameters were used to validate the MD simulations which is in reasonable agreement.

We used the NMR order parameters to estimate the conformational entropy difference between the two complexes. The backbone conformational entropy is $-T\Delta\Delta S_{bb}(R-S) = 17 \pm 5\ kJ/mol$ and corresponding for methyl groups are $-T\Delta\Delta S_{sc}(R-S) = -5 \pm 6\ kJ/mol$. This results in a total conformational entropy difference estimated from NMR to $-T\Delta\Delta S_{NMR}(R-S) = 12 \pm 8\ kJ/mol$. The conformational entropy calculated from the MD simulation for both the protein and ligand added up to $-T\Delta\Delta S_{MD,conf}(R-S) = 11 \pm 5\ kJ/mol$ which is similar to what was calculated from NMR order parameters. Grid inhomogeneous

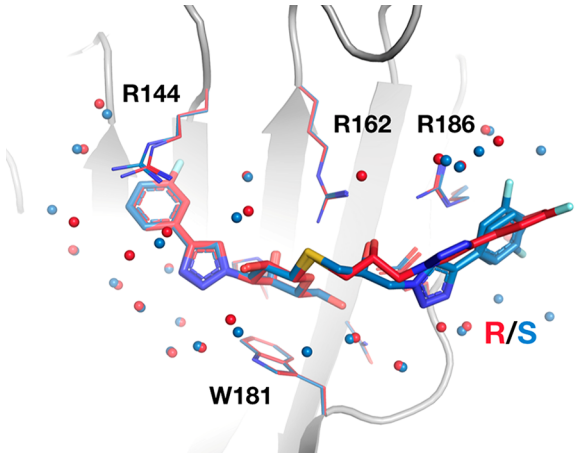


Figure 7.1: X-ray crystal structures of the ligandgalectin-3C complexes overlayed, R - Red, S - Blue. The protein backbone is in cartoon representation gray. Water molecules within 5 Å of the ligands are represented by small spheres.

solvation theory (GIST) calculations reveal similar areas of high solvation density comparing the two complexes, with subtle differences where the ligand pose is different. This results in a solvation entropy difference of $-T\Delta\Delta S_{MD,solv}(R-S) = 3 \pm 2\ kJ/mol$. Taken together the conformational and solvational entropy adds up to $13 \pm 5\ kJ/mol$ which is greater than the total taken from ITC which is $3 \pm 1\ kJ/mol$, but this is not significant to 95% confidence level. These results indicates that conformational entropy dominates over solvational entropy in determining the total entropy difference.

7.3 Paper III

Ligand binding to proteins often involve some kind of conformational change in the protein. Two common model for ligand binding which involves conformational change is induced-fit where the proteins changes conformation after ligand binding to better accommodate the ligand, and conformational selection where the protein samples a conformation which resembles the ligand bound form and the ligand selectively binds this conformation. Apo galectin-3C relaxation dispersion experiments, reveal a second conformation, which resembles the ligand bound form populated to 4%, making conformational selection a possible binding path. In order to unravel which pathway that is dominating in galectin-3C lactose binding we have measured ^{15}N NMR relaxation dispersion experiments on apo galectin-3C, as well as nine lactose concentrations covering a range of protein saturation from 7-75%. By fitting the dispersion data to a four state model, see figure 7.2, we are able to determine the rate constants for both conformational selection and induced fit. The rates reveal a 21 times higher affinity for lactose to state 2 compared to state 1. Despite the sampling of a higher affinity state in apo galectin-3C, the dominant binding pathway for the whole lactose concentration range is induced-fit. The dominance of induced fit can be explained by the higher conformational transition rates between the lactose bound forms of galectin-3C compared to apo, indicating that lactose reduces the barrier between the states.

7.4 Paper IV

In the drug design process one important parameter is the affinity of the target for the target. However, lately the kinetics of ligand binding, and especially the rate of ligand release are getting more focus since it is often a better predictor in vivo efficacy and safety. In this paper we have used ^{15}N relaxation dispersion experiments to study the on- and off-rates of six related ligands with different affinities for galectin-3C. We show that the on-rate is essentially unchanged for all ligands while the off-rate decreases with stronger binding affinity. By analysing the rates and binding affinities using a linear free energy relationship we can get a deeper understanding of the underlying energy landscape. We see that the difference in the lifetime of the complex mainly depends on the difference in free energy of the bound state and less on the height of the transition barrier.

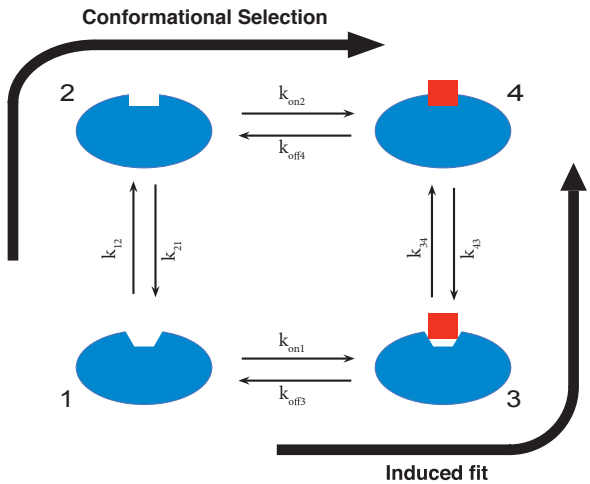


Figure 7.2: Grafic representation of the four state model from paper III where the blue shape represents the protein and the red square a ligand. The large black arrows indicates the pathway of induced fit and conformational selection. The thin black arrows indicates transition possible for each state of the protein with correpsonding transistion rates k.

Bibliography

[1] G. G. Hammes, S. J. Benkovic, and S. Hammes-Schiffer, "Flexibility, Diversity, and Cooperativity: Pillars of Enzyme Catalysis," *Biochemistry*, vol. 50, pp. 10422–10430, dec 2011.

[2] D. J. Vocadlo, G. J. Davies, R. Laine, and S. G. Withers, "Catalysis by hen egg-white lysozyme proceeds via a covalent intermediate," *Nature*, vol. 412, no. 6849, pp. 835–838, 2001.

[3] R. V. Weatherman, R. J. Fletterick, and T. S. Scanlan, "Nuclear-receptor ligands and ligand-binding domains," *Annual Review of Biochemistry*, vol. 68, no. 1, pp. 559–581, 1999. PMID: 10872460.

[4] M. Martini, M. C. D. Santis, L. Braccini, F. Gulluni, and E. Hirsch, "Pi3k/akt signaling pathway and cancer: an updated review," *Annals of Medicine*, vol. 46, no. 6, pp. 372–383, 2014.

[5] D. E. Clapham, "Calcium signaling," *Cell*, vol. 80, pp. 259–268, jan 1995.

[6] P. Visca, L. Leoni, M. J. Wilson, and I. L. Lamont, "Iron transport and regulation, cell signalling and genomics: lessons from escherichia coli and pseudomonas," *Molecular Microbiology*, vol. 45, no. 5, pp. 1177–1190, 2002.

