STRUCTURAL CHARACTERIZATION OF PROTEIN-LIGAND COMPLEXES

USING DISSOLUTION DNP ASSISTED NMR SPECTROSCOPY

A Dissertation

by

YUNYI WANG

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Christian Hilty
Committee Members,	Simon W. North
	David H. Russell
	Jim Ji
Head of Department,	Simon W. North

May 2020

Major Subject: Chemistry

ABSTRACT

Structural information on protein-ligand complexes is of key interest in pharmaceutical research, as it provides rational guidance for optimizing ligand affinity and selectivity. An efficient tool for probing protein-ligand interactions is nuclear magnetic resonance (NMR) spectroscopy. Detailed structural information related to the binding can be obtained via the observation of polarization transfer by the nuclear Overhauser effect (NOE). In this dissertation, hyperpolarization of nuclear spins by dissolution dynamic nuclear polarization (D-DNP) is combined with both ligand-observed and protein-observed NMR to enhance the sensitivity of the NOE detection. A hyperpolarized NMR experiment is developed for detecting NOEs between two ligand spins by monitoring intra-ligand polarization transfer. The measurement of a complete intra-ligand NOE build-up curve obtained from a single hyperpolarization allows the detection of binding. Cross-relaxation rates are determined between ligand proton spins, which contain distance information. In addition to the signal enhancement provided by D-DNP, the efficiency in the intra-ligand NOE measurement is increased with protein immobilization on large bead particles. This ligand-observed D-DNP approach may in the future be used for probing protein-ligand interactions in natural environments with live cells. A protein-observed NMR method based on detecting hyper-polarization transfer from ligand to protein is designed for characterizing specific inter-molecular interactions. Fast acquisition of intermolecular NOEs, followed by combining the experimental data with computational docking, allows the determination of molecu-lar structure in protein-ligand complexes. To resolve and identify specific atom-to-atom NOEs, 1D ^{1}H NMR spectroscopy is coupled with a ^{13}C single-quantum coherence selec-tion on the protein side and a selective inversion on ligand resonances. The resolution of NOE measurements is further improved through fast multi-dimensional NMR spec-

troscopy with Hadamard encoding of ligand signals. The detected intermolecular NOE contacts are used to score ligand poses generated by a docking program. A ligand structure in the active site of the protein dihydrofolate reductase is determined using experimental NOEs from a total of four DNP NMR experiments.

ACKNOWLEDGMENTS

My sincere gratitude goes first and formost for my graduate advisor, Dr. Christian Hilty, for his continued support in guiding me through each step of my Ph.D study. He introduced me to the field of NMR and helping me specialize in it. This dissertation would not have been possible without his patience in guiding me finding solutions whenever I encounter problems in the research.

I would like to thank my committee members, Dr. David Russell, Dr. Simon North, and Dr. Jim Ji for their great support and invaluable suggestions.

I would also like to thank all the members of the Hilty research group. Regularly exchanging opinions with them let me develop deeper understanding of our research field and always brought me great ideas for solving problems in the experiments.

Lastly, I would like to express my deep sense of gratitude for my parents for their unconditional support and love over the years. I want to also thank my boyfriend, Mingyuan Wang, for standing by me through both the good and bad times.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supervised by a dissertation committee consisting of Professor Christian Hilty, Professor David H. Russell, Professor Simon W. North of the Department of Chmistry and Professor Jim Ji of the Department of Electrical and Computer Engineering.

In Section 2, Dr. Mukundan Ragavan helped preparation of the isotope-labelled protein. In Section 4, Jihyun Kim assisted pulse sequence optimization and running the experiments.

All other work conducted for the dissertation was completed by the student under the supervision of the committee chair, Professoer Christian Hilty.

Funding Sources

NOMENCLATURE

1D	One-Dimensional
2D	Two-Dimensional
3D	Three-Dimensional
CORCEMA	Complete Relxation and Conformational Exchange Matrix
CSP	Chemical Shift Perturbation
D-DNP	Dissolution Dynamic Nuclear Polarization
DD	Dipole-Dipole
DHFR	Dihydrofolate Reductase
DMSO	Dimethyl Sulfoxide
DNP	Dynamic Nuclear Polarization
E. Coli	Escherichia Coli
HABA	4'-Hydroyazobenzene-2-carboxylic Acid
ILOE	Interligand Nuclear Overhauser Effect
INPHARMA	Interligand NOE for Pharmacophore Mapping
NMR	Nuclear Magnetic Resonance
NOE	Nuclear Overhauser Effect
NOESY	Nuclear Overhauser Effect Spectroscopy
OD	Optical Density
PDB	Protein Data Bank
PRE	Paramagnetic Relaxation Enhancement

PCS	Pseudocontact shift
RDC	Residual Dipolar Coupling
SBDD	Structure-Based Drug Design
SOFAST	Band-Selective Optimized-Flip-Angle Short-Triansent
RMSD	Root-Mean-Square Deviation
RDC	Residual Dipolar Coupling
SALMON	Solvent Accessibility, Ligand Binding, and Mapping of Ligand Orientation by NMR Spectroscopy
STD	Saturation Transfer Difference
THF	Tetrahydrofolic Acid
MTX	Methotrexate
TrNOE	Transferred NOE
UV-Vis	Ultraviolet-Visble
WaterLOGSY	Water-Ligand Observed via Gradient Spectroscopy

TABLE OF CONTENTS

Page

LIST OF FIGURES

FIGURE   Page

LIST OF TABLES

1. INTRODUCTION

1.1   Structure-Based Drug Design

In the past two decades, the exceptional advances in molecular biology and structure elucidation techniques, including X-ray crystallography and nuclear magnetic resonance (NMR) spectroscopy, have dramatically grown the number of three-dimensional (3D) protein structures. Until 2019, approximately 150,000 protein-only structures have been released in the Protein Data Bank, with about 10,000 structures deposited per year.[1] The increasing availability of target protein structures has led to the emergence and rapid development of a rational approach to drug discovery. Structure-based drug design (SBDD) exploits the structural information of the therapeutically relevant target to design novel drugs with improved selectivity and affinity.[2–4]

A typical SBDD process starts with the structure determination of the target protein, followed by multiple cycles of ligand screening, validation, and optimization.[5] Compounds or fragments of compounds are screened experimentally or *in silico* to identify those with some degree of binding affinity and selectivity.[6,7] The promising ligands are further evaluated and modified for better potency and selectivity in the hit-to-lead optimization process. Based on the known 3D structure of the target protein, efforts are focused on obtaining information about the specific ligand binding mode in the binding site. This information is crucial for every iteration of the optimization process by providing the structural basis for optimizing the ligand to increase interaction with the protein.

For several decades, crystallography has been the primary source of the most detailed structural information for protein-ligand complexes.[8,9] Nonetheless, producing high-quality crystals of ligand-protein complexes can be a labor-intensive task without guarantee for success, in spite of advances in high-throughput production of crystals by

1

automation.[10–13] In the situation when the crystallization is difficult, which particularly occurs in the early stages of SBDD when weakly binding ligands are involved, solution NMR spectroscopy offers another option for solving the structure of the complex in the binding site.[14–16] Structural restraints of ligand-protein binding can be obtained from NMR observables such as chemical shifts, the nuclear Overhauser effect (NOE), or paramagnetic effects.[14] Determination of the complex structure in the binding site can be accomplished based only on the NMR data if a substantial amount of structural restraints are obtained.[17] This, however, requires time-consuming processes of acquisition of spectra and subsequent data analysis. In addition to crystallography and NMR spectroscopy, computational methods, such as molecular docking, can be incorporated for the determination of the complex structure. Partial reliance on computation has become desirable considering the difficulties and costs of the experimental methods.[18,19] In the protein-ligand docking, a large number of ligand poses are generated and evaluated by predicting the binding energy, to select the most favorable binding mode of the ligand in the binding pocket. However, establishing the appropriate search algorithms and scoring functions to efficiently and accurately find the correct ligand pose in different cases remains a challenge. Recently, the combination of NMR techniques and computational docking has been proposed as a fast and reliable method for determining ligand binding modes.[20,21] Even sparse NMR data can provide enough structural restraints to select the correct ligand binding mode by filtering the docked posed or by guiding the docking process.

1.2 NMR Spectroscopy in Structural Characterization of Ligand-Protein Binding

1.2.1 Basic NMR Theory

In the presence of an external magnetic field, the interaction of the nuclear spin magnetic moment and the applied magnetic field results in the splitting of spin states

with the energy difference proportional to the field strength. The energy splitting is associated with a difference in the number of spins for the different spin states, which can be quantified by the nuclear spin polarization P. For spin $-1/2$ nuclei, the resulting two Zeeman energy levels are separated by the energy difference ΔE as

$$\Delta E = \hbar \gamma B_0. \tag{1.1}$$

B_0 is the magnetic field strength, γ is the gyromagnetic ratio, and $\hbar$ is the reduced Plank constant. At equilibrium, the spin polarization can be calculated by the Boltzmann distribution,

$$P = \frac{n_{-1/2} - n_{1/2}}{n_{-1/2} + n_{1/2}} = \frac{1 - e^{-\Delta E/(k_B T)}}{1 + e^{-\Delta E/(k_B T)}} = \tanh\left(\frac{\Delta E}{2k_B T}\right) = \tanh\left(\frac{\hbar \gamma B_0}{2k_B T}\right), \tag{1.2}$$

where $n_{-1/2}$ and $n_{1/2}$ represent the numbers of spins for the two spin states, respectively. T is the temperature, and k_B is the Boltzmann constant. The signal measured in an NMR experiment is directly determined by the bulk nuclear magnetization in the sample, which is proportional to both the spin density and the spin polarization.

1.2.2 Effects of Ligand Binding on NMR Observables

NMR spectroscopy studies protein-ligand interactions through changes of NMR parameters upon ligand binding. It can be applied both for ligand screening and also for structural analysis. The binding process for the most straightforward one-site binding is a dynamic equilibrium between the free ligand L, the free protein E and the complex (EL). It can be described by:

$$\mathrm{E + L} \underset{k_\mathrm{off}}{\overset{k_\mathrm{on}}{\rightleftharpoons}} \mathrm{EL}, \tag{1.3}$$
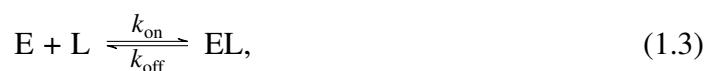
where k_{on} and k_{off} represent the association and dissociation rate constants, respectively. The binding affinity of a ligand to the receptor is characterized by the dissociation constant K_{D}, which is also the ratio of the association and dissociation rates:

$$K_{\text{D}} = \frac{[\text{E}][\text{L}]}{[\text{EL}]} = \frac{k_{\text{off}}}{k_{\text{on}}}. \tag{1.4}$$

Here, [E], [L], and [EL] are concentrations of the free protein, free ligand, and the complex at equilibrium. Typical K_{D} values for drug molecules are in the range of 10^{-10} to 10^{-8} M. At the early stage of drug discovery, the K_{D} values are usually in the range of 10^{-8} to 10^{-5} M for ligands from high-throughput screening. The more weakly binding fragment hits (MW 150–250 Da) from fragment-based screening[22] have K_{D} values varying from 10^{-5} to 10^{-3} M.[23]

At equilibrium, both the ligand and protein molecules undergo exchange between their free state and bound states. The changes in molecular environments lead to the concurrent changes in the relevant NMR parameters, allowing for the detection of binding or for providing information about the binding structure. Firstly, chemical shifts, for the ligand as well as binding sites in the protein, are expected to be affected. The significant changes in the chemical environment for the bound ligand spins and the protein spins that have close contact with the ligand will affect the electron density distribution. The resulting changes in the local magnetic fields lead to the chemical shift differences for the spins located within the binding interface. As a result, examining changes in the chemical shifts for different residues in the protein is a straightforward approach for predicting the ligand binding site.

The exchange rate between the free and bound form of the protein can be written as $k_{\text{ex}}^{(\text{protein})} = k_{\text{on}}[\text{L}] + k_{\text{off}}$, or for the ligand as $k_{\text{ex}}^{(\text{ligand})} = k_{\text{on}}[\text{E}] + k_{\text{off}}$. For NMR observation, three exchange regimes can be defined based on the comparison between the exchange

rate and the difference in the resonance frequencies of the bound and free forms of the ligand or protein, $|\Delta\nu|$. The situation of $k_{\text{ex}} \gg |\Delta\nu|$ is considered as the fast exchange, while $k_{\text{ex}} \ll |\Delta\nu|$ and $k_{\text{ex}} \sim |\Delta\nu|$ define the conditions of slow exchange and intermediate change, respectively.

Significant changes in the spin relaxation also occur for the small molecule ligand upon binding to the receptor. These changes are a manifestation of the differences in the rotational diffusion of the molecule. For calculating spin relaxation, molecular motion is usually characterized by the correlation time τ_c, which is defined as the average time for a particle to rotate by one radian. Under the assumption that molecules are rigid spheres, the rotational correlation time is proportional to the molecular weight according to the Stokes-Einstein equation.[24] Two major mechanisms that account for the nuclear spin relaxation are the dipole-dipole (DD) interaction and chemical shift anisotropy (CSA).[24] In both mechanisms, the magnetic field fluctuations caused by molecular motions can induce NMR transitions, which are responsible for the relaxation processes. The quantitative description of the effect of molecular motions on relaxation rates involves a spectral density function $J(\omega)$:

$$J(\omega) = \frac{\tau_c}{(1 + \omega^2 \tau_c^2)},$$
(1.5)

which describes the frequency distribution of the motions. For example, the longitudinal ^{1}H-^{1}H cross-relaxation rate, which is due to the DD interaction of the nuclear spins, is[25]

$$\sigma^{\text{NOE}} = \frac{1}{10} \frac{\hbar^2 \mu_0^2 \gamma^4}{r^6} \big(- J(0) + 6J(2\omega) \big).$$
(1.6)

Here, μ_0 represents the vacuum permeability, γ is the gyromagnetic ratio of ^{1}H, and r is the distance between the two protons. This cross-relaxation is the origin of the nuclear Overhauser effect (NOE), which appears as a small change of the polarization

of one spin when the polarization of its neighboring spin is perturbed from thermal equilibrium. The involvement of the $J(0)$ term in the expression of the longitudinal cross-relaxation rate leads to the sensitive dependence of the NOE on the correlation time τ_c. In addition, the sign of σ^{NOE} changes from positive for rapidly tumbling small molecules to negative for slowly tumbling macromolecules. These features, along with the $1/r^6$ distance dependence, make the detection of the NOE highly useful for the structural characterization of protein-ligand binding.

1.2.3   Chemical Shift-Based Approaches

When a ligand is added to the protein solution, the chemical shifts of protein spins located in the binding site are affected. The method of investigating the chemical shift changes upon binding is termed as chemical shift perturbation (CSP).[26] The highly crowded ^{1}H spectrum of a protein leads to the necessity for isotope labeling of the protein and multi-dimensional NMR measurements for the observation of CSP. In a standard experiment, a series of 2D HSQC spectra are recorded when the protein is titrated with the ligand to monitor the chemical shift changes. The experiment is applicable to both tightly and weakly binding ligands. For high-affinity ligands in the slow exchange limit, distinct NMR signals from the free and bound protein can be observed, while only changes in the signal intensities of free and bound protein peaks are expected with the ligand titrated in. In the case of fast exchange, by contrast, a single set of signals can be observed at the weighted average of the chemical shifts from the two exchanging states, with a gradual shift from chemical shifts corresponding to the free form to those of the bound form during the titration.

In either the fast exchange or slow exchange regimes, the dissociation constant K_{D} can be determined using the observed chemical shift changes during titration. If assignments are known, the mapped chemical shift changes can also be used for localization of the

binding interface on the protein. The peaks with the most significant change are likely located in the binding site. However, it is crucial to distinguish the chemical shift changes caused by direct interaction from those originating from binding-induced conformational change.[27,28] For the attempts of using CSP to derive more detailed structural information about the ligand binding mode, a semi-quantitative strategy has been developed, which consists of comparing the chemical shift changes for a series of related ligands.[29–31]

1.2.4 NOE-Based Approaches

Intermolecular NOEs between protein and ligand. Distance restraints for ligand-protein binding derived from ^{1}H-^{1}H intermolecular NOEs between protein and ligand can be used directly as the basis for the calculation of the complex structure. Especially when the structure of the target protein is available, a limited number of intermolecular NOEs is sufficient for the determination of the ligand binding mode. Selective measurement of only the intermolecular NOEs can be accomplished by isotope labeling of one component of the complex, followed by pulse sequences equipped with isotope filtering or editing elements.[16,32] For example, a 3D ^{13}C /^{15}N-filtered, ^{13}C-edited NOESY measurement using the sample containing ^{13}C,^{15}N-labeled protein and unlabeled ligand allows the detection of exclusively the intermolecular NOEs between protein and ligand (Figure 1.1a). Typically, a relatively high concentration (> 1 mM) of the ^{13}C or ^{15}N labeled target protein is needed. Protein chemical shift assignments are also necessary for the identification of individual NOE peaks.

A second type of NOE-based NMR experiment involving the detection of only the ligand signals exists, which relies on the fast exchange of ligand between the free and bound forms. Ligand observation substantially reduces the consumption of protein, and does not require isotope labeling or any assignment information of the protein. In the this experiment, information about the bound form can be accessed by observing the signals

from the free and bound ligands, which are collapsed into a single peak. This chemical shift is close to that of the free ligand, since the ligand is usually in large excess over the protein. The averaging effect also exists for other NMR parameters, such as R_1 and R_2 relaxation, and the NOE. The ligand resonances are strongly affected by the properties of the bound ligand, even though it only exists as a small fraction of the total ligand population. The transverse relaxation rate R_2 is commonly used for detection of binding in ligand screening. The parameter does not allow structural information to be extracted. Most ligand-observed experiments intended to provide structural information about the binding interaction rely on observation of magnetization transfer through the NOE.

Intraligand polarization transfer. The transferred NOE (trNOE) method[33] is based on the change in NOE between ligand spins, when the ligand binds to the protein. Usually, intraligand cross-peaks are observed in 2D ^{1}H-^{1}H NOESY spectra. In the presence of the macromolecular target, fast exchange leads to the averaging of the strong negative NOE between neighboring ligand spins in the bound form with the weak positive NOE for the same pair of spins in the free ligand. As a flexible ligand molecule in the bulk solution adopts a single conformation upon binding to the protein, the sensitive distance dependence of NOE results in information about the ligand conformation in the bound state (Figure 1.1b). The trNOE experiment is especially useful for cases involving large, highly flexible ligands such as peptides.[34,35]

Interligand polarization transfer. NOE based polarization transfer can also occur between two different ligands, when binding to the same target. Two experiments based on interligand NOEs are the protein-mediated interligand NOE for pharmacophore mapping (INPHARMA) method[36] and the interligand nuclear Overhauser effect (ILOE) method.[37,38] In INPHARMA experiments, two ligands competitively bind to the same binding site, causing an indirect polarization transfer between the two ligands. This process is mediated by transfer of polarization to and from the target protein. As illustrated
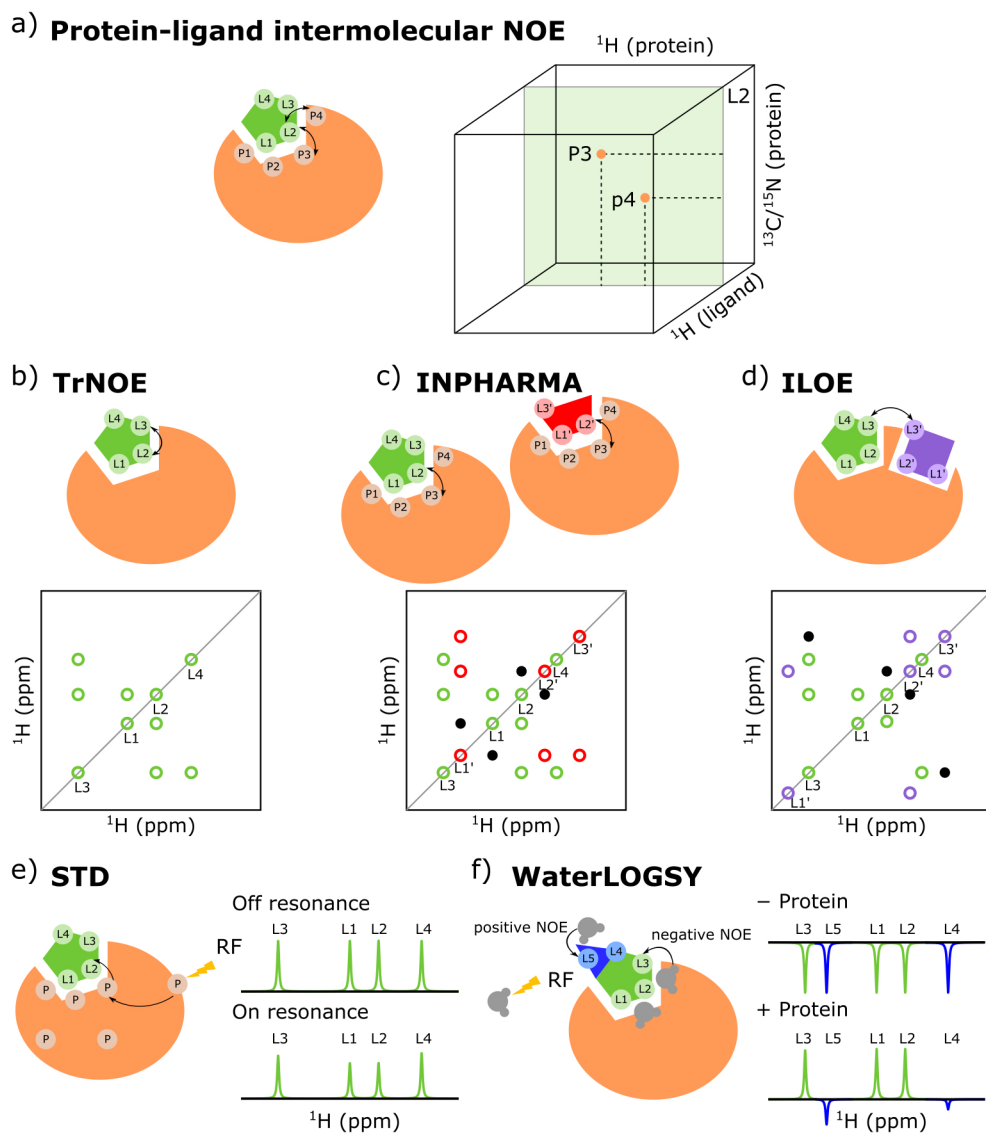
Figure 1.1: Schematic representation of different NOE-based methods for obtaining structural information about the ligand-protein binding. a) Protein-ligand intermolecular NOE. b) TrNOE. c) INPHARMA. d) ILOE. e) STD. f) WaterLOGSY.

in Figure 1.1c, an indirect intermolecular NOE cross-peak is observed between the protons H_{L2} and $H_{L2'}$, which are close to the same proton H_{P3} when the ligands are bound to the target. Through these cross-peaks, information about the relative orientation of the two ligands can be accessed. In ILOE experiments, on the other hand, both ligands bind simultaneously in proximity to each other to the same protein. Hence, direct transfer between the bound forms of the two ligands is considered as the major polarization transfer pathway. Intermolecular cross-peaks are observed for protons that have a direct close contact with each other in the binding mode, for example, the proton pairs H_{L1}-$H_{L1'}$ and H_{L2}-$H_{L2'}$ in Figure 1.1d. In both cases, if the binding mode of other ligand is known, the binding mode of one of the two ligands can be determined from the experiment. In particular, the ILOE experiment has been incorporated with the structure-activity relationships (SAR) approach.[39] The structure-guided design of novel ligands can be performed by linking the two fragment ligands that have been identified by ILOE, to produce a new ligand with high affinity.[40,41] The observation of interligand cross-peaks in the INPHARMA depends on the complex formed with either of the two ligands present at a comparable concentration, while the ILOE experiments rely on the formation of a substantial amount of the ternary complex at equilibrium. As a result, besides the basic requirements for a transferred NOE experiment, both experiments have additional constraints for the relative ligand affinities and sample concentrations.[38,42]

Saturation transfer from protein to ligand. The saturation transfer difference (STD) method[43] is a ligand-based approach for the identification of binding, which has found widespread application in drug discovery. In the STD experiment, a resonance of the target protein is selectively saturated at a frequency far from the resonance frequencies of the ligand (typically at a methyl group chemical shift in the range of -1.5 to 0 ppm). The saturation propagates within the protein through spin diffusion, then transfers to the bound ligand, and is further carried to the free ligand due to fast exchange between the

free and bound forms. The observed ligand signals thus become reduced. A reference experiment, in which the irradiation is applied off-resonance, is recorded subsequently. The difference spectrum would only yield the signals if the ligand can bind to the target protein. Structural information related to the ligand-protein interaction can be obtained by investigating the relative magnitude of the STD signals, which reflect the closeness of individual ligand protons to the protein. This approach is termed as group epitope mapping (GEM).[44] In Figure 1.1e, H_{L1} and H_{L2}, which are in closest contact with the protein, display the strongest STD signal, while H_{L4}, which is not directly in contact with the protein, gives almost no STD signal. A difference in the longitudinal relaxation times (T_1) of individual ligand protons, nonetheless, can significantly affect the magnitude of the observed STD signals, and hence this effect needs to be taken into account when determining the ligand binding mode based on STD data.[45,46]

Saturation transfer from bulk water to ligand. The water-ligand observed via gradient spectroscopy (WaterLOGSY) experiment is a ligand-observed method for probing protein-ligand interactions based on transfer of bulk water magnetization to the ligand.[47,48] A WaterLOGSY experiment begins with the inversion of the water magnetization, followed by its transfer to the bound ligand via different mechanisms. A first transfer pathway is a direct NOE between the bound ligand and water molecules immobilized at the binding interface, which have a residence time that is longer than 1 ns. Secondly, water magnetization can reach the protein through exchange between water protons and labile protein protons, and further propagate to the bound ligand through spin diffusion and intermolecular NOE. In both mechanisms, the slowly tumbling protein-ligand complex is involved, yielding an observed negative NOE correlation between the ligand and water (positive waterLOGSY signal). The small molecules that only interact with bulk water, on the other hand, exhibit a positive NOE. A WaterLOGSY experiment usually does not provide direct information about the proximity of individual ligand protons to the binding

interface. Nonetheless, it is possible to determine the orientation of the bound ligand by distinguishing solvent-accessible ligand protons from buried protons. The solvent accessibility, ligand binding, and mapping of ligand orientation by NMR spectroscopy (SALMON) approach[49,50] observes positive NOEs (negative waterLOGSY signals) for the highly solvent exposed ligand ^{1}H (H_{L4} and H_{L5} in Figure 1.1f), the internal motions of which are not hindered when binding to the macromolecule. In contrast, negative NOEs (positive waterLOGSY signals) are observed for ligand protons (H_{L1}, H_{L2} and H_{L3}) that are buried in the binding pocket.

Methods for characterizing ligand binding to membrane proteins. Ligand-observed approaches are not limited by the size of the protein. This is in contrast to protein-observed methods, which are usually applicable only to protein smaller than 30 kDa.[23,51] Since the correlation time increases proportionally with the molecular weight, ligand detected methods can achieve even larger ligand-to-protein ratios for large molecular weight (MW) targets. Ligand-observed experiments are therefore especially suitable for the study of membrane proteins, for example, G-protein-coupled receptors (GPCR).[52,53] Recently, applications of the ligand-observed NMR approaches have been extended to receptors on living cells. An example is the structural characterization of cyclopeptide ligands binding to integrin $\alpha v \beta 3$ and CD13 on different tumor cells using 2D tr-NOESY.[54] In another application, structural information about the ligand binding mode obtained from ligand-based NMR experiments with living cells has contributed to the design of novel ligands that show improved affinity for Siglec-2 (CD22)[55] and CXCR4.[56]

1.2.5 Paramagnetism-Based approaches

Magnetic properties of nuclei can be strongly affected by the presence of nearby unpaired electrons due to the large electron spin magnetic moment. This feature leads to an alternative approach for structural characterization of protein-ligand binding. The

protein or the ligand molecule can be coupled with a paramagnetic tag, either a radical label or a paramagnetic metal ion. Long-range distances between the paramagnetic center and nuclei can then be measured through the paramagnetic effect. Three types of paramagnetic effects are commonly used, namely paramagnetic relaxation enhancement (PRE), pseudocontact shift (PCS), and residual dipolar coupling (RDC). These effects are measured as differences in relaxation rates, chemical shifts, and coupling constants, respectively, between the samples in the paramagnetic and diamagnetic states.

PRE is manifested as the enhancement of relaxation rates of a nuclear spin caused by the dipolar interaction between the nuclear spin and a nearby unpaired electron spin. PREs exist in any paramagnetic system. They induce line broadening in the NMR spectrum, which has been widely used in ligand screening.[57–59] Both PCSs and RDCs depend on the nonvanishing magnetic susceptibility anisotropy tensor ($\Delta\chi$) associated with the paramagnetic center. To cause these effects, usually a lanthanide ion[60] is tagged to either the ligand[60,61] or the protein.[62,63] The PCS is manifested as a change in chemical shift, which can provide structural information by giving the relative position (both the distance and angle) of the nuclear spins with respect to the paramagnetic center. A large $\Delta\chi$-tensor causes a partial molecular alignment in the external magnetic field, leading to the observation of a change in peak splitting. The RDCs can provide structural information about the orientation of coupled spin pairs. Among the three paramagnetic effects, PCS has become the most popular for structure determination of protein-ligand complexes[60,64] due to the straightforward measurement of the chemical shift changes and the ability to obtain both distance and angular restraints.[65] A distance dependence on r^{-3} of PCS, in contrast to an r^{-6} dependence of PRE, enables the measurement of long-range distances up to 40 Å. In the case when it is possible to attach paramagnetic metal ions at different sites of the target protein, one at a time, the locations of the ligand spins can be unambiguously determined.[66,67]

1.3 Combined NMR and Molecular Docking

1.3.1 Protein-Ligand Docking

Computational methods play an increasingly important role in drug discovery. Molecular docking of small molecules to the protein binding pocket allows for the determination of ligand binding modes.[68,69] It has been widely applied in structure-based drug design since the creation of the first docking program DOCK in the 1980s.[70]

Protein-ligand docking requires the known 3D structure of the protein target with its active site identified. Successful docking relies on the performance of two major components. These include a search algorithm for efficient sampling the configurational and conformational degrees of freedom for the ligand in the active site, and a scoring function to correctly rank the predicted docking models with no ambiguity.[71] In early reported docking methods, usually, all degrees of freedom of the ligand are explored, while the protein is treated as a rigid body. However, the flexible protein molecule can undergo structural changes upon ligand binding, which may range from small side-chain conformational changes to substantial domain rearrangements. Therefore, the treatment of protein flexibility can have a major impact on the accuracy of the docking results. Recent docking programs have incorporated degrees of flexibility for the protein binding site, while maintaining computational efficiency.[72–74] More than 60 docking programs based on different search algorithms and scoring functions have been developed in the last two decades, including DOCK,[70,75] Autodock,[76,77] Autodock vina,[78] GOLD[79], and others. Recent evaluations of the performance of docking programs[80–82] have revealed that efficient ligand conformational sampling can be ensured in most circumstances. However, the ability of scoring functions to consistently differentiate the correct pose still remains a challenge.[83] To overcome this limitation, experimental constraints from NMR spectroscopy, which contain direct structural information on the protein-ligand complex,

can be incorporated with the docking to improve the accuracy of pose assessment. The NMR-derived information that is used to guide the docking process or filter the docked poses can be based on NOEs, chemical shifts or parmagnetic effects.

1.3.2   Use of Chemical Shift Data in Docking

Chemical shift perturbation (CSP) data has been extensively exploited in combination with molecular docking. Mapping the HN chemical shift changes during ligand titration measured in 2D ^{1}H-^{15}N HSQC allows the identification of the ligand binding site. Localizing the binding interface is an essential step required before the molecular docking process. In one approach, the location of ligand binding is identified through chemical shift perturbations, which are assumed to be primarily due to electron ring currents of aromatic ligands.[84] The probability of finding the center of the ring is represented by a dot sphere centered on each perturbed HN. The ligand binding site is then identified as a region with high dot density.

For assisting in a structure calculation of the binding interface, a common strategy involves using CSP defined distance restraints to guide the docking process. In the HADDOCK[85] and LIGDOCK[86] programs, CSPs are converted into distance restraints between atoms of the protein and the ligand. The docking process is driven by the combination a defined penalty function from the violation of the experimental CSPs and the force field-based energy.

Another strategy relies on the comparison of simulated CSPs from pre-docked structures with experimental data. In the AutoDockFilter program, the magnitude of the CSP is assumed to be proportional to the distance between an observed amide proton and the nearest ligand atom.[87] The calculated pseudodistance is utilized for post-filtering of the docking models. In this approach, however, the chemical nature of the ligand is not taken into account. For example, it is not considered that ring currents from an aromatic

ring cause larger chemical shift changes compared to other effects. A more sophisticated simulation of the CSPs based on the ring current effects was developed by McCoy and Wyss.[88] Following this protocol, simulated and experimental CSPs were compared for determining the ligand binding modes in a few protein-ligand systems.[89–91] Rather than considering the complex-induced chemical shift changes for protein protons, Wang *et al.* developed a protocol for determining the binding site structure based on chemical shift changes for ligand protons.[92,93]

1.3.3 Use of NOE Data in Docking

Methods that use NOE-based experimental results to assist molecular docking can be divided into two categories. Methods in a first category define distance restraints from NOE data, while the second category relies on back-calculation of the NOE results based on pre-docked structures.

In the first category, distance restraints between protein and ligand protons are extracted from observed protein-ligand intermolecular NOEs. The NOESY crosspeak intensities are usually converted into upper and lower bounds on the proton-proton distances. These derived distance restraints can be included in computational docking methods, for example serving as ambiguous interaction restraints in HADDOCK and LIGDOCK. Even sparse NOE data can provide sufficient information for the determination of a ligand binding mode. The structures of small molecules binding to heat shock protein 90 (Hsp90),[94] as well as to disulfide-dithiol oxidoreductase (DsbA),[95] were successfully solved by approximately 20 intermolecular NOEs.

In the second category, the NOE data is back-calculated based on the pre-docked structures by simulating the entire polarization transfer process for all involved protein and ligand spins.[96] Moseley *et al.* developed the Complete Relaxation and Conformational Exchange Matrix Analysis (CORCEMA) to study the trNOE data by predicting intensities

of NOESY cross-peaks at different mixing times.[97,98] In CORCEMA, free ligand, free protein, bound ligand and bound protein spins are treated separately. The time evolution of NMR signal intensities for all species are simulated using a system of first-order differential equations. The full dynamic matrix composed of all rate constants in this system of equations is defined by the sum of a relaxation matrix and a kinetic matrix. In the relaxation matrix, the diagonal elements are auto-relaxation rates ρ_i for individual proton spins, which quantify the dependence of the magnetization evolution of the I_i spin on its own deviation from equilibrium. The off-diagonal elements are cross-relaxation rates σ_{ij} accounting for exchange of magnetization between two protons. The kinetic matrix describes the kinetics of the reversible binding process for chemical exchange between the free and bound states of both the protein and ligand.

The CORCEMA theory can be applied to calculate time dependent NMR signal intensities in all types of NOE-based experiments. For different NOE-based experiments, specific modifications should be made for the simulation process, which include the selection of protons involved in the calculation, as well as which spins are initially perturbed from thermal polarization. Jayalakshmi *et al.* extended the application of the CORCEMA theory for quantitative analysis of STD data. In the calculation, protein protons are divided into two classes: protons experiencing direct RF irradiation, and the remaining protons receiving saturation indirectly through spin diffusion.[99] Based on the theory, an STD-NMR Intensity-restrained CORCEMA Optimization (SICO) procedure was developed, which couples the CORCEMA-based calculation with the simulated annealing to refine the ligand bound conformation in the binding pocket.[100] The ligand torsion angles are modified to minimize a scoring function describing the agreement between the experimental STD data and CORCEMA prediction.

When two different ligands are involved, both the INPHARMA and ILOE experiments provide the relative orientation of the two ligands. As described above, the binding mode

of one ligand can be derived if the binding mode of the other ligand is known. Recently, methods have been developed to combine the INPHARMA results with molecular docking and full relaxation and exchange calculation to determine the binding modes of both ligands simultensouly.[42,101–103] In the calculation, poses for each of the two ligands are generated by docking, and a pair of correct ligand structures is selected by comparison of the predicted and experimental INPHARMA NOEs. For the protein kinase A (PKA) and two inhibitors, the accuracy of binding mode prediction was improved compared to using exclusively an energy-based function for pose ranking.[104]

1.4 Dissolution DNP Enhanced NMR in Structural Characterization of Ligand-Protein Interactions

1.4.1 Dissolution DNP

Although NMR spectroscopy can provide detailed structural information about protein-ligand complexes, a major limitation is low sensitivity. The sensitivity of NMR signals is directly related to the nuclear spin polarization P. For example, at the temperature of 300 K and an applied magnetic field of 9.4 T, the equilibrium polarization for ^{1}H spins is 3×10^{-5}. This number, which is much smaller than unity, illustrates a potential for substantial sensitivity improment. According to Equation 1.2, the equilibrium spin polarization can be increased with a stronger magnetic field or lower temperature. Nonetheless, even the highest magnetic field available in a superconducting magnet is not able to provide an enhancement in polarization of more than a few times, and low temperature measurements are often not applicable for biological studies.

However, the spin polarization can be increased transiently with methods termed as hyperpolarization. Several hyperpolarization techniques have been developed, including dynamic nuclear polarization (DNP),[105–107] parahydrogen induced polarization (PHIP),[108–110] spin-exchange optical pumping (SEOP).[111] Among them, DNP is a versatile

method, which is capable of polarizing many nuclei, such as ^{1}H, ^{2}H, ^{19}F, ^{13}C, ^{15}N, ^{29}Si, ^{31}P,[112] in a wide range of molecules.

For liquid state NMR measurements, dissolution DNP (D-DNP) was first described by Ardenkjare-Larsen *et al.* in 2003. This method can achieve over 10,000 times of sensitivity enhancement compared to typical NMR signals acquired at thermal polarization.[113] Hyperpolarization of the nuclear spins is generated in the solid state at a temperature of about 1 K. This process relies on polarization transfer from the unpaired electron spins of free radicals to the nuclear spins while microwave irradiation is saturating an electron spin transition. For efficient polarization transfer, the molecule of interest and the radicals are prepared in a glass-forming solvent to assure homogeneous mixing. The frozen sample is then dissolved by a pre-heated solvent, and rapidly injected into the NMR spectrometer for liquid state NMR experiments. D-DNP has initially been designed for use with magnetic resonance imaging (MRI) for *in vivo* study of biological processes.[114,115] A variety of D-DNP NMR experiments have also been developed for the investigation of chemical reactions or processes by monitoring the time evolution of NMR signals,[116–119] as well as the study of spin dynamics by measuring spin relaxation[120,121] or polarization transfer.[122–125]

A limitation of the D-DNP technique is the non-renewability of the hyperpolarization, which prevents the application of conventional multi-dimensional experiments used for obtaining structural information. Progress has been made to accomplish the fast acquisition of correlation spectra. An ultrafast 2D NMR method with single-scan acquisition utilizing pulsed-field gradients for spatial encoding was demonstrated by Frydman and coworkers.[126,127] In another approach, 2D NMR spectra are acquired sequentially with variable flip-angle excitations to maintain the same consumption of hyperpolarization in each scan.[128] Schanda *et al.* have introduced the 2D band-selective optimized-flip-angle short-transient (SOFAST) experiment, which allows two-

dimensional NMR spectroscopy with short interscan delays.[129,130] This method has recently been applied for acquiring 2D NMR spectra of proteins receiving polarization transfer from hyperpolarized water.[131–133] Other approaches for fast multi-dimensional signal acquisition exploit sparse sampling, which allows obtaining correlation spectra from a small number of points in the indirect dimension. These methods include nonuniform sampling (NUS),[134] as well as Hadamard spectroscopy.[135] In the latter, only several chemical shifts of interest are encoded with selective radiofrequency pulses according to a Hadamard matrix, and subsequently reconstructed into indirect spectral information. The small number of required data acquisitions makes it well suited for use with hyperpolarization.[136]

1.4.2 Investigating Protein-Ligand Interactions Using Transfer of Hyperpolarization

NOE is proportional to the deviation from equilibrium polarization of the source spin. In a traditional NOE experiment, the maximum deviation can be twice the equilibrium value from inversion of the source spin. D-DNP can prepare a hyperpolarized source spin that displays a deviation from equilibrium that is orders of magnitude larger, leading to an enhancement of the NOE. For the study of protein ligand interactions, the small molecule ligand may be hyperpolarized and injected into a protein solution for NMR data acquisition. In principle, almost all types of conventional NOE-based NMR experiments for structure elucidation of protein-ligand complexes can be modified to incorporate the D-DNP technique. These experiments provide structure-related information through monitoring of the hyperpolarization transfer by intermolecular or intramolecular NOE.

Lee and coauthors have developed a method using hyperpolarization with protein-mediated interligand NOEs,[137] which is similar to the INPHARMA experiment. This method was demonstrated with protein kinase A (PKA) and two competing ligands. [1]H

spins of ligand 1 were hyperpolarized using D-DNP. Polarization transfer to ligand 2 mediated by protein ^{1}H spins in the binding pocket was monitored by measuring time-dependent signal changes for both ligands by multi-scan NMR with small flip angle excitation. The difference in the build-up rates for individual ligand 2 spins reveals their relative proximity to the protein binding interface, which provides structural information for the binding mode of ligand 2. This experiment does not provide direct ^{1}H-^{1}H correlation information between protons of the two ligands, since the origins of the hyperpolarization from ligand 1 can not be distinguished. Hence, knowledge about relative orientation of the two ligands is not directly accessible. Still, the possibility of predicting binding modes of both ligands using computational methods were demonstrated.