[7] N. MacAulay, S. Hamann, and T. Zeuthen, "Water transport in the brain: Role of cotransporters," *Neuroscience*, vol. 129, no. 4, pp. 1029–1042, 2004.

[8] J. M. . T. Berg J. L. ; Stryer, L., *Biochemistry*. Freeman, 6th ed., 2007.

[9] V. Ramakrishnan, "Ribosome structure and the mechanism of translation," 2002.

[10] C. J. Oldfield and A. K. Dunker, "Intrinsically disordered proteins and intrinsically disordered protein regions," *Annual Review of Biochemistry*, vol. 83, no. 1, pp. 553–584, 2014.

[11] J. Sevcik, Z. Dauter, V. S. Lamzin, and K. S. Wilson, "Ribonuclease from *Streptomyces aureofaciens* at Atomic Resolution," *Acta Crystallographica Section D*, vol. 52, pp. 327–344, Mar 1996.

[12] Y.-L. Zhang and Z.-Y. Zhang, "Low-Affinity Binding Determined by Titration Calorimetry Using a High-Affinity Coupling Ligand: A Thermodynamic Study of Ligand Binding to Protein Tyrosine Phosphatase 1B," *Analytical Biochemistry*, vol. 261, pp. 139–148, aug 1998.

[13] H.-X. Zhou and M. K. Gilson, "Theory of Free Energy and Entropy in Noncovalent Binding," *Chemical Reviews*, vol. 109, pp. 4092–4107, sep 2009.

[14] P. W. Atkins and J. De Paula, *Atkins' Physical chemistry*. Oxford; New York: Oxford University Press, 2006.

[15] A. Cooper and C. M. Johnson, *Introduction to Microcalorimetry and Biomolecular Energetics*, pp. 109–124. Totowa, NJ: Humana Press, 1994.

[16] G. Klebe, *Drug Design: Methodology, concepts, and mode-of-action*. Springer, 2013.

[17] A. C. A. Roque, "Ligand-macromolecular interactions in drug discovery," *Clifton, NJ 2010; p*, vol. 572, 2010.

[18] K. Dill and S. Bromberg, *Molecular Driving Forces: Statistical Thermodynamics in Biology, Chemistry, Physics and Nanoscience*. second edi ed., 2011.

[19] B. Wienen-Schmidt, T. Wulsdorf, H. R. A. Jonker, K. Saxena, D. Kudlinzki, V. Linhard, S. Sreeramulu, A. Heine, H. Schwalbe, and G. Klebe, "On the Implication of Water on Fragment-to-Ligand Growth in Kinase Binding Thermodynamics," *ChemMedChem*, vol. 13, no. 18, pp. 1988–1996, 2018.

[20] G. A. Holdgate, A. Tunnicliffe, W. H. Ward, S. A. Weston, G. Rosenbrock, P. T. Barth, I. W. Taylor, R. A. Pauptit, and D. Timms, "The entropic penalty of ordered water accounts for weaker binding of the

antibiotic novobiocin to a resistant mutant of dna gyrase: a thermodynamic and crystallographic study," *Biochemistry*, vol. 36, no. 32, pp. 9663–9673, 1997.

[21] C. Diehl, O. Engström, T. Delaine, M. Håkansson, S. Genheden, K. Modig, H. Leffler, U. Ryde, U. J. Nilsson, and A. M., "Protein Flexibility and Conformational Entropy in Ligand Design Targeting the Carbohydrate Recognition Domain of Galectin-3," *J. AM. CHEM. SOC.*, vol. 132, no. 41, pp. 14577–14589, 2010.

[22] C. Diehl, S. Genheden, K. Modig, U. Ryde, and M. Akke, "Conformational entropy changes upon lactose binding to the carbohydrate recognition domain of galectin-3," *J Biomol NMR*, vol. 45, no. 1-2, pp. 157–169, 2009.

[23] S.-R. Tzeng and C. G. Kalodimos, "Protein activity regulation by conformational entropy," *Nature*, vol. 488, no. 7410, pp. 236–240, 2012.

[24] K. H. DuBay and P. L. Geissler, "Calculation of Proteins' Total Side-Chain Torsional Entropy and Its Influence on Protein–Ligand Interactions," *Journal of Molecular Biology*, vol. 391, pp. 484–497, aug 2009.

[25] L. Amzel, "Calculation of entropy changes in biological processes: Folding, binding, and oligomerization," *Methods in Enzymology*, vol. 323, pp. 167–177, jan 2000.

[26] A. V. Finkelstein and J. Janin, "The price of lost freedom: entropy of bimolecular complex formation," *Protein Engineering, Design and Selection*, vol. 3, pp. 1–3, 10 1989.

[27] R. Lumry and S. Rajender, "Enthalpy–entropy compensation phenomena in water solutions of proteins and small molecules: A ubiquitous properly of water," *Biopolymers*, vol. 9, no. 10, pp. 1125–1227, 1970.

[28] A. Cooper, C. M. Johnson, J. H. Lakey, and M. Nöllmann, "Heat does not come in different colours: Entropy-enthalpy compensation, free energy windows, quantum confinement, pressure perturbation calorimetry, solvation and the multiple causes of heat capacity effects in biomolecular interactions," *Biophysical Chemistry*, 2001.

[29] M. S. Ramadan, D. F. Evans, R. Lumry, and S. Philson, "Micelle formation in hydrazine-water mixtures," *The Journal of Physical Chemistry*, vol. 89, pp. 3405–3408, jul 1985.

[30] G. Sugihara and M. Hisatomi, "Enthalpy–Entropy Compensation Phenomenon Observed for Different Surfactants in Aqueous Solution," *Journal of Colloid and Interface Science*, vol. 219, pp. 31–36, nov 1999.

[31] J. R. Beasley, D. F. Doyle, L. Chen, D. S. Cohen, B. R. Fine, and G. J. Pielak, "Searching for quantitative entropy-enthalpy compensation among protein variants," *Proteins: Structure, Function, and Bioinformatics*, vol. 49, pp. 398–402, nov 2002.

[32] B. Lee, "Solvent reorganization contribution to the transfer thermodynamics of small nonpolar molecules," *Biopolymers*, vol. 31, pp. 993–1008, jul 1991.

[33] 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*, vol. 141, pp. 2012–2026, feb 2019.

[34] T. S. G. Olsson, J. E. Ladbury, W. R. Pitt, and M. A. Williams, "Extent of enthalpy–entropy compensation in protein–ligand interactions," *Protein Science*, vol. 20, pp. 1607–1618, sep 2011.

[35] K. Peterson, R. Kumar, O. Stenström, P. Verma, P. Verma, M. Håkansson, B. Kahl-Knutsson, F. Zetterberg, H. Leffler, M. Akke, D. Logan, and U. Nilsson, "Systematic Tuning of Fluoro-galectin-3 Interactions Provides Thiodigalactoside Derivatives with Single-Digit nM Affinity and High Selectivity," *Journal of Medicinal Chemistry*, vol. 61, no. 3, 2018.

[36] E. M. Raggett, G. Bainbridge, L. J. A. Evans, A. Cooper, and J. H. Lakey, "Discovery of critical Tol A-binding residues in the bactericidal toxin colicin N: a biophysical approach," *Molecular Microbiology*, vol. 28, pp. 1335–1343, jun 1998.

[37] V. Lafont, A. A. Armstrong, H. Ohtaka, Y. Kiso, L. Mario Amzel, and E. Freire, "Compensating Enthalpic and Entropic Changes Hinder Binding Affinity Optimization," *Chemical Biology & Drug Design*, vol. 69, pp. 413–422, jun 2007.

[38] B. Breiten, M. R. Lockett, W. Sherman, S. Fujita, M. Al-Sayah, H. Lange, C. M. Bowers, A. Heroux, G. Krilov, and G. M. Whitesides, "Water Networks Contribute to Enthalpy/Entropy Compensation in Protein–Ligand Binding," *Journal of the American Chemical Society*, vol. 135, pp. 15579–15584, oct 2013.

[39] E. Grunwald and C. Steel, "Solvent Reorganization and Thermodynamic Enthalpy-Entropy Compensation," *Journal of the American Chemical Society*, vol. 117, pp. 5687–5692, may 1995.

[40] M. V. Rekharsky, T. Mori, C. Yang, Y. H. Ko, N. Selvapalam, H. Kim, D. Sobransingh, A. E. Kaifer, S. Liu, and L. Isaacs, "A synthetic host-guest system achieves avidin-biotin affinity by overcoming enthalpy–entropy compensation," *Proceedings of the National Academy of Sciences*, vol. 104, no. 52, pp. 20737–20742, 2007.

[41] C.-E. Chang and M. K. Gilson, "Free Energy, Entropy, and Induced Fit in HostGuest Recognition: Calculations with the Second-Generation Mining Minima Algorithm," *Journal of the American Chemical Society*, vol. 126, pp. 13156–13164, oct 2004.

[42] J. D. Dunitz, "Win some, lose some: enthalpy-entropy compensation in weak intermolecular interactions," *Chemistry & Biology*, vol. 2, pp. 709–712, nov 1995.

[43] E. Fischer, "Einfluß der configuration auf die wirkung der enzyme," *Ber. Dtsch. Chem. Ges.*, vol. 27, p. 2985–2993, 1894.

[44] O. F. Lange, N.-A. Lakomek, C. Farès, G. F. Schröder, K. F. A. Walter, S. Becker, J. Meiler, H. Grubmüller, C. Griesinger, and B. L. de Groot, "Recognition Dynamics Up to Microseconds Revealed from an RDC-Derived Ubiquitin Ensemble in Solution," *Science*, vol. 320, pp. 1471 LP – 1475, jun 2008.

[45] S. Brasselet, E. J. G. Peterman, A. Miyawaki, and W. E. Moerner, "Single-Molecule Fluorescence Resonant Energy Transfer in Calcium Concentration Dependent Cameleon," *The Journal of Physical Chemistry B*, vol. 104, pp. 3676–3682, apr 2000.

[46] P. Hinterdorfer and Y. F. Dufrêne, "Detection and localization of single molecular recognition events using atomic force microscopy," *Nature Methods*, vol. 3, no. 5, pp. 347–355, 2006.

[47] H. Tidow and P. Nissen, "Structural diversity of calmodulin binding to its target sites," *The FEBS Journal*, vol. 280, pp. 5551–5565, nov 2013.

[48] B. Ma, M. Shatsky, H. J. Wolfson, and R. Nussinov, "Multiple diverse ligands binding at a single protein site: A matter of pre-existing populations," *Protein Science*, vol. 11, no. 2, pp. 184–197, 2002.

[49] K. Chakrabarti, R. Agafonov, F. Pontiggia, R. Otten, M. Higgins, G. Schertler, D. Oprian, and D. Kern, "Conformational Selection in a Protein-Protein Interaction Revealed by Dynamic Pathway Analysis," *Cell Reports*, vol. 14, pp. 32–42, jan 2016.

[50] K. A. Henzler-Wildman, V. Thai, M. Lei, M. Ott, M. Wolf-Watz, T. Fenn, E. Pozharski, M. A. Wilson, G. A. Petsko, M. Karplus, C. G. Hübner, and D. Kern, "Intrinsic motions along an enzymatic reaction trajectory," *Nature*, vol. 450, p. 838, nov 2007.

[51] E. Kim, S. Lee, A. Jeon, J. M. Choi, H.-S. Lee, S. Hohng, and H.-S. Kim, "A single-molecule dissection of ligand binding to a protein with intrinsic dynamics," *Nature Chemical Biology*, vol. 9, p. 313, mar 2013.

[52] J. Monod, J. Wyman, and J.-P. Changeux, "On the nature of allosteric transitions: A plausible model," *Journal of Molecular Biology*, vol. 12, no. 1, pp. 88–118, 1965.

[53] D. E. Koshland, E. Henry, and J. Hofrichter, "Application of a Theory of Enzyme Specificity to Protein Synthesis," *Proceedings of the National Academy of Sciences*, vol. 44, pp. 98–104, feb 1958.

[54] D. D. Boehr, D. McElheny, H. J. Dyson, and P. E. Wright, "The Dynamic Energy Landscape of Dihydrofolate Reductase Catalysis," *Science*, vol. 313, pp. 1638 LP – 1642, sep 2006.

[55] J. Gsponer, J. Christodoulou, A. Cavalli, J. M. Bui, B. Richter, C. M. Dobson, and M. Vendruscolo, "A Coupled Equilibrium Shift Mechanism in Calmodulin-Mediated Signal Transduction," *Structure*, vol. 16, pp. 736–746, may 2008.

[56] M. Ghoneim and M. Spies, "Direct Correlation of DNA Binding and Single Protein Domain Motion via Dual Illumination Fluorescence Microscopy," *Nano Letters*, vol. 14, pp. 5920–5931, oct 2014.

[57] F. Paul, F. Noé, and T. R. Weikl, "Identifying Conformational-Selection and Induced-Fit Aspects in the Binding-Induced Folding of PMI from Markov State Modeling of Atomistic Simulations," *The Journal of Physical Chemistry B*, vol. 122, pp. 5649–5656, may 2018.

[58] J. Dogan, S. Gianni, and P. Jemth, "The binding mechanisms of intrinsically disordered proteins," *Physical Chemistry Chemical Physics*, vol. 16, no. 14, pp. 6323–6331, 2014.

[59] R. Grünberg, J. Leckner, and M. Nilges, "Complementarity of Structure Ensembles in Protein-Protein Binding," *Structure*, vol. 12, pp. 2125–2136, dec 2004.

[60] T. Wlodarski and B. Zagrovic, "Conformational selection and induced fit mechanism underlie specificity in noncovalent interactions with ubiquitin," *Proceedings of the National Academy of Sciences*, vol. 106, pp. 19346 LP – 19351, nov 2009.