An implementation of combining the WaterLOGSY experiment with the D-DNP method was reported by Chappuis *et. al.*,[124] In a demonstration with the protein DOt1L and its ligand, water molecules were hyperpolarized on their ^{1}H spins and injected into a solution containing the protein and ligand. Enhancement was observed for both the protein and ligand signals. Hyperpolarization of water rather than the ligand makes this approach applicable without the need to polarize different kinds of ligand molecules. Recently, the method of investigating polarization transfer from water to ligand has been integrated with the multi-dimensional NMR to characterize the interaction between a peptide and the liposomes.[138] Hyperpolarization transfer directly from water protons to the peptide amide protons leads to the enhanced signals observed in a fast 2D COSY spectrum of the peptide. Differences in the signal losses in the presence of liposomes allowed the identification of solvent shielded sites of the peptide, providing information about the orientation of the peptide interacting with the liposome. This approach is similar to the SALMON (solvent accessibility, ligand binding, and mapping of ligand orientation by NMR spectroscopy) method[49,50] described previously to derive ligand binding epitope by identification of solvent-accessible protons. Theoretically, the method of using hyperpolarization transfer

from water to map ligand orientation can also be applied to a protein-ligand system. However, direct structural information about the interaction is still challenging to extract due to the complexity introduced by coexisting polarization transfer pathways, including different NOEs and chemical exchange.

Another ligand-observed NOE based method, the trNOE experiment, is also compatible with D-DNP. Unlike INPHARMA and WaterLOGSY that investigate polarization transfer between different molecules, the trNOE experiment relies on observation of intraligand NOEs. The non-selective hyperpolarization of the ligand spins by DNP requires additional selection of specific ligand peaks, for example by saturation. In Chapter 2, we develop a method using hyperpolarized trNOE to study interactions between immobilized avidin and its ligand 4′-hydroxyazobenzene-2-carboxylic acid (HABA).

Besides the ligand-observed NMR experiments, the large polarization provided by the dissolution DNP can also be beneficial for NMR experiments based on the observation of protein signals. The mixing of the hyperpolarized ligand with the protein solution leads to the polarization transfer from the ligand to the protein through intermolecular NOE. Min *et al.* observed the buildup of enhanced protein signals after the hyperpolarized ligand benzamidine was injected into to the solution of its target protein trypsin.[123] The "fingerprint" of the enhanced protein signals differed from the pattern of the protein signals record at thermal equilibrium, while it matched the frequency profile of an STD experiment. The selective enhancement of the protein signals originates exclusively from the hyperpolarized ligand ^{1}H spins that are in close contact with the protein. However, the crowded protein ^{1}H signals obtained in 1D NMR does not allow identification of individual protein resonances. In Chapters 3-5, we demonstrate the direct observation of polarization transfer from the hyperpolarized ligand folic acid to the protein dihydrofolate reductase (DHFR). Using 1D ^{1}H NMR with a ^{13}C filter or rapid multi-dimensional NMR, resolved NOE signals can be obtained, which carry essential correlation information. We

then develop a method for combining the experimental NOE restraints with computational

docking for the determination of the ligand binding mode.

2. AMPLIFICATION OF NUCLEAR OVERHAUSER EFFECT BY HYPERPOLARIZATION FOR SCREENING OF LIGAND BINDING TO IMMOBILIZED TARGET PROTEINS

2.1 Introduction

Nuclear magnetic resonance (NMR) spectroscopy plays an essential role in the drug discovery process for the characterization of protein-ligand interactions. It enables both the fast screening for ligand binding, as well as the determination of full binding epitope structures at atomic resolution.[139,140] A limitation for the application of NMR spectroscopy is the comparably large amount of protein and ligand that is needed. Observing NMR signals of the small-molecule ligand instead of the protein target reduces the demand for the target, with the additional advantage that there is no limitation on its molecular weight. Ligand-observed measurements rely on fast exchange between the free and bound forms of the ligand, which causes averaging of observable parameters that change upon binding. Such parameters include chemical shift, spin relaxation rates, as well as cross-relaxation rates that may be manifested as cross-peaks in nuclear Overhauser effect (NOE) spectroscopy.[51,141,142]

The amount of protein needed in ligand observed NMR experiments can be further reduced by immobilizing the target proteins on a solid support, so that they can be easily reused. The target immobilized NMR screening (TINS) method capitalizes on this option by immobilizing the protein on Sepharose beads. Binding is detected by line broadening of the ligand signal caused by a change in T_2 relaxation. With this method, a library of 2000 compounds can be screened using only 3-5 mg of the target.[143] This method has been applied to thermostabilized (StaR) G protein-coupled receptors (GPCRs), which are integral membrane proteins,[144,145] in addition to water soluble proteins.

24

The T_2 relaxation parameter provides for a robust identification of binding, but does not contain detailed information on molecular structure. The bound ligand structure can be accessed through transferred NOEs (trNOEs), which are also measured solely from the ligand signals.[33,146] When the requirement of fast exchange is met, NOE signals are averaged over the free and bound forms of the ligand. Due to the significantly longer rotational correlation time of the bound ligand, strong negative NOEs are visible in these signals even at low protein concentration. TrNOE spectroscopy has been applied to purified targets including membrane proteins,[147] as well as to targets on living cells.[54]

NOEs cause only a fractional change in the observed signal from a spin upon perturbation of the equilibrium Zeeman population distribution of a nearby spin, exacerbating the sensitivity limitations of NMR. The signal in the observation of the NOE can be increased by hyperpolarization of nuclear spins, such as with the dynamic nuclear polarization (DNP) technique.[113] Specifically, dissolution DNP allows the generation of a large non-Boltzmann polarization for typical ligand molecules in solution. The resulting deviation from the equilibrium population distribution is several orders of magnitude larger than can be achieved with saturation or inversion pulses in a traditional NOESY experiment. With the sensitivity enhancement gained in dissolution DNP experiments, the observation of NOE signals becomes possible in a single scan under nearphysiological conditions. In our previous work, we have obtained structural information on binding interfaces by observation NOE based polarization transfer from hyperpolarized ligands to the protein, as well as among competitively binding ligands.[123,137,148]

Here, we introduce a method for the measurement of trNOE for a ligand binding to an immobilized target. We demonstrate that this method can detect the protein-ligand interaction, and also provide pertinent structural information on the ligand binding mode through the determination of cross-relaxation rates between ligand ^{1}H spins.

2.2 Experimental Section

Avidin coated polystyrene beads (0.5% w/v in PBS buffer, 6.0-8.0 μm diameter) were purchased from Spherotech (Lake Forest, IL). A concentrated stock solution of beads was prepared by centrifugation at $3000 \times g$ for 15 minutes. The binding capacity of the concentrated solution was estimated using a sepctrophotometric assay.[149]. It was assumed that the bead solution has the same binding capacity for HABA as for the strong ligand biotin. The binding capacity for biotin was determined based on the measurement of the amount of biotin left when removing the biotin-bound bead after adding a certain amount of biotin. 10 μL of 40 μM biotin solution was added to 30 μL of the bead solution, followed by incubation for 5 min and centrifugation for 5 min. The concentration of biotin left in the supernatant was quantified using the HABA/avidin colorimetric assay.[150] The supernatant was added to the HABA-avidin solution, causing the displacement of HABA from the complex. The decrease of absorbance of the HABA-avidin complex at 500 nm (A_{500}) was measured (Figure 2.1). A control experiment was performed by replacing the supernatant with 10 μM biotin solution. The binding capacity of the bead stock solution was determined as 8.9 μM. The manufacturer reported binding capacity is 0.05 nmole biotin per milligram of 0.5 % w/v bead suspension in PBS buffer, which corresponds to a bead concentration of 0.25 μM. The optical density at 600 nm OD_{600} values for the original solution and the concentrated bead solution were measured as 4.13 and 162, respectively. Therefore, the binding capacity of the concentrated bead solution according to the manufacturer is 9.8 μM, which is within 10 % of the value determined above. The optical density at 600 nm of this stock solution was measured as 162. The dilution factor for the bead suspension after each dissolution experiment was quantified by measuring OD_{600}.

DNP samples were prepared with 6.7 mM 4′-hydroxyazobenzene-2-carboxylic acid
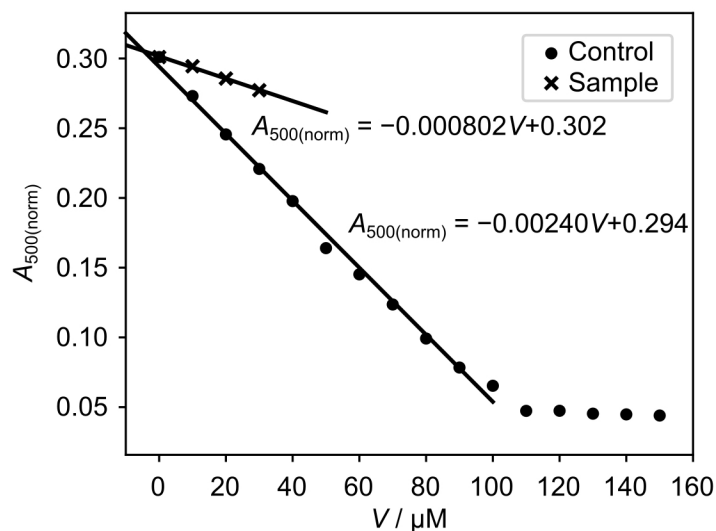
Figure 2.1: Decrease in absorbance of the HABA-avidin complex at 500 nm (A_{500}) when titrating 150 μL (V_i) HABA-avidin solution with the supernatant from the bead solution and the control biotin solution (10 μM), respectively. V_a represents the volume of the sample/control solution added. The absorbances were normalized by the total sample volumes, as $A_{500(norm)} = A_{500} \times (V_a + V_i)/V_a$. The biotin concentration in the sample supernatant was determined as 3.3 μM.

(HABA; Thermo Fisher Scientific, Fremont, CA) and 15 mM 4-hydroxy-2,2,6,6-tetramethylpiperidin-1-oxl (TEMPOL) in D_2O/DMSO−d_6 (1:1 v/v; Cambridge Isotope Laboratories, Tewksbury, MA). 10 μL aliquots were hyperpolarized in a HyperSense DNP polarizer (Oxford Instruments, Abingdon, U.K.) at a temperature of 1.4 K. Microwave irradiation was applied for 20 min at a frequency of 94.005 GHz and a power of 100 mW for ^{1}H polarization. The hyperpolarized sample was then dissolved in 4 mL phosphate buffer (50 mM, pH = 7.4), which had been heated to reach a vapor pressure of 0.6 MPa. The dissolved samples were transferred into the loop of a fast injector.[151] The samples were then injected into a 5 mm NMR tube located in a 400 MHz NMR spectrometer equipped with a broadband observe probe (Bruker Biospin, Billerica, MA). Injection occurred with a forward pressure of 1.81 MPa N_2 gas against a back pressure of 1.03 MPa. In the NMR
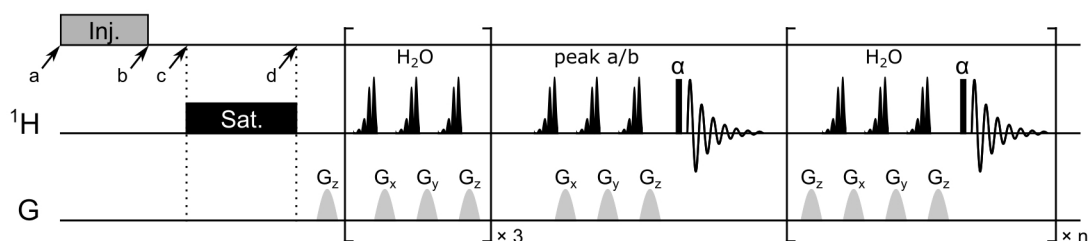
Figure 2.2: Pulse sequence for DNP NMR experiments. The hyperpolarized sample was transferred from the injection loop into the NMR instrument during the injection time of 420 ms. The NMR experiment was then triggered after a stabilization time of 100 ms, starting with a 500 ms saturation pulse applied at the resonance frequency of one ligand peak. 32 spectra were measured with small flip angle pulses, and the time interval between excitation pulses was 0.35 s. Solvent suppression was performed in each scan by selectively exciting the water resonance with EBURP shaped $\pi/2$ pulses, and dephasing it by randomized pulsed field gradients G_x, G_y and G_z. Right before the first acquisition, the selective ligand resonance was suppressed similarly by combination of EBURP shaped $\pi/2$ pulses and pulsed field gradients.

tube, the sample mixed with 40 μL phosphate buffer, concentrated solution of polystyrene beads coated with avidin, or bead solution with additional 20 μM biotin, which had been pre-loaded.

After sample injection and stabilization, multi-scan NMR experiments were conducted with small flip angle excitations following suppression of one ligand resonance (Figure 2.2). The final concentrations of HABA for each experiment were calculated by comparing ^{1}H signal integrals measured after completion of the hyperpolarized experiments with those from a reference HABA sample with known concentration. The concentrations of the avidin binding site were determined as previously described using OD_{600}.

2D NOESY spectra were acquired on a 500 MHz NMR spectrometer with a TCI cryoprope (Bruker Biospin) with varying mixing time (τ_m = 0.3, 0.5, 0.7 s). Samples containing 100 μM HABA, 100 μM HABA with 5 μM avidin and 100 μM HABA with avidin coated polystyrene beads were measured respectively. 128×4096 complex points

28

were collected in the indirect and direct dimensions with $t_{1,\mathrm{max}} = 9.8$ ms and $t_{2,\mathrm{max}} = 315$ ms, respectively. During the measurements of polystyrene beads, the sample was taken out of the spectrometer for remixing after every 32 scans. Cross-relaxation rates were determined by a linear fit of the volume ratios of the cross-peaks to the diagonal peaks ($A_{\mathrm{cross}}/A_{\mathrm{diagonal}}$) vs. τ_{m} curve. The concentration of the avidin was determined based on the molar extinction coefficient of 25,000 cm^{-1}M^{-1} per binding site at 282 nm.[150]

The data from the DNP NMR experiments were processed with an exponential window function (4 Hz line broadening) and Fourier transformed using Topspin software (Bruker Biospin). The baseline was corrected with a fifth-order polynomial. Integrals of peaks from H$_{\mathrm{a}}$ and H$_{\mathrm{b}}$ were calculated. The Time evolution of the two signals can be described by the Solomon equations for a system consisting of two spins:[152]

$$\frac{\mathrm{d}(I_{\mathrm{a}} - I_{\mathrm{a(eq)}})}{\mathrm{d}\,t} = -r_{\mathrm{a}}(I_{\mathrm{a}} - I_{\mathrm{a(eq)}}) - \sigma(I_{\mathrm{b}} - I_{\mathrm{b(eq)}}) \tag{2.1}$$

$$\frac{\mathrm{d}(I_{\mathrm{b}} - I_{\mathrm{b(eq)}})}{\mathrm{d}\,t} = -r_{\mathrm{b}}(I_{\mathrm{b}} - I_{\mathrm{b(eq)}}) - \sigma(I_{\mathrm{a}} - I_{\mathrm{a(eq)}}) \tag{2.2}$$

I_{a}, I_{b} represent signal integrals for spin H$_{\mathrm{a}}$ and H$_{\mathrm{b}}$ and $I_{\mathrm{a(eq)}}$, $I_{\mathrm{b(eq)}}$ are the integrals at thermal equilibrium. Time dependent signal integrals of Ha and Hb were fitted by repeated numerical solution of Equation 2.1 and 2.2 for periods between two successive pulses. An additional signal depletion of a factor of $\cos(\alpha)$ was included for each small flip angle excitation.[125] The fitted parameters include the auto-relaxation rates r_{a} and r_{b} for spins H$_{\mathrm{a}}$ and H$_{\mathrm{b}}$, and the cross-relaxation rate σ between the spins.

For the trials with avidin coated beads, σ_{b}, was calculated with the fitted σ, already determined σ_{f}, and the bound fraction of ligand, X_{b}, using $\sigma_{\mathrm{b}} = (\sigma - X_{\mathrm{f}}\sigma_{\mathrm{f}})/X_{\mathrm{b}}$. Here, X_{b} was determined with known concentrations of HABA, $[L]_{\mathrm{t}}$, and avidin binding sites, $[P]_{\mathrm{t}}$ the dissociation constant K_{d}, using $X_{\mathrm{b}} = ([P]_{\mathrm{t}} + [L]_{\mathrm{t}} + K_{\mathrm{d}} - \sqrt{([P]_{\mathrm{t}} + [L]_{\mathrm{t}} + K_{\mathrm{d}})^2 - 4[P]_{\mathrm{t}}[L]_{\mathrm{t}}})/2[L]_{\mathrm{t}}$.

2.3 Results and Discussion

The ^{1}H spins of 4'-hydroxyazobenzene-2-carboxylic acid (HABA), were hyperpolarized using dissolution DNP. The binding to the protein was characterized by intraligand trNOE measurements utilizing the nuclear spin hyperpolarization. Two multi-scan NMR data sets, each acquired from a single hyperpolarization experiment, are shown in Figure 2.3. The spectra in Figure 2.3a and b are from a solution containing polystyrene beads coated with avidin. For comparison, spectra from a second experiment without the bead particles are shown in Figure 2.3c and d.

Detection of the intraligand NOE was accomplished by observing polarization transfer to a pre-suppressed ligand spin, which was Ha in both experiments shown in Figure 2.3. The polarization originated from a neighboring hyperpolarized spin in the target molecule. The D-DNP technique hyperpolarizes all spins of the same type of nucleus in a molecule, therefore the suppression of the target spin signal from its highly enhanced state was necessary. This suppression was achieved before the acquisition of the first transient, by irradiating the ligand signal for H_a with both a continuous saturation pulse and a series of selective shaped pulses. After suppression, a transient build-up of NOE signal was observed. The suppression of one ligand signal is confirmed in the first scan (Figure 2.3b and d, top). Since the NOE with a $1/r^6$ dependence is highly sensitive to the distance between the participating spins, in the following only polarization transfer between spin pair H_a and H_b is considered (see ligand structure in Figure 2.3e).

When only the ligand is included in the experiment and no protein containing beads are present, the signal of H_a at first rapidly reduces to negative signals, reaches a negative maximum (Figure 2.3d, bottom) and slowly approaches to the thermal equilibrium in the end. Meanwhile, the signal from H_b decays exponentially from its originally hyperpolarized state. In contrast, in the presence of protein coated beads, the signal for H_a
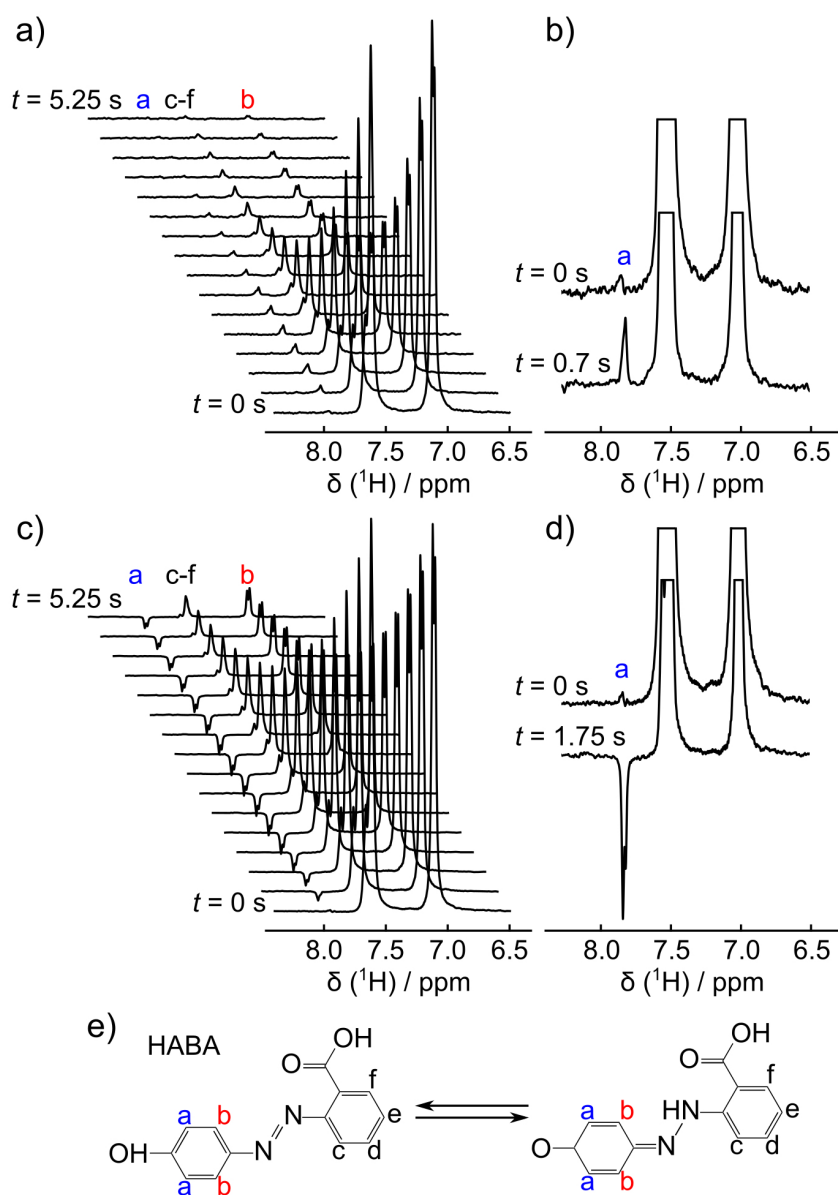
Figure 2.3: a) Time-resolved series of 1H NMR spectra of hyperpolarized HABA (63.5 μM) with avidin coated polystyrene beads (0.83 μM binding site) after suppression of H_a. The spectra were acquired with a 22.5° flip angle excitation pulse at intervals of 0.35 s. The spectra for the first 5.25 s from a total of 10.85 s acquired are shown. b) 1D spectra from (a), including the first scan taken right after complete signal suppression at $t = 0$ s, and the spectrum showing maximum NOE taken at $t = 0.7$ s. c) The same panels as in (a), from only hyperpolarized HABA (65.7 μM) without beads. d) 1D spectra from (c), including the first scan at $t = 0$ s and the scan showing the largest negative NOE signal taken at $t = 1.75$ s. e) Tautomers of HABA. The free ligand exists in the azo form shown to the left, while the bound ligand is the hydrazone tautomer shown to the right.

31

increases until a positive maximum is reached (Figure 2.3b, bottom), and then gradually returns to its thermal equilibrium value. This qualitative difference is attributed to the change in sign of the cross-relaxation rate (σ), which gives rise to the NOE. The small molecule ligand alone exhibits a short correlation time for its rotational motion, resulting in a positive cross-relaxation rate and negative NOE build-up. When the ligand binds to the larger receptor, it takes on the motional properties of the macromolecule within the residence time of its bound state, giving a large correlation time and hence a positive NOE build-up.

An average signal enhancement factor of 3300 for free ligand ^{1}H spins in the 400 MHz NMR spectrometer was observed (Table 2.1). When the protein coated beads were added, the increased T_1 relaxation led to a lower enhancement factor, which was approximately 2000. Still, the significant signal enhancement allowed single-scan detection of the NOE signal, which appeared as only 2% of the signal from the unsuppressed hyperpolarized NOE source spin, for ligand concentration of 63.5 µM and binding site concentration of 0.83 µM. The maximum NOE signal observed (Figure 2b) during the acquisition gave a signal-to-noise ratio of 17. Maintaining a similar bound ratio for the ligand, the detection limit would be around 10 µM ligand with a binding site concentration of 0.2 µM on beads. Here, the DNP NMR measurements were performed on a 400 MHz spectrometer with a room-temperature probe. A cryoprobe would provide an additional 3-4 times of enhancement that is independent of the polarization.[153]

Quantitative values for the auto-relaxation rates of spin H_a and H_b, as well as cross-relaxation rates between the two spins were obtained by fitting the time evolution of the NMR signals using the Solomon equations (Equation 2.1 and 2.2 in Experimental Section).[152] The fitting results for hyperpolarized experiments with only HABA and those with HABA and pre-loaded avidin coated beads are given in Figure 2.4a and b, respectively. For each sample condition, two sets of data are shown for experiments with

32
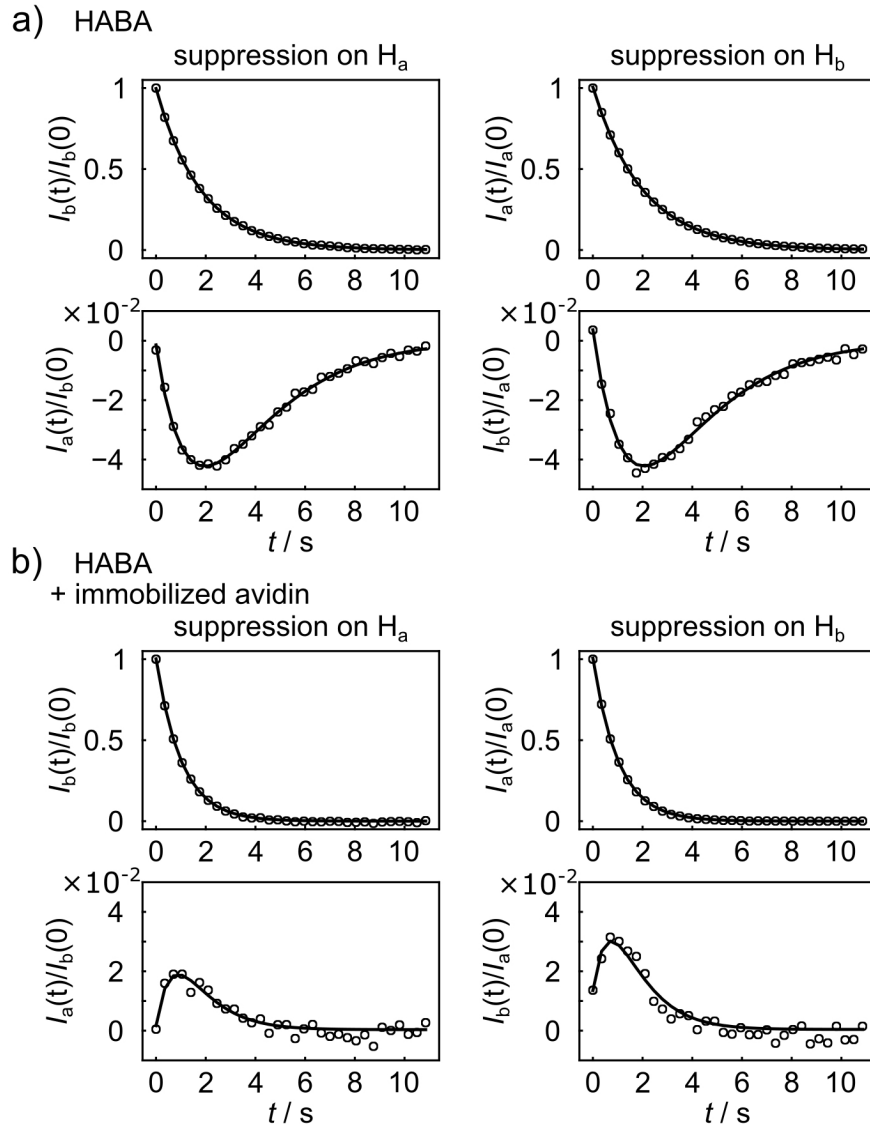
Figure 2.4: Time dependence of ^{1}H signal integrals of H_a and H_b demon-strate polarization transfer to a pre-suppressed target spin (H_a or H_b) from the other unsuppressed source spin. The integral values, $I_a(t)$ and $I_b(t)$ were normalized by the integral from the unsuppressed peak in the first scan as $I_a(0)$ for H_b suppression or $I_b(0)$ for H_a suppression. a) Hyperpolarized HABA. The left panels show the data set with signal suppression on H_a. The fitted cross-relaxation rate is $\sigma = 0.057$ s^{-1}. The right panels give the results for the data with signal suppression on Hb, $\sigma = 0.059$ s^{-1}. b) Hyperpolarized HABA with avidin coated beads. The fitted cross-relaxation rates are $\sigma = -0.055$ s^{-1} with H_a pre-suppressed (left) and $\sigma = -0.074$ s^{-1} with H_b pre-suppressed (right).

Exp no.	c_{HABA} (μM)	$c_{\text{binding site}}$ (μM)	X_b	Peak suppressed	ϵ (a/b)	r_a (s^{-1})	r_b (s^{-1})	σ (s^{-1})	σ_b (s^{-1})
1	66.9	-	-	a	3.26E+03	0.28	0.34	0.066	-
2	65.7	-	-	a	3.67E+03	0.23	0.33	0.057	-
3	66.6	-	-	a	2.84E+03	0.22	0.34	0.056	-
4	69.7	-	-	b	3.12E+03	0.27	0.27	0.058	-
5	69.7	-	-	b	3.06E+03	0.28	0.28	0.056	-
6	67.6	-	-	b	4.01E+03	0.27	0.28	0.059	-
7	63.5	0.83	0.12	a	2.11E+03	1.00	0.75	−0.055	−9.48
8	90.4	0.83	0.009	a	2.59E+03	0.87	0.67	−0.032	−8.48
9	64.7	1.15	0.016	b	1.74E+03	0.76	1.00	−0.074	−10.45
10	68.4	0.90	0.012	b	2.07E+03	0.67	0.84	−0.045	−8.13
11	73.9	0.54	0.007	a	2.74E+03	0.24	0.45	0.025	-

Table 2.1: Experimental parameters and fitted results of the DNP trNOE experiments with hyperpolarized HABA. Exp no. 1-6: HABA only; Exp no. 7-10: HABA mixed with preloaded avidin coated polystyrene particles; Exp no. 11: HABA with preloaded beads and biotin. The enhancement factor ϵ for the unsuppressed signal (a or b) was determined by comparing the peak integral from the first scan of the DNP experiment with the reference spectrum measured under thermal polarization for the unsuppressed peak. r_a, r_b, and σ were obtained from fitting the time evolution σ_f peak integrals simultaneously for signal a and b using the Solomon equations (Equation 2.1 and 2.2). σ_b was determined when immobilized protein is involved, based on the overall cross-relaxation rate σ with determined cross-relaxation rate for free ligand σ_f and the bound fraction X_b (Formula in Experimental Section).

suppression pulses applied on spin H_a or H_b. Despite the choice of the pre-suppressed peak, similar trends are observed in the decay curves of the source spin signal, as well as in the build-up curves of the NOE signal. The fitted parameters along with the experimental conditions for a total of six NMR measurements with only hyperpolarized HABA and a total of four experiments with hyperpolarized HABA and preloaded avidin coated beads are summarized in Table 2.1. Among the 6 trials with no bead present, consistent cross-relaxation rates, a cross-relaxation rate $\sigma_f = 0.059 \pm 0.004$ s^{-1} for free HABA was obtained, where the range indicates the standard deviation of the results from the individual measurements. Since this cross-relaxation occurs between two spins in the same molecule, it is, as expected, independent of the concentration. For the four trials with the avidin
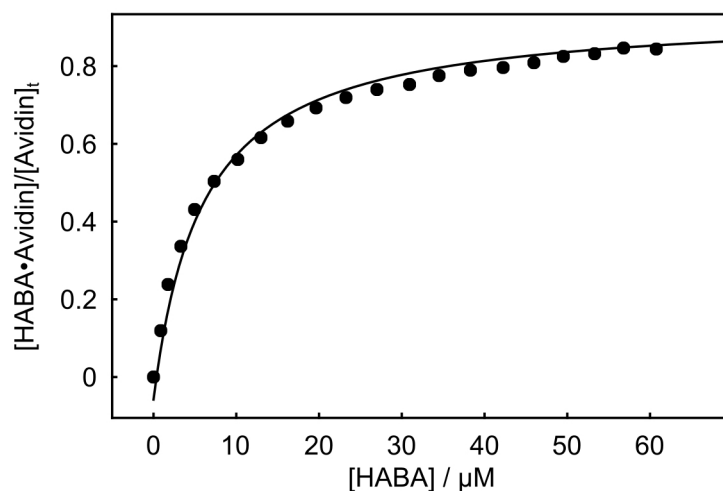
Figure 2.5: Titration of avidin with HABA. The affinity was determined by measuring the absorbance at 500 nm using UV/Vis spectrophotometry. A molar extinction coefficient of 35500 $cm^{-1}M^{-1}$ for the HABA-avidin complex[155] was used for calculating the complex concentration [HABA·Avidin]. The fraction of bound protein was fitted with the equation $[\text{HABA} \cdot \text{Avidin}]/[\text{Avidin}]_t = [\text{HABA}]/([\text{HABA}] + K_d)$. The dissociation constant K_d was determined as 5.9 ± 0.6 μM.

coated beads, the directly derived σ values vary for the different samples, ranging from $\sigma = -0.032\,\text{s}^{-1}$ to $-0.074\,\text{s}^{-1}$. Here, σ is the weighted average of the cross-relaxation rates in the free (σ_f) and bound forms (σ_b). Hence, it depends on the bound fraction of the ligand X_b, which was calculated for each sample with known ligand and protein concentrations and the dissociation constant K_d (see Experimental Section). K_d was determined by UV/Vis photometry as 5.9 ± 0.6 μM (Figure 2.5), in agreement with the previously reported value of 6 μM.[154] Using the above determined σ_f, combined with X_b, the cross-relaxation rate for the bound form was found to be $\sigma_{b(p-b)} = -9.1 \pm 1.1\,\text{s}^{-1}$.

This result is close to the value of $-7.47\,\text{s}^{-1}$ determined by a series of non-hyperpolarized 2D ^{1}H-^{1}H NOESY measurements at 500 MHz (Figure 2.7 and 2.8). The degree of deviation in the cross-relaxation rates caused by different field strengths can be estimated using using the following equations, which take into account only the dipolar
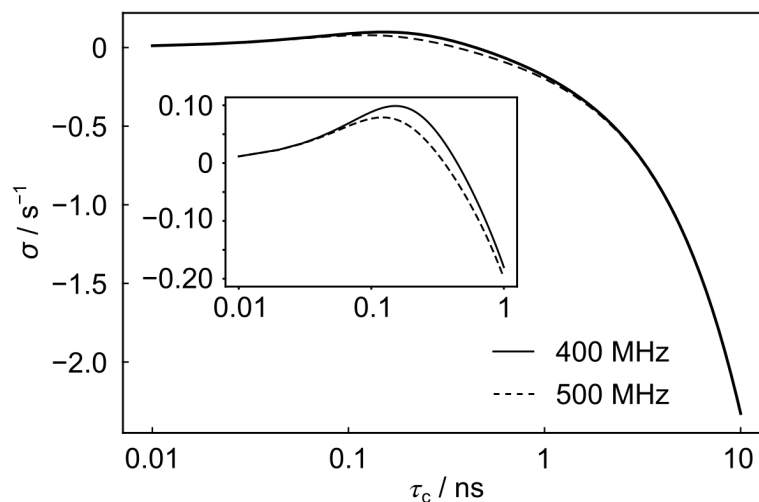
Figure 2.6: Calculated cross-relaxation rate for a ^{1}H-^{1}H dipolar spin system at 400 MHz and 500 MHz. Equation 2.3 and 2.4 were used to perform the calculations with r = 2.5 Å.

relaxation in a two-spin system.[24]

$$\sigma = \left(\frac{\mu_0}{4\pi}\right)^2 \frac{\gamma_H^4 \hbar^2}{10} \frac{1}{r^6}[(-J(0) + 6J(2\omega)] \tag{2.3}$$

$$J(\omega) = \frac{\tau_c}{1 + (\omega\tau_c)^2} \tag{2.4}$$

γ_H is the gyromagnetic ratio of the proton and r is the spin-spin distance. σ values two different field strengths corresponding to ^{1}H frequencies (ω) of $2\pi \times 400$ Hz and $2\pi \times 500$ Hz were simulated with the molecular correlation time (τ_c) changing from 0.01 ns to 10 ns (Figure 2.6). The difference in cross-relaxation rates between 400 MHz and 500 MHz is estimated as less than 10% in the small molecule and less than 1% in the macromolecule. The sample used for the conventional measurement contained 100 µM ligand with the bound ratio of 0.021, compared with the averaged 70 µM ligand with the ligand bound ratio of 0.012 in the hyperpolarized experiments. Still, rather than a single experiment with a polarization time of approximately 20 min followed by an NMR acquisition

System	σ / s^{-1} (DNP, 400 MHz)	σ / s^{-1} (thermal, 500 MHz)
free HABA	0.059 ± 0.004	0.066
bound HABA on avidin	-	-0.83
bound HABA on immobilized avidin	-9.15 ± 1.05	-7.52

Table 2.2: Cross-relaxation rates σ between spin H$_a$ and H$_a$ of HABA in deferent forms determined from the DNP trNOE experiments at 400 MHz and non-hyperpolarized 2D NOESY experiments measured at 500 MHz. The error ranges indicate standard deviations from repetitions summarized in Table 2.1.

time of 10 seconds, the non-hyperpolarized measurement required a titration including four experiments of 2 hours each, using a cryoprobe for sensitivity enhancement. The conventional experiment required removing the sample from the magnet every 30 minutes to resuspend beads. Partial precipitation, however, may still lead to an underestimation of the NOE the conventional measurement.

Since the trNOE experiment is conducted with the ligand in large excess, there is a possibility for a contribution by nonspecific binding to the protein or bead surface.[157] In order to quantify this contribution to the change in the cross-relaxation rate, the trNOE experiment was repeated with a bead suspension that was pre-treated with biotin. Biotin is a high-affinity ligand for avidin ($K_d = 10^{-15}$ M^{-1}), which blocks its binding site.[154] The preloaded solution was prepared as 20 μM biotin with 8.9 μM binding site, to ensure full saturation of the specific binding site. With the biotin treated beads, the cross-relaxation rate between spins of H$_a$ and H$_b$ of hyperpolarized HABA was determined to be $\sigma_c = 0.017$ s^{-1} (Figure 2.9). This value is smaller than the cross-relaxation rate $\sigma_f = 0.059$ s^{-1} for free HABA, indicating the existence of non-specific binding. It is however still a positive number, showing that the specific binding makes a major contribution to the observed negative trNOE in the HABA-avidin on the bead system. This result was compared with

37
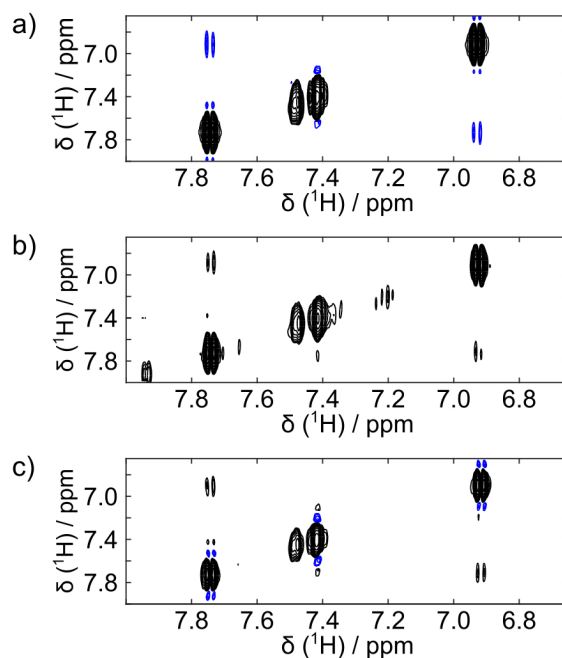
Figure 2.7: ^{1}H-^{1}H NOESY spectrum of a) 100 μM HABA showing positive NOE b) 100 μM HABA with avidin (20 μM binding site) showing negative NOE c) 100 μM HABA with avidin immobilized on polystyrene beads (2.25 μM binding site) showing negative NOE. Mixing time of 500 ms, and 128×4096 complex points were collected.

the result from a hyperpolarized experiment including beads without biotin treatment (Exp no. 7 in Table 2.1), which had similar protein and ligand concentrations as in the control experiment. Contributions from free ligand and specific binding ($\sigma = 0.017$ s^{-1}) were subtracted from the observed cross-relaxation rate of -0.055 s^{-1} and no effects on the specific binding fraction from non-specific binding was assumed ($X_b = 0.012$). The cross-relaxation rate between H_a and H_b in only specifically bound HABA, was determined as $\sigma_{b(p\text{-}b,\text{specific})} = (\sigma - \sigma_c)/X_b = -6.0$ s^{-1}.

The relationship between the cross-relaxation rate and molecular motions can be quantified using relaxation theory. Dipolar relaxation in a rigid body was assumed for calculating the rotational correlation time for the HABA bound to the immobilized protein on the bead surface. The measured cross-relaxation rate of -9.1 s^{-1} for overall binding
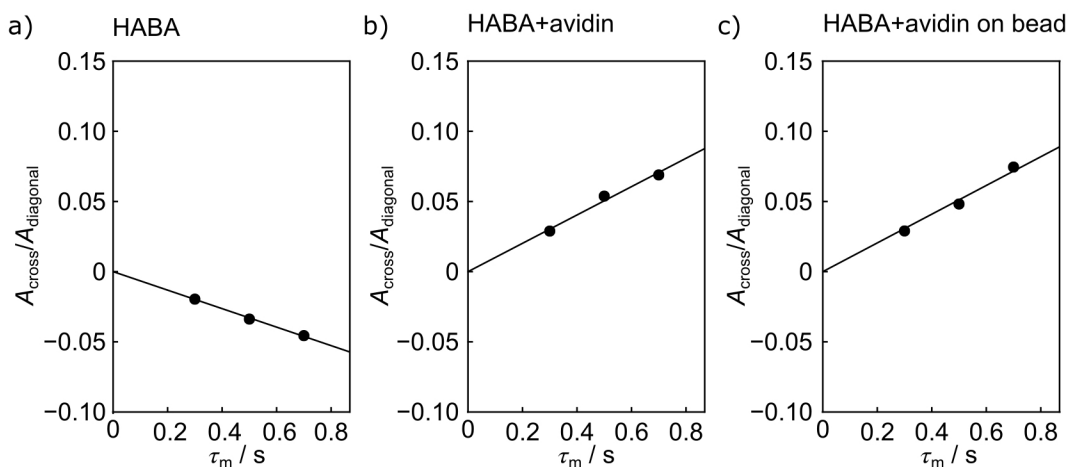
Figure 2.8: NOE build-up curves in initial regime measured from ^{1}H-^{1}H NOESY spectra with mixing times of 300 ms, 500 ms, and 700 ms. A_{cross} represents the average integral of the cross peaks between two HABA proton spins, H_a and H_b, while $A_{diagonal}$ is the average integral of the two diagonal peaks for spin H_a and spin H_b. Cross-relaxation rates are obtained by a linear fit of $A_{cross}/A_{diagonal}(\tau_{mix})$.[156] Sample conditions and corresponding fitted cross-relaxation rates are a) 100 μM HABA, $\sigma = 0.066$ s^{-1}; b) 100 μM, HABA with avidin (20 μM binding site), $\sigma = -0.101$ s^{-1}; c) 100 μM HABA with avidin immobilized on polystyrene beads (2.25 μM binding site),$\sigma = -0.102$ s^{-1}.

between spin H_a and H_b with a distance of 250 pm,[158] corresponds to a correlation time of $\tau_c = 39$ ns. The determined cross-relaxation rate of -6.0 s^{-1} for only specific binding leads to $\tau_c = 26$ ns. The correlation time for a rigid polystyrene bead with a diameter of 7 μm by Stokes' law would be on the second time scale, however, the correlation time for the bound ligand on the bead surface should be limited to the lifetime of the bound form. With the above determined K_d, and assuming the diffusion limited $k_{on} = 1 \times 10^8$ M^{-1}s^{-1}, this lifetime is 1.6 ms. The rigid-body correlation times in the nanosecond range calculated above are much smaller than this value, indicating the existence of internal motions.