[61] G. G. Hammes, Y.-C. Chang, and T. G. Oas, "Conformational selection or induced fit: a flux description of reaction mechanism.," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 106, pp. 13737–41, aug 2009.

[62] J. P. Hughes, S. Rees, S. B. Kalindjian, and K. L. Philpott, "Principles of early drug discovery," *British Journal of Pharmacology*, vol. 162, pp. 1239–1249, mar 2011.

[63] J.-P. Renaud, C.-w. Chung, U. H. Danielson, U. Egner, M. Hennig, R. E. Hubbard, and H. Nar, "Biophysics in drug discovery: impact, challenges and opportunities," *Nature Reviews Drug Discovery*, vol. 15, p. 679, aug 2016.

[64] T. L. Blundell, H. Jhoti, and C. Abell, "High-throughput crystallography for lead discovery in drug design," *Nature Reviews Drug Discovery*, vol. 1, no. 1, pp. 45–54, 2002.

[65] G. E. de Kloe, D. Bailey, R. Leurs, and I. J. de Esch, "Transforming fragments into candidates: small becomes big in medicinal chemistry," *Drug Discovery Today*, vol. 14, pp. 630–646, jul 2009.

[66] C. A. Lipinski, F. Lombardo, B. W. Dominy, and P. J. Feeney, "Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings," *Advanced Drug Delivery Reviews*, vol. 23, pp. 3–25, jan 1997.

[67] G. Klebe, "Applying thermodynamic profiling in lead finding and optimization," *Nature Reviews Drug Discovery*, vol. 14, p. 95, jan 2015.

[68] E. Freire, "Do enthalpy and entropy distinguish first in class from best in class?," *Drug Discov Today*, vol. 13, no. 19-20, pp. 869–874, 2008.

[69] S. H. Barondes, V. Castronovo, D. Cooper, R. D. Cummings, K. Drickamer, T. Feizi, M. A. Gitt, J. Hirabayashi, C. Hughes, and K.-i. Kasai, "Galectins: a family of animal beta-galactoside-binding lectins.," *Cell*, vol. 76, no. 4, pp. 597–8, 1994.

[70] J. Dumic, S. Dabelic, and M. Flögel, "Galectin-3: An open-ended story," *Biochimica et Biophysica Acta (BBA)*, vol. 1760, pp. 616–635, 2006.

[71] J. Hirabayashi and K.-i. Kasai, "The family of metazoan metal-independent β-galactoside-binding lectins: structure, function and molecular evolution," *Glycobiology*, vol. 3, pp. 297–304, 08 1993.

[72] H. Leffler, S. Carlsson, M. Hedlund, Y. Qian, and F. Poirier, "Introduction to galectins," *Glycoconjugate J.*, vol. 19, pp. 433–440, 2004.

[73] A. Raz, G. Pazerini, and P. Carmi, "Identification of the metastasis-associated, galactoside-binding lectin as a chimeric gene product with homology to an ige-binding protein," *Cancer Research*, vol. 49, no. 13, pp. 3489–3493, 1989.

[74] J. Seetharaman, A. Kanigsberg, R. Slaaby, H. Leffler, S. H. Barondes, and J. M. Rini, "X-ray Crystal Structire of the Human Galectin-3 Carbohydrate Recognition Domain at 2.1-Å Resolution," *J. Biol. Chem.*, vol. 273, pp. 13047–13052, 1998.

[75] S. M. Massa, D. N. W. Cooper, H. Leffler, and S. H. Barondes, "L-29, an endogenous lectin, binds to glycoconjugate ligands with positive cooperativity," *Biochemistry*, vol. 32, no. 1, pp. 260–267, 1993.

[76] J. Ochieng, B. Green, S. Evans, O. James, and P. Warfield, "Modulation of the biological functions of galectin-3 by matrix metalloproteinases," *Biochimica et Biophysica Acta (BBA) - General Subjects*, vol. 1379, no. 1, pp. 97 – 106, 1998.

[77] T. J. Flotte, T. Springer, and G. Thorbecke, "Dendritic cell and macrophage staining by monoclonal antibodies in tissue sections and

epidermal sheets.," *The American journal of pathology*, vol. 111, no. 1, p. 112, 1983.

[78] M. M. Lotz, C. W. Andrews, C. A. Korzelius, E. C. Lee, G. D. Steele, A. Clarke, and A. M. Mercurio, "Decreased expression of mac-2 (carbohydrate binding protein 35) and loss of its nuclear localization are associated with the neoplastic progression of colon carcinoma.," *Proceedings of the National Academy of Sciences*, vol. 90, no. 8, pp. 3466–3470, 1993.

[79] S. K. Gupta, S. Masinick, M. Garrett, and L. D. Hazlett, "Pseudomonas aeruginosa lipopolysaccharide binds galectin-3 and other human corneal epithelial proteins.," *Infection and immunity*, vol. 65, no. 7, pp. 2747–2753, 1997.

[80] S. Heilmann, T. Hummel, F. Margolis, M. Kasper, and M. Witt, "Immunohistochemical distribution of galectin-1, galectin-3, and olfactory marker protein in human olfactory epithelium," *Histochemistry and cell biology*, vol. 113, no. 3, pp. 241–245, 2000.

[81] Q. Bao and R. C. Hughes, "Galectin-3 expression and effects on cyst enlargement and tubulogenesis in kidney epithelial mdck cells cultured in three-dimensional matrices in vitro," *Journal of cell science*, vol. 108, no. 8, pp. 2791–2800, 1995.

[82] M. KASPER and R. C. HUGHES, "Immunocytochemical evidence for a modulation of galectin 3 (mac-2), a carbohydrate binding protein, in pulmonary fibrosis," *The Journal of pathology*, vol. 179, no. 3, pp. 309–316, 1996.

[83] D. M. S. Villa-Verde, E. Silva-Monteiro, M. G. Jasiulionis, D. A. Farias-de Oliveira, R. R. Brentani, W. Savino, and R. Chammas, "Galectin-3 modulates carbohydrate-dependent thymocyte interactions with the thymic microenvironment," *European journal of immunology*, vol. 32, no. 5, pp. 1434–1444, 2002.

[84] V. Castronovo, F. A. Van den Brûle, P. Jackers, N. Clausse, F.-T. LIU, C. Gillet, and M. E. Sobel, "Decreased expression of galectin-3 is associated with progression of human breast cancer," *The Journal of pathology*, vol. 179, no. 1, pp. 43–48, 1996.

[85] C. Schaffert, P. M. Pour, and W. G. Chaney, "Localization of galectin-3 in normal and diseased pancreatic tissue," *International Journal of Pancreatology*, vol. 23, pp. 1–9, Feb 1998.

[86] I. K. Moutsatsos, J. M. Davis, and J. L. Wang, "Endogenous lectins from cultured cells: subcellular localization of carbohydrate-binding protein 35 in 3t3 fibroblasts.," *The Journal of Cell Biology*, vol. 102, no. 2, pp. 477–483, 1986.