Firstly, to evaluate local motions of the bound ligand, HABA bound to purified avidin in solution was analyzed. The cross-relaxation rate between the spins H_a and H_b in the complex was determined from a series of NOESY spectra as $\sigma_{b(p)} = -0.82$ s^{-1} (Table 2.2
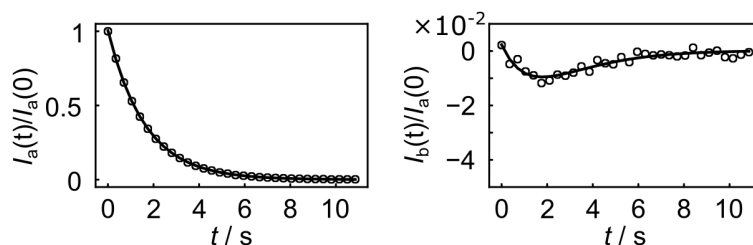
Figure 2.9: Fitting of the time-dependent signal integrals of the unsuppressed peak (left) and the pre-suppressed peak (right) in the control experiment with hyperpolarized HABA (58 µM) and biotin-capped avidin beads (0.8 µM binding site) after suppression of spin H_b. The fitted cross-relaxation rate between H_a and H_b is $\sigma = 0.017$ s^{-1}.

and Figures 2.7 and 2.8). The corresponding rigid-body rotational correlation time of τ_c = 3.6 ns is also smaller than the correlation time of 21 ns estimated by Stokes' law for a 66 kDa protein.[24] It can be speculated that this comparably low value may be related to the tautomerization of the ligand that occurs upon binding (Figures 2.3).[158,159] Secondly, the protein motion may be weakly restricted to the bead surface, and hence only partially acquire the motional properties of the large beads. Despite that the ligand does not assume the largest possible correlation time based on the rigid-body assumption, the trNOE with bead linked protein is enhanced 10-fold compared to protein in solution. In other cases, for example if the protein is embedded in a cell membrane as in ref. [54], more rigidly restricted protein motions and a concomitantly even more efficient NOE are expected.

Immobilized proteins provide significant advantages for the detection of ligand binding, because the large size of the beads increases the overall correlation time and hence the sensitivity of NOE signals. The polystyrene beads of micrometer size can be easily separated from the solution by centrifugation. The resulting possibility of reusing the sample improves the potential for applications to high-throughput screening. The signal enhancement provided by D-DNP can overcome limitations that are introduced by a reduced target density in immobilized samples. It allows detection of the trNOE within

40

a few seconds with the target at high nanomolar concentrations. The short measurement time required with D-DNP, combined with the rapid injection system,[151] ensures efficient mixing during the entire NMR acquisition, which is a major concern for NMR experiments involving heterogenous samples. Since the hyperpolarization enhances exclusively the ligand signal, background signals from other compounds including the bead support, which are not hyperpolarized, can be neglected.

In addition to the detection of binding, the determined cross-relaxation rates contain information about distances between ligand spins in the bound form. Like in conventional trNOE experiments, direct calculation of inter-proton distances from the measured cross-relaxation rates may not be possible due to the complexity in modeling internal motions and the possibility of non-specific binding. However, comparison between cross-relaxation rates determined for different pairs of ligand protons can provide information about the bound ligand conformation. In the D-DNP experiment described here, signal suppression was applied to only one of the ligand spins, so that polarization transfer from all other ligand ^{1}H spins to the target spin is observed. If more than one proton is in range to serve as the NOE source, this method does not supply the necessary information to distinguish the polarization origin. In this case, however, fast multidimensional NMR approaches may be introduced for encoding different ligand frequencies, for example, using Hadamard spectroscopy.[135,136]

2.4 Conclusions

In summary, the enhancement provided by hyperpolarization through D-DNP, combined with an increase in the efficiency of the NOE in an immobilized target, allows detection of trNOE at a target concentration in the submicromolar range, in a single scan. Besides the detection of binding, structure related information can be obtained from the cross-relaxation rates between proton pairs. This experiment demonstrates that the

large signal enhancement and short acquisition time make the dissolution DNP technique especially suitable for NMR experiments in heterogeneous environments.

3. SITE SPECIFIC POLARIZATION TRANSFER FROM A HYPERPOLARIZED LIGAND OF DIHYDROFOLATE REDUCTASE[1]

3.1 Introduction

Observing spin polarization transfer based on the nuclear Overhauser effect (NOE) is one of the most direct ways to confirm the existence of an intermolecular interaction. Applied to proteins, NOE transfer allows the determination of protein–protein binding interfaces,[160,161] as well as the identification of binding pockets for protein–ligand interaction.[162,163] The NOE is typically manifested as a small fractional change in signal obtained from a nuclear spin after perturbation of the equilibrium Zeeman population of another, nearby spin. In the case of intermolecular interactions, the magnitude of the spin polarization transferred due to the NOE is further reduced if the binding sites on target molecules are only fractionally occupied. Consequently, observation of intermolecular NOEs can be subject to important sensitivity limitations. A powerful way of increasing NOE intensity, however, is through hyperpolarization. A hyperpolarized source spin exhibits a deviation from equilibrium spin polarization that is orders of magnitude larger than that of the simple population inversion achievable with the application of a radio-frequency pulse.

The spin polarization-induced nuclear Overhauser effect (SPINOE) from hyperpolarized xenon has been used to enhance signals of hydrophobic cavities in proteins.[164] Surface accessibility of tryptophan residues furthermore can be studied in detail using chemically induced dynamic nuclear polarization, by a cyclic reaction with flavin.[165] In this case, a change in solvent accessibility during protein folding has

been found. Polarization transfer from molecules that are directly hyperpolarized on proton spins using dissolution dynamic nuclear polarization[113] (D-DNP) is also possible. Ultrafast 2D NMR spectra of polarization transferred from water to exchangeable amide protons in proteins were reported.[166] In our previous work, we have demonstrated that polarization transfer from a specifically binding hyperpolarized ligand shows a spectrum similar to the frequency dependence of the transfer of saturation from the protein to the ligand.[123] Using D-DNP hyperpolarization, a two-step polarization transfer between competitively binding ligands, mediated by the protein, can be observed similar to the Interligand NOE for Pharmacophore Mapping (INPHARMA) experiment.[42,137] While in the latter case the measured interligand NOE transfer rates are indicative of ligand orientation, additional structural information would be available from spectra of protein resonances, provided that they can be obtained with sufficient resolution.

Dihydrofolate reductase (DHFR) is an essential enzyme in both prokaryotes and eukaryotes, which reduces dihydrofolic acid (DHF) to tetrahydrofolic acid (THF) in the presence of the cofactor, dihydronicotinamide adenine dinucleotide phosphate (NADPH). Owing to its function of maintaining the cellular levels of THF and its derivatives, DHFR is an important enzyme involved in the folate cycle. This cycle produces precursors for purine and thymidylate synthesis. Hence, DHFR serves as a classic drug target and one of the most well studied enzymes.[167] Established antifolate drugs include antibacterial compounds trimethoprim (TMP)[168] and the anticancer agent methotrexate (MTX).[169] While *Escherichia coli* (*E. coli*) DHFR has been most extensively studied, the amino acids required for catalysis and the general features of the secondary structures are conserved. For example, both the E. coli and the human protein have the binding site located at the junction of two subdomains.[170,171]. Structural characterization by X-ray crystallography[172–177] and NMR[178,179] have revealed interactions of *E. coli* DHFR with various ligands.

Here, we used *E. coli* DHFR as a model protein to introduce a series of D-DNP NMR experiments that allow for the site specific resolution of NOE signals transferred from the ligand to the protein, by employing indirect selection based on ^{13}C chemical shift. We demonstrate selectivity by mapping side-chain resonances of the binding pocket of DHFR for the ligand folic acid, and discuss the utility of this method for obtaining limited structural information on the binding site.

3.2 Experimental Section

DHFR from *E. coli* was expressed from a plasmid pET-Duet-1 in *E. coli* BL21(DE3) cells. For unlabeled DHFR, cells were grown in LB medium and induced with 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) at 37 °C. Uniformly ^{13}C/^{15}N-labeled DHFR was prepared using a protocol modified from Marley *et al.*[180] Briefly, transformed cells were inoculated in 4 × 1 L LB medium and grown overnight at 37 °C and 250 rpm in an incubator shaker (Brunswick Instruments, New Brunswick, NJ) to an optical density of 0.6. Cells were centrifuged for 15 min at 5000×g and 4 °C, resuspended in 200 mL of M9 minimum medium (without glucose and ammonium chloride) and centrifuged again. The resuspension and centrifugation procedure was repeated a total of three times, in order to remove unlabeled growth medium components. Cells were subsequently resuspended in 1 L of M9 minimal medium containing 3 g of ^{13}C$_6$-glucose and 1 g of ^{15}N-ammonium chloride and incubated at 37 °C and 250 rpm for 1 h. Protein expression was induced by addition of 1 mL of 1 M IPTG. After induction, cells were grown for an additional 20 h at 16 °C. Cells were harvested by centrifugation (5000×g, 15 min, 4 °C), resuspended in 50 mL of buffer A (20 mM sodium phosphate, 0.5 M NaCl, 5 mM imidazole, pH 7.4) and lysed by sonication on ice for 10 min. Cell lysate was centrifuged (10,000×g, 1 h, 4 °C), and the supernatant was loaded onto a 5 mL HisTrap HP column (GE Healthcare, Pittsburgh, PA). After washing with buffer A, the protein was eluted using buffer B (20

mM sodium phosphate, 0.5 M NaCl, 0.5 M imidazole, pH 7.4) with a linear gradient. Final purification was achieved by gel filtration using a Sephacryl S100 column (GE Healthcare) with 50 mM potassium phosphate, 1 mM EDTA, 1 mM 1,4-dithiothreitol (DTT), 150 mM NaCl, pH 6.8. The purified protein was concentrated to 4.5 mM using a centrifugal filter device with 3 kDa molecular weight cut-off (EMD Millipore, Billerica, MA).

NMR spectroscopy was performed using a sample of 1.5 mM DHFR in 50 mM potassium phosphate, 1 mM EDTA, 1 mM DTT, 50 mM KCl, pH 6.8/10 % D_2O, with 15 mM folic acid, at a temperature of 298 K. Backbone and side chain chemical shift assignments were mapped from Falzone et al.,[179] using spectra acquired on a 500 MHz NMR spectrometer with a TCI cryoprobe (Bruker Biospin, Billerica, MA). The spectra for assignment were HNCO ($128 \times 48 \times 1024$ complex points in ^{13}C, ^{15}N and ^{1}H dimensions with $t_{1,max}$ = 31.8, 13.9 and 81.9 ms, respectively), HNCA ($96 \times 64 \times 1024$ complex points in ^{13}C,, ^{15}N and ^{1}H dimensions with $t_{1,max}$ = 12.7, 19.7 and 81.9 ms, respectively), HNCACB ($160 \times 64 \times 1024$ complex points in ^{13}C, ^{15}N and ^{1}H dimensions with $t_{1,max}$ = 9.8, 18.6 and 81.9 ms, respectively), H(CC)(CO)NH-TOCSY ($128 \times 40 \times 2048$ complex points in ^{1}H, ^{15}N and ^{1}H dimensions with $t_{1,max}$ = 16.0, 12.3 and 163.8 ms, respectively, τ_m = 15 ms), (H)CC(CO)NH-TOCSY ($170 \times 44 \times 1024$ complex points in ^{13}C, ^{15}N and ^{1}H dimensions with $t_{1,max}$ = 10.2, 13.6 and 81.9 ms, respectively, τ_m = 15 ms), and HCCH-TOCSY ($128 \times 64 \times 2048$ complex points in ^{1}H, ^{13}C and ^{1}H dimensions with $t_{1,max}$ = 12.8, 3.4 and 163.8 ms, respectively, τ_m = 15 ms). NOEs between ligand and protein were identified from the same sample by acquiring a HSQC-NOESY ($80 \times 128 \times 2048$ complex points in ^{13}C, ^{1}H, and ^{1}H dimensions with $t_{1,max}$ = 4.2, 9.8 and 157.7 ms, respectively, τ_m = 500 ms).

For hyperpolarized NMR, aliquots of 10 μL volume were prepared containing 225 mM folic acid and 15 mM 4-hydroxy-2,2,6,6-tetramethylpiperidin-1-oxyl (TEMPOL) in D_2O/DMSO-d_6 (1:1 v/v). Aliquots were irradiated in a HyperSense DNP polarizer (Oxford

46

Instruments, Abingdon, UK) at a microwave frequency of 94.005 GHz optimized for
^{1}H polarization, microwave power of 100 mW and temperature of 1.4 K. After 20 min,
hyperpolarized samples were dissolved in 4 mL of 50 mM potassium phosphate, pH 6.8,
which had been pre-heated in a closed vessel until a pressure of 1 MPa had been reached
(time point a in Figure 3.1). Dissolved samples were taken into the loop of a sample
injector device.[151] A volume of 450 µL was injected into a 5 mm NMR tube, during 380
ms (starting from time point b, and finishing at time point c in Figure 3.1) using nitrogen
gas at a pressure of 1.81 MPa against a back pressure of 1.03 MPa. The empty NMR
tube, or the NMR tube that had been loaded with 25 µL of 4.5 mM DHFR in 50 mM
potassium phosphate, pH 6.8, had been pre-installed in a 400 MHz NMR spectrometer
equipped with a broadband observe (BBO) probe (Bruker Biospin, Billerica, MA) set to a
temperature of 298 K. In the case of the NMR tube loaded with DHFR, the hyperpolarized
folic acid mixed with the protein during injection. In both cases, NMR spectroscopy was
performed starting at time point d in Figure 3.1, immediately following a delay of 500 ms
after completion of injection.

^{1}H NMR spectra of hyperpolarized folic acid alone were acquired using a single $\pi/2$
pulse with water suppression by selective excitation. The same pulse sequence was used
for acquiring spectra of DHFR in the presence of hyperpolarized folic acid, when not
using isotope selection. NMR data with ^{13}C isotope selection were acquired using single
quantum coherence transfer[181] (Figure 3.1). For the refocusing pulse at the center of the
^{13}C coherence evolution time, a Gaussian shape with 4000 or 7650 ms duration and 1%
truncation level was used for chemical shift selection, or a hard pulse with $\gamma B_1/2\pi = 27.8$
kHz for broad band ^{13}C selection. GARP decoupling was applied on ^{13}C with $\gamma B_1/2\pi$
= 3.1 kHz during acquisition, WALTZ-16 decoupling on ^{1}H with $\gamma B_1/2\pi = 3.1$ kHz
can be turned on during the selective refocusing pulse for improved selectivity on ^{13}C.
Spectra were processed with an exponential window function (25 Hz line broadening) and

polynomial baseline correction using TopSpin 3.2 (Bruker Biospin).

3.3   Results and Discussion

A ^{1}H NMR spectrum of hyperpolarized folic acid is shown in Figure 3.2. When compared to the spectrum from a sample thermally polarized in the NMR magnet, all peaks for the ligands are enhanced significantly. The largest enhancement factors are 910 for H7, 700 for H2′/H6′, and 610 for H3′/H5′, while they vary between 100 and 200 for other folic acid peaks. For obtaining these spectra, a rapid injection system[151] was used, allowing a sample transfer time of approximately 1200 ms. Polarization loss during the sample injection period was thereby minimized.

Due to the hyperpolarization of the ligand, polarization transfer to the protein leading
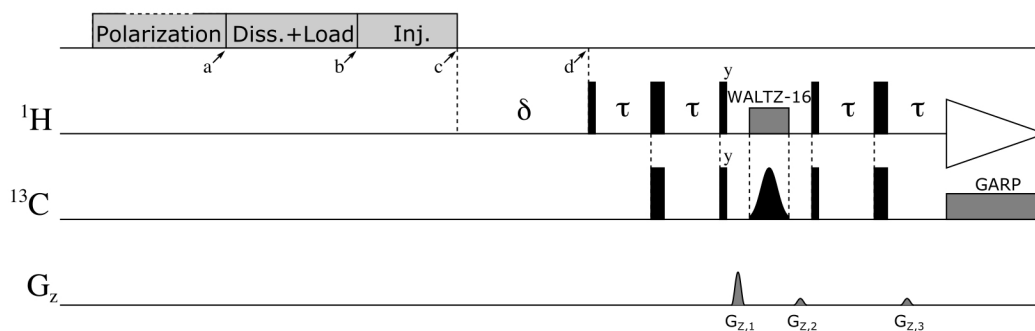
Figure 3.1: Experimental time line including sample polarization (20 min), dissolution ("diss."), loading into the injector loop, injection into the NMR instrument ("inj."; 380 ms), delay time δ = 500 ms and pulse sequence for DNP-NOE experiment. In the pulse sequence, narrow black bars represent π/2 pulses ($\gamma B_1/2\pi$ = 21.6 kHz for ^{1}H, $\gamma B_1/2\pi$ = 27.8 kHz for ^{13}C). The pulse designated with a shape at the center of the ^{13}C coherence evolution time was applied either as a selective pulse for chemical shift selection, or as a hard π pulse for broad band ^{13}C selection (see text). Simultaneous proton decoupling during the selective carbon pulse can also be introduced by applying a WALTZ-16 pulse train ($\gamma B_1/2\pi$ = 3.1 kHz) (see text). GARP decoupling ($\gamma B_1/2\pi$ = 3.1 kHz) was applied on ^{13}C during acquisition. The delay was τ = 1.79 ms. Pulsed field gradients for coherence selection were $G_{z,1}$ = 32.1 G·cm^{-1}, $G_{z,2}$ = 6.5 G·cm^{-1} and $G_{z,3}$ = 6.5 G·cm^{-1}. Reprinted with permission from [148].
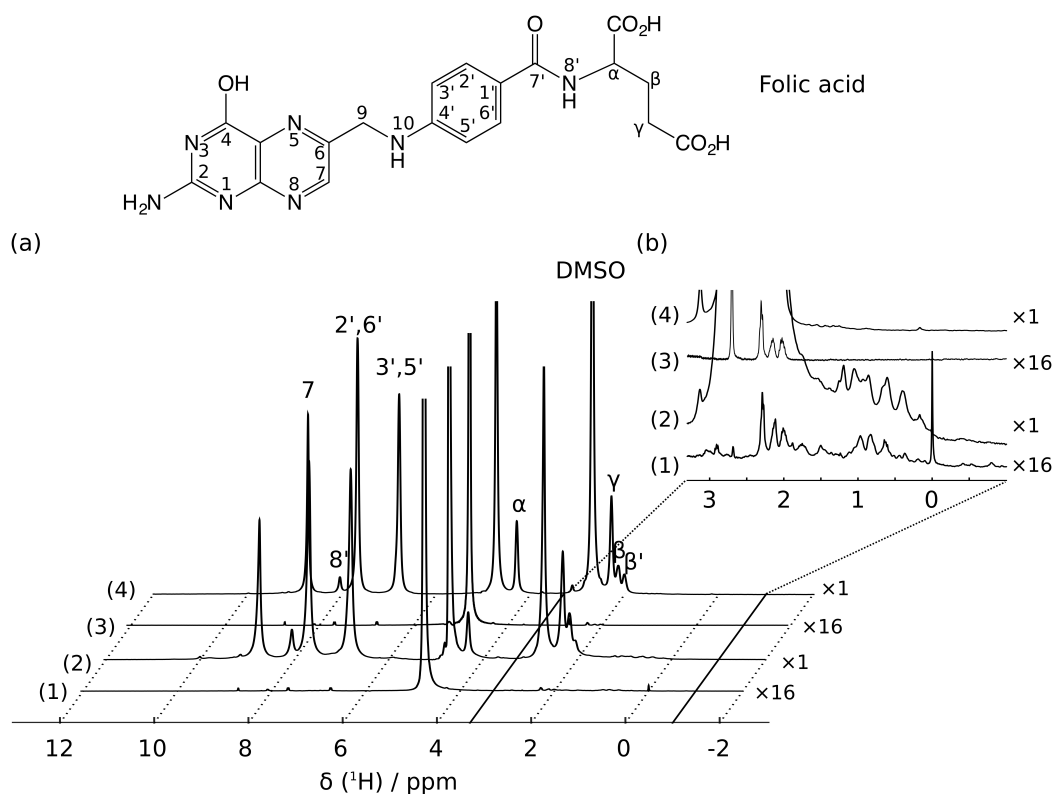
Figure 3.2: ^{1}H NMR spectra a) in full scale and b) with aliphatic region enlarged to show enhanced protein peaks acquired with 1 thermally polarized folic acid with DHFR; 2 hyperpolarized folic acid mixed with preloaded DHFR; 3 thermally polarized folic acid without DHFR and 4 hyperpolarized folic acid without DHFR. Thermal spectra 1 and 3 are rescaled as 16 times of the original intensities. Reprinted with permission from [148].

to visible signals even in a single scan can be expected.[123] Enhanced signals in spectral regions characteristic of the protein, in particular side-chain protons, are indeed readily seen upon admixing of DHFR. Since these signals are not present in the sample of folic acid alone, and are not observable in a sample of protein without hyperpolarized ligand, it is apparent that these signals originate from nuclear Overhauser (NOE) transfer from the ligand to the protein. Although polarization transferred from a ligand to the protein specifically originates at the location of the binding site, the enhanced signals are expected to stem from a sufficiently large number of protons such that their individual identification

49

is generally precluded using ^{1}H NMR alone.