[87] L. Wang, H. Inohara, K. Pienta, and A. Raz, "Galectin-3 is a nuclear matrix protein which binds rna," *Biochemical and Biophysical Research Communications*, vol. 217, no. 1, pp. 292 – 303, 1995.

[88] M. Hubert, S.-Y. Wang, J. L. Wang, A.-P. Sève, and J. Hubert, "Intranuclear distribution of galectin-3 in mouse 3t3 fibroblasts: Comparative analyses by immunofluorescence and immunoelectron microscopy," *Experimental Cell Research*, vol. 220, no. 2, pp. 397 – 406, 1995.

[89] D. Weinmann, K. Schlangen, S. André, S. Schmidt, S. M. Walzer, B. Kubista, R. Windhager, S. Toegel, and H.-J. Gabius, "Galectin-3 Induces a Pro-degradative/inflammatory Gene Signature in Human Chondrocytes, Teaming Up with Galectin-1 in Osteoarthritis Pathogenesis," *Scientific Reports*, vol. 6, p. 39112, dec 2016.

[90] Z. Cao, N. Said, S. Amin, H. K. Wu, A. Bruce, M. Garate, D. K. Hsu, I. Kuwabara, F.-T. Liu, and N. Panjwani, "Galectins-3 and-7, but not galectin-1, play a role in re-epithelialization of wounds," *Journal of Biological Chemistry*, vol. 277, no. 44, pp. 42299–42305, 2002.

[91] D. Yang and L. E. Kay, "Contributions to Conformational Entropy Arising from Bond Vector Fluctuations Measured from NMR-Derived Order Parameters: Application to Protein Folding," *J. Mol. Biol.*, vol. 263, pp. 369–382, 1996.

[92] T. Fukumori, Y. Takenaka, N. Oka, T. Yoshii, V. Hogan, H. Inohara, H.-o. Kanayama, H.-R. C. Kim, and A. Raz, "Endogenous galectin-3 determines the routing of cd95 apoptotic signaling pathways," *Cancer Research*, vol. 64, no. 10, pp. 3376–3379, 2004.

[93] S. F. Dagher, J. L. Wang, and R. J. Patterson, "Identification of galectin-3 as a factor in pre-mrna splicing," *Proceedings of the National Academy of Sciences*, vol. 92, no. 4, pp. 1213–1217, 1995.

[94] K. Fritsch, M. Mernberger, A. Nist, T. Stiewe, A. Brehm, and R. Jacob, "Galectin-3 interacts with components of the nuclear ribonucleoprotein complex," *BMC Cancer*, vol. 16, no. 1, p. 502, 2016.

[95] H.-M. Lin, R. G. Pestell, A. Raz, and H.-R. C. Kim, "Galectin-3 enhances cyclin D1 promoter activity through SP1 and a cAMP-responsive element in human breast epithelial cells," *Oncogene*, vol. 21, no. 52, pp. 8001–8010, 2002.

[96] I. Paron, A. Scaloni, A. Pines, A. Bachi, F.-T. Liu, C. Puppin, M. Pandolfi, L. Ledda, C. D. Loreto, G. Damante, and G. Tell, "Nuclear localization of galectin-3 in transformed thyroid cells: a role in transcriptional regulation," *Biochemical and Biophysical Research Communications*, vol. 302, no. 3, pp. 545 – 553, 2003.

[97] J. Ochieng, V. Furtak, and P. Lukyanov, "Extracellular functions of galectin-3," *Glycoconjugate Journal*, vol. 19, pp. 527–535, Jan 2002.

[98] R. A. Pacis, M. J. Pilat, K. J. Pienta, K. Wojno, A. Raz, V. Hogan, and C. R. Cooper, "Decreased galectin-3 expression in prostate cancer," *The Prostate*, vol. 44, no. 2, pp. 118–123, 2000.

[99] A. Danguy, I. Camby, and R. Kiss, "Galectins and cancer," *Biochimica et Biophysica Acta (BBA) - General Subjects*, vol. 1572, pp. 285–293, sep 2002.

[100] F. van den Brûle, S. Califice, and V. Castronovo, "Expression of galectins in cancer: A critical review," *Glycoconjugate Journal*, vol. 19, no. 7, pp. 537–542, 2002.

[101] F.-T. Liu and G. A. Rabinovich, "Galectins as modulators of tumour progression.," *Nature reviews. Cancer*, vol. 5, no. 1, pp. 29–41, 2005.

[102] U. C. Sharma, S. Pokharel, T. J. van Brakel, J. H. van Berlo, J. P. Cleutjens, B. Schroen, S. Andre, H. J. Crijns, H. J. Gabius, J. Maessen, and Y. M. Pinto, "Galectin-3 marks activated macrophages in failure-prone hypertrophied hearts and contributes to cardiac dysfunction," *Circulation*, vol. 110, no. 19, pp. 3121–3128, 2004.

[103] M. M. Pierce, C. Raman, and B. T. Nall, "Isothermal Titration Calorimetry of Protein–Protein Interactions," *Methods*, vol. 19, pp. 213–221, oct 1999.

[104] R. O'Brien, J. E. Ladbury, and B. Z. Chowdhry, "Isothermal titration calorimetry of biomolecules," in *Protein-Ligand Interactions: hydrodynamics and calorimetry* (S. E. Harding and B. Z. Chowdhry, eds.), ch. 10, pp. 263–286, Oxford: Oxford University press, 2001.

[105] T. Wiseman, S. Williston, J. F. Brandts, and L.-N. Lin, "Rapid measurement of binding constants and heats of binding using a new titration calorimeter," *Analytical Biochemistry*, vol. 179, pp. 131–137, may 1989.

[106] L. Freiburger, K. Auclair, and A. Mittermaier, "Global ITC fitting methods in studies of protein allostery," *Methods*, vol. 76, pp. 149–161, 2015.

[107] A. Velázquez Campoy and E. Freire, "ITC in the post-genomic era...? Priceless," *Biophysical Chemistry*, vol. 115, no. 2-3, pp. 115–124, 2005.

[108] A. Hansen and L. Kay, "Quantifying millisecond time-scale exchange in proteins by CPMG relaxation dispersion NMR spectroscopy of side-chain carbonyl groups," *J Biomol NMR*, vol. 50, no. 4, pp. 347–355, 2011.

[109] M. L. Doyle and P. Hensley, "[5] Tight ligand binding affinities determined from thermodynamic linkage to temperature by titration calorimetry," *Methods in Enzymology*, vol. 295, pp. 88–99, jan 1998.

[110] M. L. Doyle, G. Louie, P. R. Dal Monte, and T. D. Sokoloski, "[8] Tight binding affinities determined from thermodynamic linkage to protons by titration calorimetry," *Methods in Enzymology*, vol. 259, pp. 183–194, jan 1995.

[111] J. Tellinghuisen, "Isothermal titration calorimetry at very low c," *Analytical Biochemistry*, vol. 373, pp. 395–397, feb 2008.