In order to resolve overlapped protein ^{1}H peaks, one-dimensional spectra of the protein were acquired using a single-quantum filter for ^{13}C chemical shift selection. This pulse sequence, described in the Experimental section, contains a selective refocusing pulse that can be adjusted for chemical shift position and selectivity. Spectra using filters with a full width at half maximum of 690 and of 210 Hz are shown in Figure 3.3a, b, respectively. For comparison, a spectrum obtained with broad band ^{13}C filter is shown in Figure 3.3c. This spectrum is expected to contain all of the protein signals that received sufficient NOE transfer from the hyperpolarized ligand. At the same time, the ^{13}C filter removes the coherences from unlabeled compounds and in this experiment ensures that the observed signals are from the protein. From the figure, it can be seen that the signals obtained with broad selection approximately represent the combination of the sets of signals observed in the individual traces. An increased selectivity with the narrower filter is further seen in the set of spectra in Figure 3.3b, where fewer peaks are observed in each trace compared to the spectra in Figure 3.3a. The frequency selection was achieved using Gaussian shaped pulses with the duration of 4000 or 7650 ms, in the case of the longer pulse also including simultaneous ^{1}H decoupling to prevent undesired effects due to J-coupling during the pulse time (Figure 3.3d).

In order to relate observed chemical shifts to the structure, chemical shift assignments (Figure 3.4 and 3.5) were mapped from Falzone *et al.*,[179] using triple resonance experiments to adjust for the sample conditions used as described in Materials and Methods. From the crystal structure of the complex ,[177] methyl protons located within 0.7 nm of one of the three ligand protons H7, H3′/H5′ or H2′/H6′ were further identified. Vertical lines at the chemical shift positions of these nearby methyl protons are drawn in Figure 3.3a, b, and corresponding residue numbers are indicated at the top. Since the spectra acquired with selection at different ^{13}C chemical shifts show distinct patterns, it
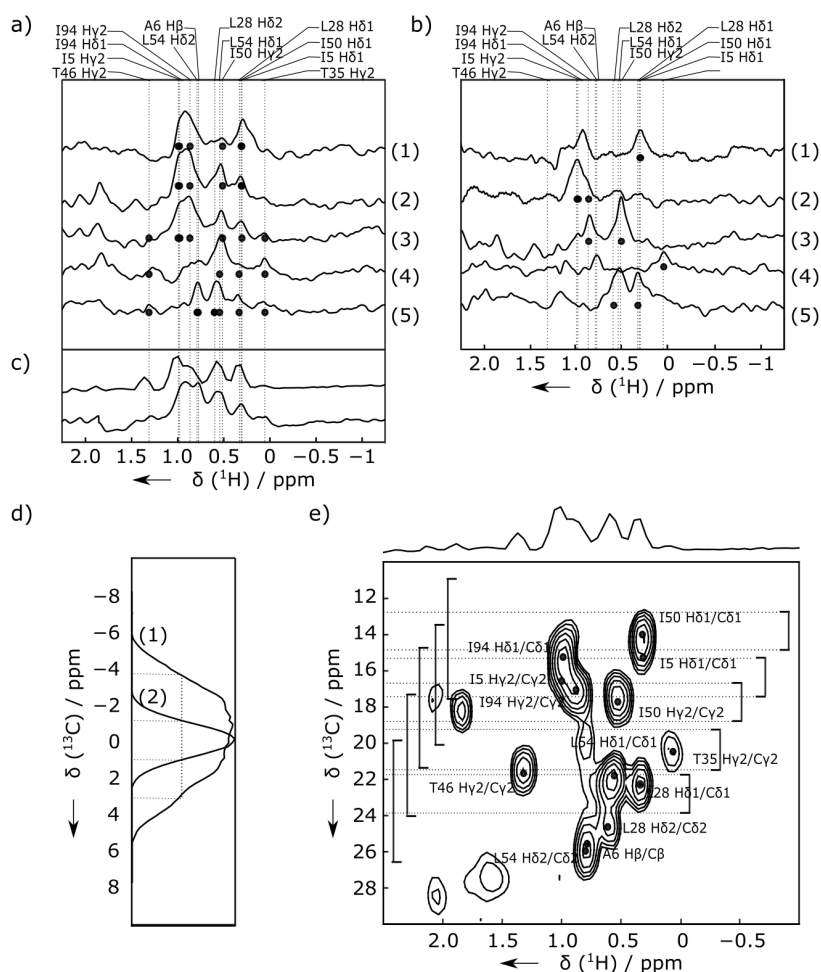
Figure 3.3: Comparison between DNP-NMR spectra and 3D filtered NOESY spectrum for observation of intermolecular NOE peaks between DHFR and folic acid. a) DNP-NMR Spectra of DHFR in the presence of hyperpolarized folic acid, acquired with chemical shift selection at ^{13}C positions of (1) 13.7 ppm; (2) 16.2 ppm; (3) 17.7 ppm; (4) 20.2 ppm; (5) 22.7 ppm and width of 690 Hz. Dotted lines represent the ^{1}H chemical shift of estimated NOE peaks based on ^{1}H–^{1}H distance (cutoff 0.7 nm) calculated from X-ray crystal structure,[177] and solid dots indicate the estimated NOE peaks within the excitation region. b) Spectra as in (a), but with selection width of 210 Hz. c) Comparison of spectra of DHFR in the presence of hyperpolarized folic acid, acquired with a hard excitation pulse (bottom) and 1D projection of the NOESY spectrum (top) from (e). d) Experimentally determined excitation profiles of the pulse sequence used for selective excitation in the DNP-NMR experiments. e) Superposition of three distinct ω_3 (ligand ^{1}H) planes (H7: 8.71 ppm, H2′/H6′: 7.65 ppm, H3′/H5′: 6.74 ppm) from the 3D HSQC-NOESY of ^{13}C,^{15}N-DHFR with unlabeled ligand. Estimated NOE peaks are calculated as in (a), and selection ranges are indicated. A 1D projection is shown at the top. Reprinted with permission from [148].

Figure 3.4: ^{1}H-^{15}N HSQC spectrum of uniformly ^{13}C/^{15}N-labeled DHFR complexed with folic acid. (256 × 2048 complex points in ^{15}N and ^{1}H dimensions). Reprinted with permission from [148].

appears to be useful to map the observed peaks to the known ^{13}C and ^{1}H chemical shifts in the protein. For example, signal intensity at 0.32 ppm, which is near the proton resonance of Hδ1 for Ile50 (Hδ1: 0.32 ppm, Cδ1: 13.98 ppm), can be predominantly observed in the first spectra in Figure 3.3a, b with selection pulse centered at 13.7 ppm. Furthermore, for almost every peak appearing in the DNP-NMR spectra between 0 and 1.5 ppm, a corresponding nearby protein proton within the selection range can be identified. In other words, there is a strong correlation between the observed signal intensities in the DNP-NMR spectra and the short ^{1}H–^{1}H distances identified from the X-ray crystal structure.

Figure 3.5: ^{1}H-^{13}C HSQC spectrum of uniformly ^{13}C/^{15}N-labeled DHFR complexed with folic acid. (128 × 1024 complex points in ^{13}C and ^{1}H dimensions). Reprinted with permission from [148].

Protein–ligand NOEs can also be determined using conventional, non-hyperpolarized NMR spectroscopy—albeit not in a single scan, but rather in a multi-dimensional data set acquired over the course of several days. Specific intermolecular protein–ligand interactions can be investigated using isotope-filtered NMR methods,[32,182] such as in a 3D ^{13}C-edited, ^{13}C/^{15}N-filtered HSQC-NOESY spectrum using uniformly ^{13}C/^{15}N-labeled protein combined with unlabeled ligand. Such a 3D HSQC-NOESY spectrum was measured with a mixing time of 500 ms. This spectrum contains exclusively intermolecular NOE peaks between the unlabeled folic acid and ^{13}C/^{15}N-labeled DHFR with protein ^{13}C resonances in the ω_1 dimension, protein ^{1}H resonances in the ω_2 dimension and ligand ^{1}H resonances in the directly acquired ω_3 dimension.

53

NOE peaks between protein proton and different ligand protons can be obtained from analysis of the $[\omega_1/\omega_2]$ planes at each folic acid ^{1}H chemical shift. Most of these cross peaks appear at the methyl group chemical shifts in the spectrum of the protein. These signals can be compared to the spectra obtained from the protein after NOE transfer from hyperpolarized ligand, which were acquired with a corresponding waiting time of 500 ms after admixing of the hyperpolarized ligand. The contact time in the DNP spectra is somewhat longer than 500 ms because of the sample mixing that occurs during part of the injection time. With the goal of comparing this 3D NOESY data with hyperpolarized spectra, it is necessary to consider the contribution of multiple ligand spins in both cases. Given that the signals from protons H7, H3$'$/H5$'$ and H2$'$/H6$'$ combined comprise 69% of the total signal intensity of folic acid in the spectrum of Figure 3.2a, the intensities in the planes of the 3D-NOESY spectrum corresponding to magnetization transferred to the ligand protons at three chemical shifts were added and shown in Figure 3.3e (individual planes are shown in Figure 3.6). No attempt was made to exactly scale the individual contributions since the integrals of these three signals vary in the 1D ^{1}H NMR spectrum of the hyperpolarized ligand. Differences arise due to different numbers of protons and other factors including differing polarization levels and varying polarization transfer efficiency.

The ^{13}C and ^{1}H chemical shift positions of the amino acid side chains in proximity to these protons are marked in Figure 3.3e, with the corresponding residue number indicated. Almost all of the observed signal intensity in the spectrum is located at the positions identified using the distance ranges described above. In order to compare the DNP hyperpolarized spectra with the conventional NOESY spectrum, the regions corresponding to the full width at half maximum of the selection ranges in the DNP hyperpolarized spectra are further indicated in Figure 3.3e. Almost all peaks shown in the DNP spectra in Figure 3.3a, b can be correlated to the NOE peaks in the NOESY spectra, within the respective selection range. These results indicate that a series of DNP NOE experiments

54
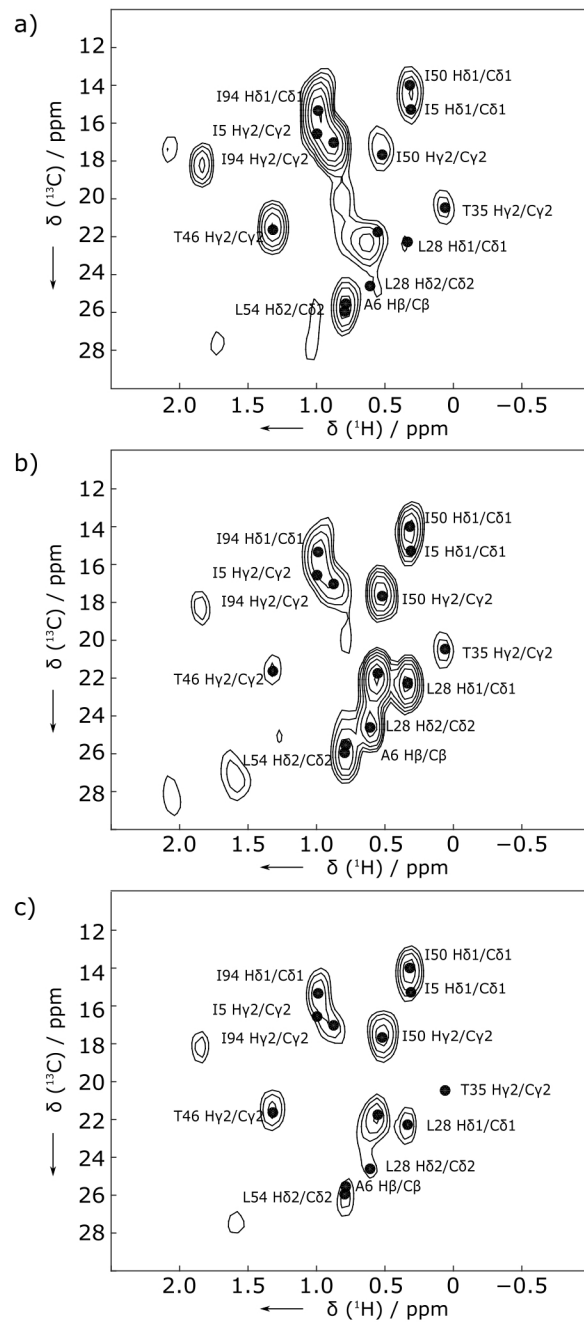
Figure 3.6: ω_3 (folic acid ^{1}H) planes from the 3D HSQC-NOESY (^{13}C filter in ω_3, τ_{m}= 500 ms) of ^{13}C/^{15}N-labeled DHFR with unlabeled ligand. a) H7: 8.71 ppm. b) H2'/H6': 7.65 ppm. c) H3'/H5': 6.74 ppm. Reprinted with permission from [148].

can provide intermolecular distance information with correlated ^{13}C chemical shift, similar to a 3D filtered HSQC-NOESY experiment.

The correlation of observed crosspeaks with the distances between protein and ligand ^{1}H can further be visualized by the structure shown in Figure 3.7 (numerical distance values are listed in Table 3.1). The DNP-NMR spectra contain intensity at the chemical shift positions of the methyl groups of 6 amino acids; Ile5, Leu28, Thr35, Ile50, Leu54,
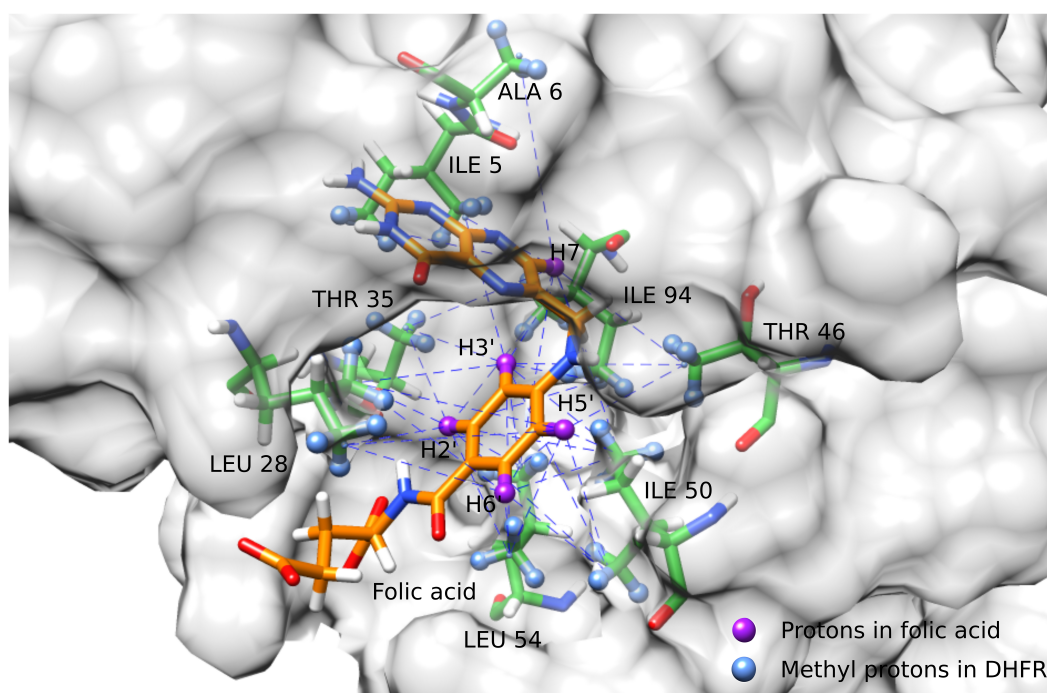
Figure 3.7: .
Close distances (cutoff: 0.7 nm) between folic acid and DHFR methyl protons are indicated with dashed lines. H7, H2′, H3′, H6′, H7′ of folic acid are shown as purple spheres. Blue spheres indicate methyl protons in DHFR that are within 0.7 nm from the labeled folic acid protons. Image created with UCSF Chimera software version 1.10.2.[183] Crystal structure of folic acid-DHFR complex [Protein Data Bank 1RE7,[177] chain A]. Close distances (cutoff: 0.7 nm) between folic acid and DHFR methyl protons are indicated with dashed lines. H7, H2′, H3′, H6′, H7′ of folic acid are shown as purple spheres. Blue spheres indicate methyl protons in DHFR that are within 0.7 nm from the labeled folic acid protons. Image created with UCSF Chimera software version 1.10.2.[183] Reprinted with permission from [148].

56

	δ(^{1}H)/ppm	δ(^{13}C)/ppm	Ligand proton	Distance/Å
Ile5 Hγ2/Cγ2	1.00	16.54	H7	4.68
			H5′	6.49
Ile5 Hδ1/Cδ1	0.31	15.26	H7	6.66
Ala6 Hβ/Cβ1	0.78	25.53	H7	6.37
Leu28 Hδ2/Cδ2	0.61	24.60	H2′	5.31
			H3′	6.12
			H5′	5.07
			H6′	3.97
Leu28 Hδ1/Cδ1	0.34	22.28	H2′	4.95
			H3′	6.31
			H6′	5.91
Thr35 Hγ2/Cγ2	0.06	20.44	H7	6.94
			H5′	6.79
			H6′	6.72
Thr46 Hγ2/Cγ2	1.32	21.62	H7	4.97
			H3′	5.17
			H5′	5.25
Ile50 Hδ1/Cδ1	0.32	13.98	H7	5.93
			H2′	5.26
			H3′	4.42
			H5′	4.34
			H6′	5.16
Ile50 Hγ2/Cγ2	0.52	17.67	H2′	4.86
			H3′	5.36
			H5′	6.93
			H6′	6.48
Leu54 Hδ1/Cδ1	0.32	13.98	H7	6.43
			H2′	6.50
			H3′	6.78
			H5′	4.38
			H6′	3.87
Leu54 Hδ2/Cδ2	0.52	17.67	H2′	5.22
			H3′	6.41
			H5′	5.90
			H6′	4.49
Ile94 Hγ2/Cγ2	0.87	17.05	H7	4.48
			H5′	5.67
			H6′	6.90
Ile94 Hδ1/Cδ1	0.99	15.34	H7	4.50
			H5′	4.62
			H6′	5.91

Table 3.1: ^{1}H and ^{13}C chemical shifts for methyl groups of DHFR located within 0.7 nm from one of the folic acid protons H2′, H3′, H5′, H6′ and H7 and all distances less than 0.7 nm between the methyl protons (center of mass of three protons) and these five ligand protons. Reprinted with permission from [148].

Ile94. Their positions in the protein–ligand complex are all in the vicinity of the ligand molecule, as shown in the figure. For example, side chains of amino acid Ile50 (Hδ1: 0.32 ppm, Cδ1: 13.98 ppm; Hγ2: 0.52 ppm, Cγ2: 17.67 ppm) are in close contact with the (p-aminobenzoyl)-glutamate tail of folic acid. Hyperpolarization likely transferred from the highly polarized phenyl protons (H2′, H3′, H5′, H6′) yielded two strong peaks around 0.32 and 0.52 ppm (Figure 3.3a, b). In Figure 3.3b, where the selectivity in the ^{13}C dimension is increased, the former peak only appeared in the first spectrum (^{13}C selection at 13.7 ppm), whereas the latter can only be observed in the third spectrum with ^{13}C selection at 17.7 ppm. This difference is in accordance with the ^{13}C chemical shifts of the two methyl groups, which are 13.98 and 17.67 ppm, respectively. Based on this and other, similar observations made by comparing Figure 3.3a, b with Figure 3.7, it becomes apparent that structural information of protein–ligand complexes is available from the D-DNP experiments.

The hyperpolarized experiment with the ^{13}C single quantum filter is akin to a doubly selective conventional NMR experiment. In the DNP experiment, a first effective selection step is for the ligand, which is hyperpolarized to a level exceeding thermal polarization by a factor of at least 10^2–10^3. Since any NOE signals originating from the hyperpolarized ligand are amplified by this factor, many signals from non-polarized sample components—foremost the signals from the thermally polarized protein present at sub-millimolar concentration—can be neglected. A second selection step is represented by the spectroscopically applied filter, which primarily retains the signal of the ^{13}C enriched protein, with a selectivity of up to 1:99 based on the natural abundance of this isotope in unlabeled compounds. Therefore, the experiment becomes sensitive to the intermolecular NOE between the ligand and the protein alone. With the chemical shift selection, residue specific identification of intermolecular NOE becomes possible—a prerequisite for the use of such NOEs in applications requiring distance constraints. Additional selectivity may be

achieved by other spectroscopic means, such as those based on J coupling constants or on coupling multiplicity. The efficiency of recording an entire dataset with correlations may further be increased by Hadamard spectroscopy,[135] which we have in the past also applied with D-DNP,[136] and similar correlations may also be obtained using single-scan ultrafast 2D approaches.[127] The traditional 3D ^{13}C-edited, ^{13}C/^{15}N-filtered HSQC-NOESY experiments shown here was acquired over a time of 38 h, while one D-DNP experiment can be performed with a polarization step requiring less than 30 min, and almost instantaneous NMR acquisition. This rapid data acquisition in particular would lend itself to future applications to samples that show only a transient stability.