[112] A. Velazquez-Campoy and E. Freire, "Isothermal titration calorimetry to determine association constants for high-affinity ligands," *Nature Protocols*, vol. 1, no. 1, pp. 186–191, 2006.

[113] B. W. Sigurskjold, "Exact Analysis of Competition Ligand Binding by Displacement Isothermal Titration Calorimetry," *Analytical Biochemistry*, vol. 277, pp. 260–266, jan 2000.

[114] D. G. Myszka, Y. N. Abdiche, F. Arisaka, O. Byron, E. Eisenstein, P. Hensley, J. A. Thomson, C. R. Lombardo, F. Schwarz, W. Stafford, and M. L. Doyle, "The ABRF-MIRG'02 study: assembly state, thermodynamic, and kinetic analysis of an enzyme/inhibitor interaction," *Journal of biomolecular techniques : JBT*, vol. 14, pp. 247–269, dec 2003.

[115] I. Wadsö, "Needs for standards in isothermal microcalorimetry," *Thermochimica Acta*, vol. 347, pp. 73–77, apr 2000.

[116] L. Baranauskienė, V. Petrikaitė, J. Matulienė, and D. Matulis, "Titration Calorimetry Standards and the Precision of Isothermal Titration Calorimetry Data," 2009.

[117] L. S. Mizoue and J. Tellinghuisen, "The role of backlash in the"first injection anomaly"in isothermal titration calorimetry," *Analytical biochemistry*, vol. 1, no. 326, pp. 125–127, 2004.

[118] E. M. Purcell, H. C. Torrey, and R. V. Pound, "Resonance Absorption by Nuclear Magnetic Moments in a Solid," *Physical Review*, vol. 69, pp. 37–38, jan 1946.

[119] F. Bloch, W. W. Hansen, and M. Packard, "Nuclear Induction," *Physical Review*, 1946.

[120] M. A. Balafar, A. R. Ramli, M. I. Saripan, and S. Mashohor, "Review of brain MRI image segmentation methods," *Artificial Intelligence Review*, vol. 33, no. 3, pp. 261–274, 2010.

[121] P. Brambilla, A. Hardan, S. U. di Nemi, J. Perez, J. C. Soares, and F. Barale, "Brain anatomy and development in autism: review of structural MRI studies," *Brain Research Bulletin*, vol. 61, pp. 557–569, oct 2003.

[122] T. K. Karamanos, V. Tugarinov, and G. M. Clore, "Unraveling the structure and dynamics of the human DNAJB6b chaperone by NMR reveals insights into Hsp40-mediated proteostasis," *Proceedings of the National Academy of Sciences*, p. 201914999, oct 2019.

[123] M. T. Colvin, R. Silvers, Q. Z. Ni, T. V. Can, I. Sergeyev, M. Rosay, K. J. Donovan, B. Michael, J. Wall, S. Linse, and R. G. Griffin, "Atomic Resolution Structure of Monomorphic Aβ42 Amyloid Fibrils," *Journal of the American Chemical Society*, vol. 138, pp. 9663–9674, aug 2016.

[124] K. Sugase, H. J. Dyson, and P. E. Wright, "Mechanism of coupled folding and binding of an intrinsically disordered protein," *Nature*, vol. 447, p. 1021, may 2007.

[125] H. Koss, M. Rance, and A. G. Palmer, "General Expressions for Carr–Purcell–Meiboom–Gill Relaxation Dispersion for N -Site Chemical Exchange," *Biochemistry*, vol. 57, pp. 4753–4763, aug 2018.

[126] J. Cavenagh and M. Akke, "May the driving force be with you - whatever it is," *Nat. struct. Biol.*, vol. 7, pp. 11–13, 2000.

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

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

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

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

[131] M. J. Grey, C. Wang, and A. G. Palmer, "Disulfide Bond Isomerization in Basic Pancreatic Trypsin Inhibitor: Multisite Chemical Exchange Quantified by CPMG Relaxation Dispersion and Chemical Shift Modeling," *Journal of the American Chemical Society*, vol. 125, pp. 14324–14335, nov 2003.

[132] J. Santoro and G. C. King, "A constant-time 2D overbodenhausen experiment for inverse correlation of isotopically enriched species," *Journal of Magnetic Resonance (1969)*, vol. 97, pp. 202–207, mar 1992.

[133] G. W. Vuister and A. Bax, "Measurement of two-bond JCOHα coupling constants in proteins uniformly enriched with13C," *Journal of Biomolecular NMR*, vol. 2, no. 4, pp. 401–405, 1992.

[134] M. Goldman, "Formal Theory of Spin–Lattice Relaxation," *Journal of Magnetic Resonance*, vol. 149, no. 2, pp. 160–187, 2001.

[135] 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, no. 5, pp. 1624–1671, 2006.

[136] G. Lipari and A. Szabo, "Model-Free Approach to the Interpretation of Nuclear Magnetic Resonance Relaxation in Macromolecules. 1. Theory and Range of Validity," *J. AM. CHEM. SOC.*, vol. 104, pp. 4546–4559, 1982.

[137] G. Lipari and A. Szabo, "Model-Free Approach to he Interpretation of Nuclear Magnetic Resonance Relaxation in Mactomolecules.

2. Analysis of Experimental Results," *J. AM. CHEM. SOC.*, vol. 104, pp. 4559–4570, 1982.

[138] 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.

[139] G. M. Clore, A. Szabo, A. Bax, L. E. Kay, P. C. Driscoll, and A. M. Gronenborn, "Deviations from the simple two-parameter model-free approach to the interpretation of nitrogen-15 nuclear magnetic relaxation of proteins," *Journal of the American Chemical Society*, vol. 112, no. 12, pp. 4989–4991, 1990.

[140] 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.

[141] D. E. Woessner, "Nuclear Spin Relaxation in Ellipsoids Undergoing Rotational Brownian Motion," *The Journal of Chemical Physics*, vol. 37, pp. 647–654, aug 1962.

[142] A. T. Namanja, X. J. Wang, B. Xu, A. Y. Mercedes-Camacho, K. A. Wilson, F. A. Etzkorn, and J. W. Peng, "Stereospecific gating of functional motions in Pin1," *Proceedings of the National Academy of Sciences*, vol. 108, pp. 12289–12294, jul 2011.

[143] D. M. Francis, B. Różycki, D. Koveal, G. Hummer, R. Page, and W. Peti, "Structural basis of p38α regulation by hematopoietic tyrosine phosphatase," *Nat Chem Biol*, vol. 7, no. 12, pp. 916–924, 2011.

[144] A. V. Zhuravleva, D. M. Korzhnev, E. Kupce, A. S. Arseniev, M. Billeter, and V. Y. Orekhov, "Gated Electron Transfers and Electron Pathways in Azurin: A NMR Dynamic Study at Multiple Fields and Temperatures," *J Mol Biol*, vol. 342, no. 5, pp. 1599–1611, 2004.