3.4   Conclusions

While traditional high-resolution NMR spectroscopy provides detailed structural information on biological macromolecules, these studies typically rely on multi-dimensional data sets that require acquisition times on the order of days. Applications that involve more rapid changes in the samples under study are often precluded. On the other hand, recent developments in hyperpolarization have demonstrated the potential of D-DNP for the study of fast processes by acquiring highly sensitive spectra of various nuclei. In the most basic implementation, D-DNP data however provide one-dimensional spectra that do not contain the resolution necessary for the identification of most individual resonances in biological macromolecules. Here, we demonstrate the combination of rapid acquisition using DNP-NMR, with an experiment providing structural information through [^{13}C, ^{1}H] single quantum selection in a protein of molecular weight of 18.8 kDa, which is too large for resolution of most individual peaks using homonuclear spectroscopy alone. This and other heteronuclear selection mechanisms provide a means for obtaining structural constraints in biomolecules using hyperpolarization by D-DNP.

4. DETERMINATION OF LIGAND BINDING EPITOPE STRUCTURES USING POLARIZATION TRANSFER FROM HYPERPOLARIZED LIGANDS[1]

4.1 Introduction

Structure-based drug design (SBDD)[5] is a rational strategy for optimizing drug candidate molecules on the basis of ligand–target complex structures. Often, the structure of the macromolecular target is available, for example, from X-ray crystallography. The task then becomes determining the binding mode of the ligand at the binding site. Computational docking of small molecules into the macromolecular binding site is a high-throughput technique for this task and has become one of the most widely used approaches in drug development for lead optimization.[18,68] Often, in a docking protocol, a large number of ligand trial poses with varying conformations and orientations are generated. Subsequently, one or multiple scoring functions are used to rank the ligand poses and to select the one that best represents the native structure. The currently applied scoring functions are generally force field-based, empirical, or knowledge-based.[185] However, these energy-based scoring functions cannot always successfully discriminate the correct pose from alternative biding modes, leading to multiple solutions. Most of the scoring functions are not able to provide correct rankings of candidate poses, with poor correlations between the pose score and the structural root-mean-square deviation (RMSD) or experimental binding affinity.[82,186] Most docking simulations used in SBDD now still rely on a rigid receptor conformation, considering the resources and efforts needed to account for many degrees of freedom. Nevertheless, with the understanding of the significance of considering protein flexibility, more and more docking programs now

---

[1] Reprinted with permission from "Wang, Y.; Hilty, C. Site Determination of Ligand Binding Epitope Structures Using Polarization Transfer from Hyperpolarized Ligands. *J. Med. Chem.* **2019**, 62(5), 2419-2427.".[184] Copyright (2019) American Chemical Society

allow limited conformational changes for selected receptor side chains, which would be expected to improve the accuracy of the results for systems with only small displacements upon binding.[187–189] For systems with significant structural flexibility, the variety of receptor conformations can first be sampled using molecular dynamics.[190,191] One possible method for improving the accuracy of computational docking is to combine it with experimental structural constraints, such as those obtained by nuclear magnetic resonance (NMR) spectroscopy. Such constraints can be based on the intermolecular nuclear Overhauser effect (NOE),[104,192,193] chemical shift perturbations,[89,186] or saturation transfer difference.[194] In particular, [^{1}H,^{1}H]-NOE-derived distance information forms the basis for many structural determinations of macromolecules by NMR, aided by the selectivity imparted by an r^{-6} distance dependence.[195] NOEs in biological macromolecules are most typically measured in two- or three-dimensional NMR experiments. T he acquisition of multidimensional NMR spectra, however, is time-consuming and when used for ligand screening reduces throughput.

Hyperpolarization techniques can enhance NMR sensitivity by orders of magnitude through the use of a non-equilibrium spin population and thereby dramatically reduce experimental time. Hyperpolarized spin states in the liquid state, such as those generated using dissolution dynamic nuclear polarization (D-DNP), can be applied to the sensitive NMR detection on time scales of seconds or less. Transient interligand NOEs originating from D-DNP hyperpolarization have previously been exploited to characterize epitopes of competitively binding ligands.[137] Specific transfer of hyperpolarization from the ligand to the protein can also be directly observed.[123,148] The hyperpolarized signal is large enough to observe transient NOEs in a single scan. A challenge for the determination of distance information, however, lies in the fact that a single scan experiment is not directly compatible with the acquisition of indirect time dimensions in typical multidimensional NMR experiments. Two-dimensional spectra can still be obtained but require a distribution

of signals in a spatial dimension[127] or time sequential excitations on a single sample.[128]

The enzyme dihydrofolate reductase (DHFR) catalyzes the reduction of dihydrofolate (DHF) to tetrahydrofolate (THF). THF is a vital coenzyme involved in the purine and thymidylate synthesis and thus plays an essential role in cell proliferation.[167] As the enzyme that controls the production of THF, DHFR is a target for several drugs, including the anticancer drug methotrexate (MTX)[169] and the antibiotic drug trimethoprim (TMP).[168] Numerous structural studies have investigated how various ligands bind to DHFR from different species and how binding affects the enzymatic function, which could further provide guidance for rational structure-based drug design.[196,197] DHFR–ligand complex structures, in particular, have been determined by X-ray crystallography[177,198] and NMR spectroscopy.[179,199,200]

The authors of this study have previously explored protein-based ^{13}C chemical shift selection to identify interacting protein and ligand spins from simpler, one-dimensional hyperpolarized NOE experiments.[148] Here, we describe a method for the rapid determination of the structure of protein–ligand complexes by the combined use of computational docking methods and structural information obtained from hyperpolarized NOE experiments. With the folic acid–*Escherichia coli* DHFR complex as an example, we demonstrate that this method improves target selection compared to the use of an energy-based scoring function alone.

4.2   Experimental Section

Folic acid was purchased from Spectrum Chemicals (Gardena, CA). D_2O and dimethyl sulfoxide-d_6 (DMSO-d_6) were purchased from Cambridge Isotope Laboratories (Tewksbury, MA). Uniformly ^{13}C-labeled DHFR was expressed from plasmid pET-Duet-1 in *Escherichia coli* BL21(DE3) cells and purified as previously described.[148]

Samples (10 µL) containing 254 mM folic acid and 15 mM TEMPOL in D_2O/DMSO-

d$_6$ [1:1 (v/v)] were hyperpolarized at 1.4 K with a microwave frequency of 94.005 GHz and a power of 100 mW. A deuterated glassing matrix was chosen to avoid large solvent signals in ^{1}H measurements. Additionally, an increase in the attainable ^{1}H polarization level may be expected with deuteration.[201] After 20 min, the hyperpolarized sample was dissolved with 4 mL of preheated 50 mM potassium phosphate buffer (pH 6.8). The leading 0.42 mL of this solution was loaded into a sample loop and injected into the NMR tube, which already contained 30 μL of 3.5 mM [13]C- and ^{15}N-labeled DHFR. The injection was achieved in 0.44 s, using nitrogen gas at a pressure of 1.81 MPa, against a back pressure of 1.03 MPa.[151] Samples exited the DNP polarizer at a temperature of approximately 302 K and were without further adjustment injected into the NMR probe held at 298 K. The hyperpolarized ligand sample was mixed with the protein due to turbulence generated by the injection. The start of mixing was considered at the middle of the injection time (point a in Figure 4.1). The final concentrations of folic acid and DHFR in the NMR tube were ~1.6 and ~0.23 mM, respectively. The concentrations for each DNP–NMR experiment were determined by comparison to known standards, using ^{1}H NMR after completion of hyperpolarized experiments.

All NMR spectra were acquired using a 400 MHz NMR spectrometer equipped with a 5 mm broadband observe (BBO) probe (Bruker Biospin, Billerica, MA). Hyperpolarized ^{1}H NMR spectra were acquired using a pulse sequence with protein-side ^{13}C isotope selection using a single-quantum filter and ligand-side ^{1}H isotope selection using a selective inversion pulse. [13]C selection was set at 13.5, 15.5, and 17.5 ppm. The frequency offset from the water signal of the inversion pulse (point b in Figure 4.1) was set to 1590, 710, and 690 Hz, respectively, for the purpose of selectively inverting the polarization of folic acid protons H7, H2′/H6′, and H3′/H5′. A control experiment was also performed with the inversion pulse set off-resonance at 20000 Hz. To determine the polarization levels of ligand peaks in each experiment and the effectiveness of the selective inversion,
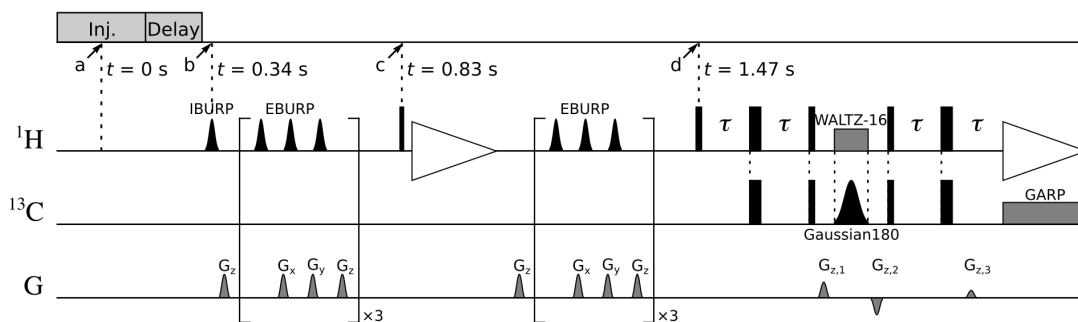
Figure 4.1: Pulse sequence for ^{13}C isotope selection using single-quantum coherence transfer with selective inversion on ligand proton peaks. The NMR experiment was triggered after an injection time of 440 ms and another 100 ms for sample stabilization. A 20 ms IBURP-shaped π pulse was included for performing selective inversion on specific ligand proton peaks followed by a pulsed field gradient G_z (44.9 G·cm^{-1}). Water suppression was then done by selectively exciting the water resonance by repeated 20 ms EBURP-shaped $\pi/2$ pulses, followed by randomized pulsed field gradients G_x, G_y, and G_z for dephasing. Acquisition on the ^{1}H channel for determining enhancements of ligand signals was performed after a 1° hard pulse. The shaped pulsed for selectively suppressing the water signal was repeated before the pulse sequence block including a single-quantum filter for ^{13}C chemical shift selection started. Delay τ was 1.79 ms. Selective inversion on ^{13}C resonances was applied by using a 7.65 ms shaped π pulse. The pulse had a Gaussian shape with a 1% truncation level and resulted in a frequency profile with a full width at half-maximum of 210 Hz. During this pulse, one complete cycle of WALTZ-16 decoupling was applied on the ^{1}H channel. The ^{1}H signal was acquired for $t_{1,\max} = 320$ ms, concomitant with the application of ^{13}C GARP decoupling with $\gamma B_1/(2\pi) = 3.3$ kHz. Pulse filed gradients $G_{z,1}$ (25.7 G·cm^{-1}, $G_{z,2}$ (-25.7 G·cm^{-1}), and $G_{z,3}$ (12.9 G·cm^{-1}) were included for coherence selection. Reprinted with permission from [184].

^{1}H signals were measured with a 1° excitation (point c in Figure 4.1).

The folic acid molecule was docked into the DHFR binding pocket using AutoDock 4.2.[77] The crystal structure of DHFR complexed with methotrexate (PDB entry 1RG7[177]) was used as the macromolecular structure after removing the ligand. A docking grid map with $60 \times 60 \times 60$ points was generated with a grid point spacing of 0.375 Å, and it was chosen to be centered to the center of the original ligand, MTX. Flexible ligand docking was performed to the rigid protein binding site using the Lamarckian genetic algorithm

(LGA); 250 ligand trial poses were generated by 250 independent runs with a maximum number of 2,500,000 energy evaluations and a maximum number of 27,000 generations. The receptor structure used in the docking along with the docked ligand poses was aligned with the structure of DHFR in PDB entry 1RE7[177] (DHFR–folic acid), so that the ligand poses can be compared with the ligand structure in the crystal structure with the RMSD values calculated.

Calculation of the NOE polarization transfer to generate simulated NOE peaks was performed for each ligand trial pose. The simulation was done by applying a complete relaxation and conformational exchange matrix analysis (CORCEMA).[97,99] The calculation was based on the two-state model of ligand–receptor reversible interactions.

$$\mathbf{E_f} + \mathbf{L_f} \underset{k_{\mathrm{off}}}{\overset{k_{\mathrm{on}}}{\rightleftharpoons}} \mathbf{E_b}\mathbf{L_b} \tag{4.1}$$

The time evolution of signal intensity for all species can be described by a set of differential equations. Both bound and free forms of 13 ligand protons and all protein protons, located within 6 Å of the ligand, were used in the calculation:

$$\frac{\mathrm{d}(\mathbf{S} - \mathbf{S}(\mathrm{e}))}{\mathrm{d}\,t} = (\mathbf{R} + \mathbf{K})(\mathbf{S} - \mathbf{S}(\mathrm{e})) \tag{4.2}$$

$\mathbf{S}$ and $\mathbf{S}(\mathrm{e})$ are vectors of time-dependent and equilibrium signal intensities for all spins of the form

$$\mathbf{S} = \begin{bmatrix} \mathbf{S}_{\mathrm{L}}^{\mathrm{f}} \\ \mathbf{S}_{\mathrm{E}}^{\mathrm{f}} \\ \mathbf{S}_{\mathrm{L}}^{\mathrm{b}} \\ \mathbf{S}_{\mathrm{E}}^{\mathrm{b}} \end{bmatrix} \tag{4.3}$$

$\mathbf{S}_{\mathrm{L}}^{\mathrm{f}}$, $\mathbf{S}_{\mathrm{E}}^{\mathrm{f}}$, $\mathbf{S}_{\mathrm{L}}^{\mathrm{b}}$ and $\mathbf{S}_{\mathrm{E}}^{\mathrm{b}}$ represent the free ligand, free protein, bound ligand, and bound protein, respectively. The kinetic matrix $\mathbf{K}$ is generated based on the two-state exchange between

free and bound species.

$$\mathbf{K} = \begin{bmatrix} k_{\mathrm{on}}[\mathrm{E_f}] \cdot \mathbf{I_L} & \mathbf{0} & -k_{\mathrm{off}} \cdot \mathbf{I_L} & \mathbf{0} \\ \mathbf{0} & k_{\mathrm{on}}[\mathrm{L_f}] \cdot \mathbf{I_E} & \mathbf{0} & -k_{\mathrm{off}} \cdot \mathbf{I_E} \\ -k_{\mathrm{on}}[\mathrm{E_f}] \cdot \mathbf{I_L} & \mathbf{0} & k_{\mathrm{off}} \cdot \mathbf{I_L} & \mathbf{0} \\ \mathbf{0} & -k_{\mathrm{on}}[\mathrm{L_f}] \cdot \mathbf{I_E} & \mathbf{0} & k_{\mathrm{off}} \cdot \mathbf{I_E} \end{bmatrix} \quad (4.4)$$

$\mathbf{I_L}$ and $\mathbf{I_E}$ are unit matrices with the dimension of the number of included ligand or protein spins. kon and koff are on and off rate constants for the protein-ligand complex. The relaxation matrix is given by

$$\mathbf{R} = \begin{bmatrix} \mathbf{R_L^f} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{R_E^f} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{R_L^b} & \boldsymbol{\sigma} \\ \mathbf{0} & \mathbf{0} & \boldsymbol{\sigma} & \mathbf{R_E^b} \end{bmatrix} \quad (4.5)$$

$\mathbf{R_L^f}$ and $\mathbf{R_L^b}$ represent the relaxation matrices for the free and bound ligand, respectively. $\mathbf{R_E^f}$ and $\mathbf{R_E^b}$ are relaxation rate matrices for the protein in the free and bound forms, and $\boldsymbol{\sigma}$ describes the cross-relaxation rate matrix between ligand and protein spins. The auto-relaxation rates of spin I_i, ρ_i, and the cross-relaxation rates between spins I_i and I_j, σ_{ij}, are included.

$$\sigma_{ij} = \frac{\mu_0^2 \gamma_{\mathrm{H}}^4 \hbar^2}{64\pi^2 r_{ij}^6} \Big(6J(2\omega) - J(0) \Big) \quad (4.6)$$

$$\rho_i = \sum_{j \neq i} \frac{\mu_0^2 \gamma_{\mathrm{H}}^4 \hbar^2}{64\pi^2 r_{ij}^6} \Big(6J(2\omega) + 3J(\omega) - J(0) \Big) + \rho_i^* \quad (4.7)$$

$$J(\omega) = \frac{2}{5} \frac{\tau_c}{(1 + \omega^2 \tau_c^2)} \quad (4.8)$$

ρ_i^* is the leakage rate caused by relaxation processes other than ^{1}H-^{1}H dipolar interaction between the spins in the matrix.

The effect of internal motions was considered for the methyl groups. The spectral

densities for intramethyl, methyl-methyl and methyl-other proton dipolar interactions were calculated by a modified Lipari-Szabo model-free approach.[202,203]

$$J(\omega) = \frac{2}{5} \left(\frac{S_i S_j \tau_c}{(1 + \omega^2 \tau_c^2)} + \frac{(1 - S_i S_j)\tau}{(1 + \omega^2 \tau^2)} \right) \tag{4.9}$$

$$\frac{1}{\tau} = \frac{1}{\tau_c} + \frac{1}{\tau_{\text{int}}} \tag{4.10}$$

τ_c and τ_{int} are the correlation times for the molecule and for the internal motions. S_i and S_j are order parameters for spin i and j.

The on rate, k_{on}, was estimated to be 10^8 M^{-1}s^{-1}, assuming that the association step is diffusion-controlled. The off rate, k_{off}, was calculated as 280 s^{-1}, with dissociation constant K_D measured to be 2.8 μM by isothermal titration calorimetry (ITC). The self-relaxation rates of free ligand protons were measured by inversion recovery experiments with a 2.5 mM folic acid solution in phosphate buffer (50 mM, pH 6.8). The cross-relaxation rates between free ligand protons were determined by fitting the NOE buildup curves after measuring ^{1}H–^{1}H NOESY spectra of the same folic acid solution with varying mixing times. The nonhyperpolarized NMR experiments and the ITC experiments were performed at 298 K. Other relaxation rates were calculated on the basis of dipolar interactions using intramolecular and intermolecular ^{1}H–^{1}H distances between protons i and j (r_{ij}) in each given ligand pose. All equivalent protons, including the methyl protons, were considered explicitly in the calculation. The overall rotational correlation time of DHFR, τ_c, was estimated using the Stokes law for a rigid sphere in water at 298 K.[24] The internal correlation time for the methyl groups, τ_{int}, was set to 0.005 ns.[204,205] The product of order parameters $S_i S_j$ was set to 0.25 for internal methyl interactions.[206] For external methyl interactions, the order parameter S_i was set to 0.39 for methyl protons[204] and 0.85 for other protons.[205]

Among all protein protons located within 6 Å of the ligand, which were included in the calculation, those protein spins located at the 1 Å thick outside shell were considered to have an additional leakage rate ρ^*(shell) for compensating relaxation caused by interaction with ^{1}H spins that are outside of the 6 Å sphere. For all 250 trial poses, R_1 relaxation rates for each protein spin were calculated by adding the contributions of auto- and cross-relaxation. ρ^*(shell) = 0.13 s^{-1} was estimated from the difference in average R_1 relaxation rates for the spins inside of the 5 Å sphere and those in the shell.

The calculated signal intensities for all species were defined as $S_i = [i]En_i$, where $[i]$ and En_i are the concentration and spin polarization enhancement factor of species i, respectively. At $t = 0$, $En_i(0) = 1$ for each protein spin. For the ligand spins, the $En_i(0)$ values were determined from the first scan of the DNP-NMR experiment. It was assumed that kinetic equilibrium was reached rapidly, so that the concentrations of the ligand and protein in the free and bound forms were kept as constant values and were calculated on the basis of the K_D value and total protein and ligand concentrations. Total protein and ligand concentrations for each experiment were determined by comparison to known standards, using ^{1}H NMR after completion of hyperpolarized experiments. In the first scan of the D-DNP NMR experiment, the acquisition was performed after a 1° hard pulse. Enhancements of all ligand signals were determined by comparing the D-DNP ligand NMR spectra ($t = 0.34$ s) with spectra measured at thermal equilibrium after each experiment (Table 4.1). Initial enhancement levels ($t = 0$) when the protein and ligand start to mix were calculated with the R_1 relaxation rates of each ligand ^{1}H signal measured by multiscan D-DNP NMR experiments with small flip angle acquisitions. For those peaks (H7, H2′/H6′, or H3′/H5′) that were inverted, their initial enhancements were estimated using the average relative ratio to other peaks (H7, H2′/H6′, or H3′/H5′) in non-inverted experiments. Enhancement levels of H9, which is close to the water resonance and hence cannot be observed in the D-DNP NMR spectra, were set by averaging enhancements of

other CH_2 protons, $H\beta$, $H\beta'$, and $H\gamma$.