[145] N. A. Farrow, R. Muhandiram, A. U. Singer, S. M. Pascal, C. M. Kay, G. Gish, S. E. Shoelson, T. Pawson, J. D. Forman-Kay, and L. E. Kay, "Backbone Dynamics of a Free and a Phosphopeptide-Complexed Src Homology 2 Domain Studied by 15N NMR Relaxation," *Biochemistry*, vol. 33, no. 19, pp. 5984–6003, 1994.

[146] Y. Ivarsson, C. Travaglini-Allocatelli, P. Jemth, F. Malatesta, M. Brunori, and S. Gianni, "An on-pathway intermediate in the folding of a PDZ domain.," *The Journal of biological chemistry*, vol. 282, pp. 8568–72, mar 2007.

[147] M. Onitsuka, H. Kamikubo, Y. Yamazaki, and M. Kataoka, "Mechanism of induced folding: Both folding before binding and binding before folding can be realized in staphylococcal nuclease mutants," *Proteins: Structure, Function, and Bioinformatics*, vol. 72, pp. 837–847, aug 2008.

[148] A. G. I. I. I. Palmer, "NMR Characterization of the Dynamics of Biomacromolecules," *Chemical Reviews*, vol. 104, no. 8, pp. 3623–3640, 2004.

[149] K. A. Henzler-Wildman, M. Lei, V. Thai, S. J. Kerns, M. Karplus, and D. Kern, "A hierarchy of timescales in protein dynamics is linked to enzyme catalysis," *Nature*, vol. 450, p. 913, nov 2007.

[150] M. Wolf-Watz, V. Thai, K. Henzler-Wildman, G. Hadjipavlou, E. Z. Eisenmesser, and D. Kern, "Linkage between dynamics and catalysis in a thermophilic-mesophilic enzyme pair," *Nature Structural & Molecular Biology*, vol. 11, no. 10, pp. 945–949, 2004.

[151] M. L. Gill, R. A. Byrd, and I. I. I. Palmer Arthur G., "Dynamics of GCN4 facilitate DNA interaction: a model-free analysis of an intrinsically disordered region," *Physical Chemistry Chemical Physics*, vol. 18, no. 8, pp. 5839–5849, 2016.

[152] H. Sun, D. Long, R. Brüschweiler, and V. Tugarinov, "Carbon Relaxation in 13Cα–Hα and 13Cα–Dα Spin Pairs as a Probe of Backbone Dynamics in Proteins," *The Journal of Physical Chemistry B*, vol. 117, pp. 1308–1320, feb 2013.

[153] A. G. Palmer, M. Rance, and P. E. Wright, "Intramolecular motions of a zinc finger DNA-binding domain from Xfin characterized by proton-detected natural abundance carbon-13 heteronuclear NMR spectroscopy," *Journal of the American Chemical Society*, vol. 113, pp. 4371–4380, jun 1991.

[154] K. T. Dayie and G. Wagner, "Carbonyl carbon probe of local mobility in 13C, 15N-enriched proteins using high-resolution nuclear magnetic resonance," *Journal of the American Chemical Society*, vol. 119, no. 33, pp. 7797–7806, 1997.

[155] J. Engelke and H. Rüterjans, "Backbone dynamics of proteins derived from carbonyl carbon relaxation times at 500, 600 and 800 mhz: Application to ribonuclease t1," *Journal of Biomolecular NMR*, vol. 9, pp. 63–78, Jan 1997.

[156] G. N. B. Yip and E. R. P. Zuiderweg, "A phase cycle scheme that significantly suppresses offset-dependent artifacts in the R2-CPMG 15N relaxation experiment," *Journal of Magnetic Resonance*, vol. 171, no. 1, pp. 25–36, 2004.

[157] L. E. Kay, D. A. Torchia, and A. Bax, "Backbone dynamics of proteins as studied by nitrogen-15 inverse detected heteronuclear nmr spectroscopy: application to staphylococcal nuclease," *Biochemistry*, vol. 28, no. 23, pp. 8972–8979, 1989.

[158] 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, no. 2, pp. 139–147, 1999.

[159] Y. Hiyama, C. H. Niu, J. V. Silverton, A. Bavoso, and D. A. Torchia, "Determination of 15N chemical shift tensor via 15N-2H dipolar coupling in Boc-glycylglycyl[15N glycine]benzyl ester," *Journal of the American Chemical Society*, vol. 110, no. 8, pp. 2378–2383, 1988.

[160] E. D'Auvergne and P. Gooley, "The use of model selection in the model-free analysis of protein dynamics," *J Biomol NMR*, vol. 25, no. 1, pp. 25–39, 2003.

[161] E. D'Auvergne and P. Gooley, "Optimisation of NMR dynamic models II. A new methodology for the dual optimisation of the model-free parameters and the Brownian rotational diffusion tensor," *J Biomol NMR*, vol. 40, no. 2, pp. 121–133, 2008.

[162] E. D'Auvergne and P. Gooley, "Optimisation of NMR dynamic models I. Minimisation algorithms and their performance within the model-free and Brownian rotational diffusion spaces," *J Biomol NMR*, vol. 40, no. 2, pp. 107–119, 2008.

[163] H. Akaike, "A new look at the statistical model identification," *Automatic Control, IEEE Transactions on*, vol. 19, no. 6, pp. 716–723, 1974.

[164] O. Millet, D. R. Muhandiram, N. R. Skrynnikov, and L. E. Kay, "Deuterium Spin Probes of Side-Chain Dynamics in Proteins. 1. Measurement of Five Relaxation Rates per Deuteron in 13C-Labeled and Fractionally 2H-Enriched Proteins in Solution," *Journal of the American Chemical Society*, vol. 124, no. 22, pp. 6439–6448, 2002.

[165] D. R. Muhandiram, T. Yamazaki, B. D. Sykes, and L. E. Kay, "Measurement of 2H T1 and T1.rho. Relaxation Times in Uniformly 13C-Labeled and Fractionally 2H-Labeled Proteins in Solution," *Journal of the American Chemical Society*, vol. 117, no. 46, pp. 11536–11544, 1995.

[166] C. Jin, J. J. Prompers, and R. Brüschweiler, "Cross-correlation suppressed t1 and noe experiments for protein side-chain 13ch2 groups," *Journal of Biomolecular NMR*, vol. 26, pp. 241–247, Jul 2003.

[167] K. Pervushin, G. Wider, and K. Wüthrich, "Deuterium Relaxation in a Uniformly 15N-Labeled Homeodomain and Its DNA Complex1," *Journal of the American Chemical Society*, vol. 119, pp. 3842–3843, apr 1997.

[168] D. Yang, A. Mittermaier, Y.-K. Mok, and L. E. Kay, "A study of protein side-chain dynamics from new 2h auto-correlation and 13c cross-correlation nmr experiments: application to the n-terminal sh3 domain from drk11edited by p. e. wright," *Journal of Molecular Biology*, vol. 276, no. 5, pp. 939 – 954, 1998.

[169] J. Boyd, "Measurement of 15n relaxation data from the side chains of asparagine and glutamine residues in proteins," *Journal of Magnetic Resonance, Series B*, vol. 107, no. 3, pp. 279 – 285, 1995.

[170] J. Iwahara, Y.-S. Jung, and G. M. Clore, "Heteronuclear NMR Spectroscopy for Lysine NH3 Groups in Proteins: Unique Effect of Water Exchange on 15N Transverse Relaxation," *Journal of the American Chemical Society*, vol. 129, pp. 2971–2980, mar 2007.

[171] M. Akke, N. J. Skelton, J. Kördel, A. G. I. I. I. Palmer, and W. J. Chazin, "Effects of Ion Binding on the Backbone Dynamics of Calbindin D9k Determined by 15N NMR Relaxation," *Biochemistry*, vol. 32, pp. 9832–9844, 1993.

[172] D.-W. Li and R. Brüschweiler, "A Dictionary for Protein Side-Chain Entropies from NMR Order Parameters," *Journal of the American Chemical Society*, vol. 131, no. 21, pp. 7226–7227, 2009.

[173] J. J. Prompers and R. Brüschweiler, "Thermodynamic Interpretation of NMR Relaxation Parameters in Proteins in the Presence of Motional Correlations," *The Journal of Physical Chemistry B*, vol. 104, no. 47, pp. 11416–11424, 2000.

[174] M. Philippopoulos, A. M. Mandel, A. G. Palmer III, and C. Lim, "Accuracy and precision of NMR relaxation experiments and MD simulations for characterizing protein dynamics," *Proteins: Structure, Function, and Bioinformatics*, vol. 28, pp. 481–493, aug 1997.

[175] D. C. Chatfield, A. Szabo, and B. R. Brooks, "Molecular Dynamics of Staphylococcal Nuclease: Comparison of Simulation with 15N and 13C NMR Relaxation Data," *Journal of the American Chemical Society*, vol. 120, pp. 5301–5311, jun 1998.

[176] 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.

[177] D. A. Case, "Molecular Dynamics and NMR Spin Relaxation in Proteins," *Acc. Chem. Res.*, vol. 35, pp. 325–331, 2002.

[178] 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*, vol. 124, pp. 4522–4534, apr 2002.

[179] S. Brüschweiler, P. Schanda, K. Kloiber, B. Brutscher, G. Kontaxis, R. Konrat, and M. Tollinger, "Direct Observation of the Dynamic Process Underlying Allosteric Signal Transmission," *Journal of the American Chemical Society*, vol. 131, no. 8, pp. 3063–3068, 2009.

[180] N. Popovych, S. Sun, R. H. Ebright, and C. G. Kalodimos, "Dynamically driven protein allostery," *Nature Structural & Molecular Biology*, vol. 13, no. 9, pp. 831–838, 2006.

[181] E. Z. Eisenmesser, O. Millet, W. Labeikovsky, D. M. Korzhnev, M. Wolf-Watz, D. A. Bosco, J. J. Skalicky, L. E. Kay, and D. Kern, "Intrinsic dynamics of an enzyme underlies catalysis," *Nature*, vol. 438, no. 7064, pp. 117–121, 2005.

[182] J. P. Loria, M. Rance, and A. G. Palmer, "A relaxation-compensated Carr Purcell Meiboom Gill sequence for characterizing chemical exchange by NMR spectroscopy," *Journal of the American Chemical Society*, vol. 121, no. 10, pp. 2331–2332, 1999.

[183] H. Y. Carr and E. M. Purcell, "Effects of diffusion on free precession in nuclear magnetic resonance experiments," *Physical Review*, vol. 94, no. 3, pp. 630–638, 1954.

[184] S. Meiboom and D. Gill, "Modified spin-echo method for measuring nuclear relaxation times," *Review of Scientific Instruments*, vol. 29, no. 8, pp. 688–691, 1958.

[185] I. R. Kleckner and M. P. Foster, "An introduction to NMR-based approaches for measuring protein dynamics," *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*, vol. 1814, no. 8, pp. 942–968, 2011.

[186] C. Charlier, S. F. Cousin, and F. Ferrage, "Protein dynamics from nuclear magnetic relaxation," *Chemical Society Reviews*, vol. 45, no. 9, pp. 2410–2422, 2016.

Scientific publications

My contributions to the papers

Co-authors are abbreviated as follows:
Carl Diehl (CD), Kristofer Modig (KM), Mikael Akke (MA), Magdalena Riad (MR), Sven Wernersson (SW).

Paper I: Systematic Tuning of Fluoro-galectin-3 Interactions Provides Thiodigalactoside Derivatives with Single-Digit nM Affinity and High Selectivity

I developed a protocol to perform competitive ITC on Galectin-3C. I also performed all the ITC experiments and did all the data fitting to the thermograms. I also contributed to the writing of the article.

Paper II: Interplay between Conformational Entropy and Solvation Entropy in ProteinLigand Binding

I did the ITC experiments and fitting of the thermogram and analysis. I wrote the program for fitting the ITC thermograms. I also performed the NMR experiments, assigned the protein spectra and analysed the data. I also contributed in writing the paper.

Paper III: Resolving Conformational Selection and Induced Fit Pathways of Protein-Ligand Binding Using NMR Relaxation Dispersion

CD and I did the NMR experiments. I did some rewriting of script originally written by KM and MA. I analysed the data and together with MA and KM wrote the paper.

Paper IV: Linear Free Energy Relationships of Ligand Binding Kinetics and Affinity Determined by Protein NMR Relaxation Dispersion

I did together with CD all experiments, I analysed the data and contributed in writing the paper together with MA, CD and KM.

This is not a novel, short story or fable, so if that is what you are looking for, please put this book down before you get disappointed. It is not a book containing recipes (even if that would be fun to make one day), not about how to renovate a boat, nor is it a memoir, even though it contains parts of the authors life. This book contains a few facts, some scientific results, and perhaps one or two speculations.

Yes, you guessed: this is a PhD thesis. If you are interested in theology, economics or the social sciences, this is not the book for you. So what kind of thesis is this? It is the kind containing strange words you don't hear every day, figures of proteins and equations with signs you might not have seen before. Now, the majority of humanity is probably deterred from ever opening this book, but just to exclude a few more, here's another scary word: ENTROPY.

So, if you for some reason received this book, but have no plan whatsoever to read it, except for perhaps the acknowledgements, I have a few suggestions as to what you could do with it. First of all, if you are mentioned in the acknowledgements and for some reason want to keep it, I suggest tearing it out together with the cover, and staple them together and save. This way, it is much easier to bring when you are moving. Otherwise, it should work just fine as kindle, or perhaps it could be used to prop up something, or as a paperweight. For those of you who do plan to read the book, or parts of it, I wish you happy reading!

Printed by Media-Tryck, Lund 2019 NORDIC SWAN ECOLABEL 3041 0903

ISBN 978-91-7422-706-2

Biophysical Chemistry
Department of Chemistry
Faculty of Engineering (LTH)
Lund University