Exp.	Enhancement			Factor					Concentration (mM)	
Number	H7	H8	H2′/H6′	H3′/H5′	Hα	Hγ	Hβ	Hβ′	Folic acid	DHFR
1	766	233	941	870	541	311	230	640	2.24	0.236
2	49*	292	1323	1420	851	450	313	1056	2.02	0.218
3	1248	124	117*	1039	946	472	338	804	2.08	0.233
4	1157	305	1113	385*	712	381	288	790	1.79	0.216
5	1488	571	1670	1550	890	511	394	1608	1.52	0.252
6	229*	268	1431	1636	971	614	382	1259	2.33	0.228
7	1378	172	309*	1140	1024	566	393	1081	1.99	0.242
8	2043	652	1764	233*	1153	565	383	992	2.07	0.220
9	1033	231	1064	1002	619	339	257	636	2.08	0.269
10	18*	527	1673	1789	1007	508	353	1513	1.80	0.220
11	1368	117	−221*	1248	1009	548	455	1353	1.50	0.257
12	1425	339	1241	129*	849	535	346	997	2.44	0.238

Table 4.1: Signal enhancements of 8 ligand peaks and sample concentrations for 12 DNP-NMR experiments. * indicates the peak that was selectively inverted. The enhancement was determined by comparing the ligand ^1H signals measured at thermal polarization with hyperpolarized ^1H signals measured with a 1° flip angle pulse before each scan with ^{13}C selection. Here, the 12 experiments correspond to the 12 ^1H spectra are shown in Figure 4.2. Reprinted with permission from [184].

The time evolution of signal intensities for all spins was simulated by numerically solving the matrix differential equation using the ode45 solver in MATLAB. The occurrence of ligand–protein mixing was set as time zero. Selective inversion on individual ligand proton peaks occurred at 0.34 s, which was estimated as 220 ms corresponding to half of the injection time, 100 ms of stabilization time, and 20 ms of duration for the selective inversion pulse. In the simulation, as a result, the signal of the selected inverted peak was multiplied by an additional factor finv. Because the inversion is not perfect, finv is not −1. For the nine experiments with an inversion pulse, f_{inv} values were optimized until

the measured enhancement level of the selective ligand was fit best. Simulated intensities for each NOE peak in the D-DNP NMR experiment were obtained at 1.47 s, where the acquisition was performed. The effect of ^{13}C single-quantum coherence selection was considered by multiplying the simulated intensities with an additional coefficient determined by the excitation profile of the 7650 μs Gaussian pulse.

NOE peaks in the protein spectra measured as the second scan in the 12 D-DNP experiments were fitted with Lorentzian line shapes, leading to 28 individual peaks (Figure 4.2). To compare the integrals of these measured NOE peaks with the simulated results, a normalization factor needs to be taken into consideration as $S_{\mathrm{Sim},i} N = S_{\mathrm{Exp},i}$. The same pulse sequence with the ^{13}C selection at 11.4 ppm was applied after the D-DNP experiments. The normalization factor was calculated by

$$N = \frac{S_{\mathrm{Exp},i}}{c_{\mathrm{methyl}} \cdot 1} = \frac{S_{\mathrm{Exp},i}}{3 c_{\mathrm{protein}} \cdot 1} \tag{4.11}$$

For each trial pose, the simulated signals of the 28 NOE signals were compared with experimental integrals by calculating the NOE Score function, defined in the text.

The leakage rate ρ^* was adjusted with a step size of 0.05 s^{-1} to obtain the lowest average NOE Score for all 250 poses (Figure 4.3).

4.3   Results and Discussion

To observe polarization transfer through intermolecular NOEs, two one-dimensional (1D) ^{1}H NMR spectra were acquired after mixing of DHFR protein with the hyperpolarized ligand folic acid. The first scan was for obtaining a spectrum of the hyperpolarized folic acid immediately after injection of the sample into the NMR spectrometer, acquired with a 1° hard pulse for excitation (Figure 4.4a, top left). This spectrum can be used to calibrate signal intensities, which depend on the magnitude of initial hyperpolarization of the ligand. From the spectrum in Figure 4.4a, significant
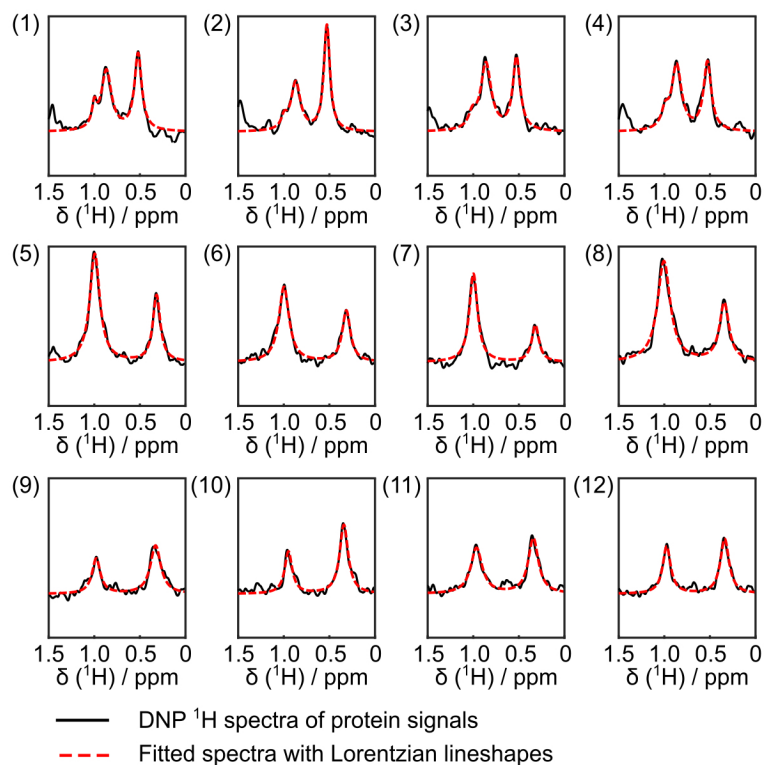
Figure 4.2: [^1]H NMR spectra of DHFR measured after admixing DNP hyperpolarized folic acid (black lines). Peaks in the spectra were fitted with Lorentzian line shapes (red dotted lines). Each of the 12 panels shows data from a separate experiment measured with a ^{13}C filter at 17.5 ppm (1-4), 15.5 ppm (5-8) and 13.5 ppm (9-12). Selective inversion pulses were applied at the beginning of the pulse sequence at the resonance of folic acid H7 (2, 6, 10), H2′/H6′ (3, 7, 11) and H3′/H5′ (4, 8, 12), respectively. Spectra 1, 5 and 9 were measured without any inversion pulse. Spectra are zoomed in to the methyl group region. Reprinted with permission from [184].

enhancement of ligand signals due to hyperpolarization was observed (top left) by comparison with a reference spectrum acquired after the decay of hyperpolarization (bottom left). Average enhancement factors among 12 experiments were determined as 1320 ± 350 for H7, 1360 ± 300 for H2′/H6′, and 1300 ± 320 for H3′/H5′. With a rapid sample injection system,[151] which minimizes the transfer time, the total time starting from the dissolution to this NMR scan was 1.6 s. This timing mitigated polarization loss during the injection. It enabled us to obtain a signal enhancement of >1000 for the five protons
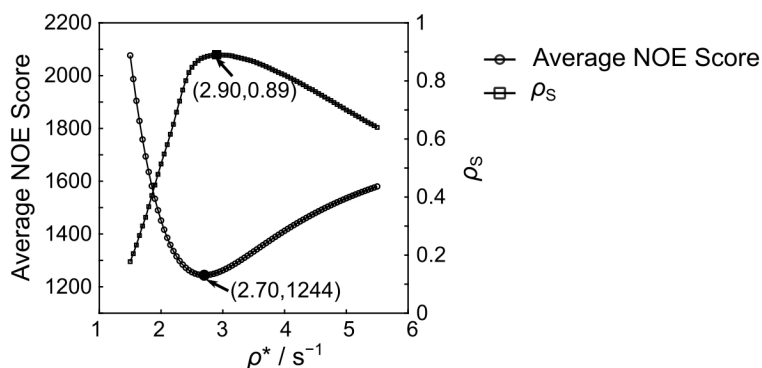
Figure 4.3: Plot of average NOE Scores among 250 poses, and Spearman's correlation coefficient (ρ_S) between the NOE Score and the RMSD value for 5 protons, as a function of leakage rate (ρ^*) chosen for the calculation. The lowest average NOE Score of 1244 was attained with a leakage rate of 2.7 s^{-1}. The largest correlation coefficient of 0.89 was reached with a leakage rate of 2.9 s^{-1}. Reprinted with permission from [184].

employed in the subsequent analysis, all of which have high-field T_1 relaxation times in the range of 0.8–2 s. Signals of other peaks with shorter relaxation times between 0.4 and 0.6 s were also observed but were not included in the analysis.

The second scan was measured after a time delay of 1.47 s, allowing for the transfer of hyperpolarization to the protein. To observe individual enhanced protein peaks, the pulse sequence for this scan contained a single-quantum filter with ^{13}C chemical shift selection. The resulting ^{1}H spectrum in Figure 4.4a (top right) generates only signals of protein ^{1}H spins that are correlated to ^{13}C spins within a specific range of chemical shifts, and only signals of proton spins located close enough to the hyperpolarized ligand spins are efficiently enhanced. Compared with the reference spectrum measured with the same ^{13}C filter at thermal polarization (bottom right), the D-DNP NMR spectrum shows two resolved peaks in the methyl group region strongly enhanced. The enhancement levels of >1000 for the ligand ^{1}H spins allow the observation of the enhanced protein signals in the single-scan NMR spectrum. Without hyperpolarization, 10^6 scans would be needed to obtain data of the same quality using a 400 MHz NMR spectrometer, which with a recycle

72
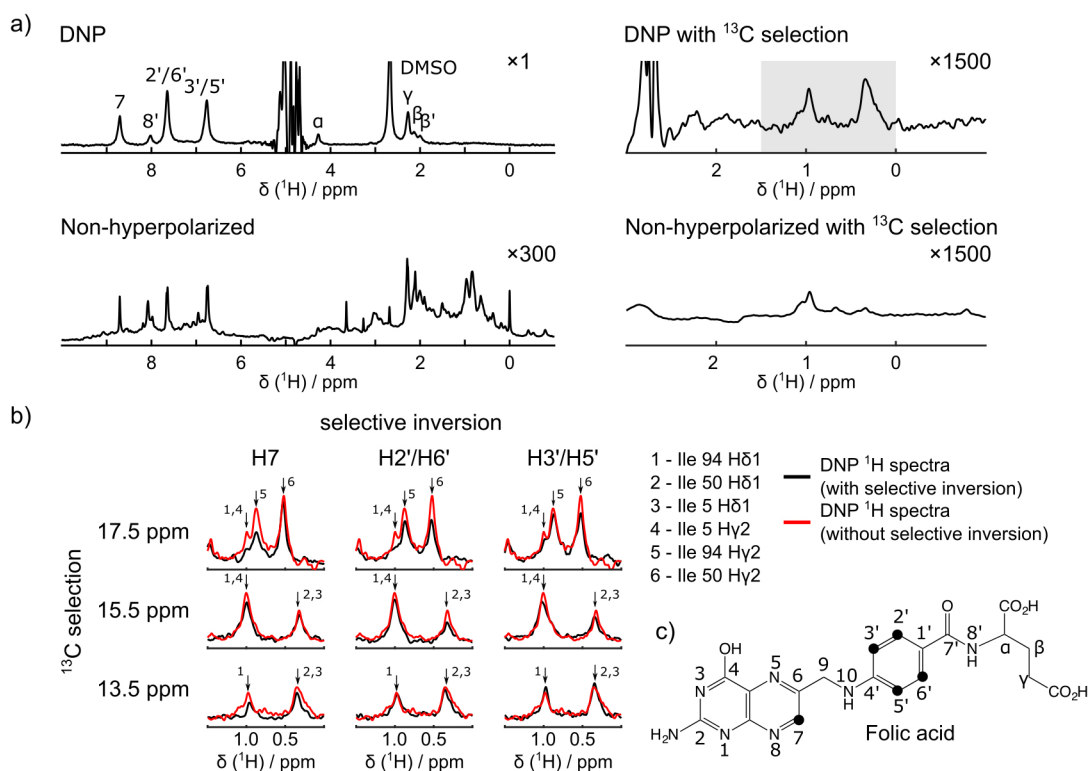
Figure 4.4: a) ^{1}H spectra of folic acid and DHFR: top left, hyperpolarized ligand with protein measured with 1° excitation; top right, hyperpolarized ligand with protein with ^{13}C selection at 13.5 ppm; bottom left, ligand with protein at thermal polarization, averaged by 32 scans; bottom right, ligand with protein at thermal polarization with ^{13}C selection at 13.5 ppm, averaged by 256 scans. b) D-DNP NMR spectra (black) with a ^{13}C selection filter at 13.5, 15.5, and 17.5 ppm, with a close-up on the DHFR methyl region. Selective inversions were applied at folic acid protons H7, H2′/H6′, and H3′/H5′ for panels 1–3, respectively. Spectra with no selective inversion (red) are superimposed in all panels. The spectra displayed were normalized by the enhancement factors of the ligand and the concentration of the protein. These two parameters are listed for each experiment in Table 4.1. They exhibit a variation of approximately $\pm 25\%$ for the enhancement factor and $\pm 7\%$ for the protein concentration. c) Structure of folic acid. Protons that were selectively inverted are marked. Reprinted with permission from [184].

delay of 1 s would require 11 days.

Data sets, as shown in Figure 4.4b, were measured with the selection at three different ^{13}C chemical shifts. In addition, an inversion pulse on a specific ^{1}H ligand resonance was

73

applied in some of the experiments. The reduction in the intensity of a protein signal due to inversion of a specific ligand signal allows one to distinguish the origins of the transferred polarization. In total, 12 D-DNP NMR ^{1}H spectra were acquired, and 28 NOE peaks were observed. The 28 NOE peaks were assigned to six methyl groups from three amino acids, which are within the excitation region of the Gaussian-shaped selective pulse on the ^{13}C channel. Via comparison of the spectra acquired with and without the selective inversion, it becomes evident which protein signals receive polarization transfer from each individual ligand spin. For example, the NOE signals for Ile 94 Hδ1 (peak 1), Ile 5 Hγ2 (peak 4), and Ile 94 Hγ2 (peak 5) show the greatest signal loss when H7 is inverted, while a smaller change is observed with inversion of H3$'$/H5$'$. For the Ile 50 Hγ2 peak (peak 6), greater signal reduction can be seen with H2$'$/H6$'$ inversion, while there is less signal loss for H3$'$/H5$'$ inversion and almost no change for H7 inversion. This signal intensity variation is consistent with the NOE cross-peaks observed in a conventional filtered three-dimensional (3D) NOESY spectrum described in ref[148], in which larger NOE cross-peaks are present for Ile 94 Hδ1, Ile 5 Hγ2, and Ile 94 Hγ2 in the ligand H7 plane, and intense NOE cross-peaks of Ile 50 Hγ2 are seen in the ligand H2$'$/H6$'$ plane. This multidimensional NOESY experiment was conducted on a 500 MHz NMR spectrometer with a cryoprobe for 3 days. The observed decrease in the intensity of the signal in the hyperpolarized spectra upon inversion of a ligand signal can be used as an indicator of the proximity of the respective protons. Although all ligand protons are hyperpolarized simultaneously, the selective inversion at the beginning of the experiment makes it possible to distinguish contributions from different ligand protons.

A limited number of NOE cross-peaks, such as those obtained in Figure 4.4, are alone not sufficient to calculate the molecular structure. Combining similar numbers of experimental restraints with computational docking methods, however, has been proven to be successful in determining the binding epitope structure.[94,95] Here, 250 poses of folic

acid with different ligand conformations and orientations were generated using the docking program AutoDock.[77] For docking, the X-ray crystal structure of the DHFR–methotrexate (MTX) complex determined by Sawaya *et al.* [Protein Data Bank (PDB) entry 1RG7[177]] was used. A structure with a ligand different from the folic acid used in the experiments was purposefully chosen. The protein structure remained static during the docking procedure, while the ligand orientation and bond angles were changed. After completion of the docking calculation, the polarization transfer process in the D-DNP experiment was simulated for each pose by a complete relaxation and conformational exchange matrix analysis (CORCEMA) as described by Moseley *et al.*[97,99] Thirteen folic acid proton spins and all DHFR proton spins that are within 6 Å of the ligand were included in the calculation. The spins in free and bound forms of protein and ligand were treated separately. A leakage rate, ρ^*, was added to the autorelaxation term for each bound ligand spin and all the protein spins, to compensate for relaxation processes other than the ^{1}H–^{1}H dipolar interaction, such as dipolar interaction with ^{13}C, chemical shift anisotropy relaxation, paramagnetic relaxation due to the radicals, and exchange of amide protons with the solvent. These simulations resulted in the predicted time dependence of signal intensities for each of the ligand and protein peaks, in the presence or absence of selective inversion applied on ligand spins. The simulation results for two of the calculated poses are shown in Figure 4.5. In the simulations with no inversion pulse, it can be seen that the methyl proton signals of the protein build up immediately after mixing and eventually drop toward the equilibrium. Inverting the polarization of ligand ^{1}H resonances generates distinct effects on different DHFR methyl peaks as evidenced by these calculations. For the pose in Figure 4.5a, inversion of the ligand H7 proton leads to a marked decrease in the intensity of the signal for Ile94 Hγ2, Ile94 Hδ1, and Ile5 Hγ2, while inversion of ligand H2′/H6′ or H3′/H5′ has minor effects on those three methyl protons. The different response of the same protein peaks with inversion of individual ligand protons reflects
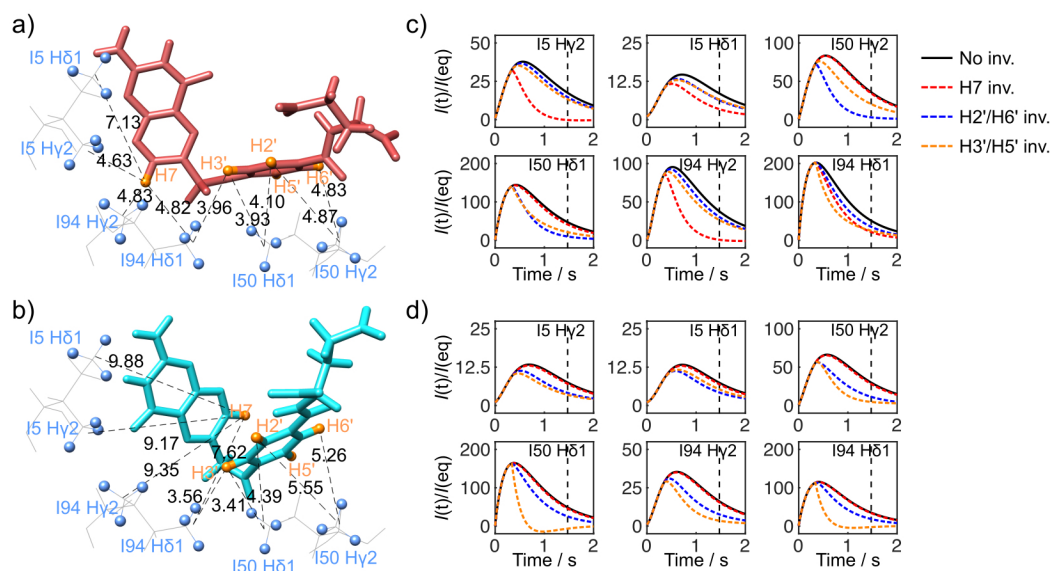
Figure 4.5: Simulation of polarization transfer for two ligand trial poses. a) and b) Structures of these two poses along with three nearby amino acids in DHFR that were observed in the D-DNP NMR experiments. Selected intermolecular ^{1}H–^{1}H distances between folic acid and DHFR are shown with dashed lines and indicated with numbers in units of angstroms. The protein structure is from PDB entry 1RG7.[177] c) and d) Evolution of simulated enhancement factors for the signals of six protein methyl protons based on ligand poses in panels a and b, respectively. The solid line represents the experiments with no selective inversion. Selective inversions at H7 (red dotted), H2′/H6′ (blue dotted), and H3′/H5′ (orange dotted) were applied at 0.34 s. The vertical black dotted line indicates the time ($t = 1.47$ s) when the acquisition started. Reprinted with permission from [184].

corresponding structural information. In this particular pose, Ile94 and Ile5 are closer to H7 than the other protons inverted (H7-Ile5 Hγ2, 4.63 Å; H7-Ile94 Hδ1, 4.82 Å; H7-Ile94 Hγ2, 4.83 Å).

Upon comparison of the simulated curves of the protein signal intensity versus time for the poses in panels a and b of Figure 4.5, distinct differences can be seen. For the latter pose, simulated signal intensities of Ile94 Hγ2, Ile94 Hδ1, and Ile5 Hγ2 build up more slowly, and inverting the polarization of ligand H7 has no obvious effects on these signals. In this pose, the pteridine ring of the ligand is flipped compared to the pose in

Figure 4.5a. The H7 atom is then located farther from I5 and I94 (H7-Ile5 Hγ2, 9.17 Å; H7-Ile94 Hδ1, 7.62 Å; H7-Ile94 Hγ2, 9.35 Å). The calculated results for these two ligand poses illustrate that the simulation of NOE peaks based on polarization transfer reflects the structural properties of the ligand–protein complex in each pose. Therefore, the simulation can be used to establish a relationship between the structure and the experimental results.

To allow for a comparison with the experimental D-DNP NMR data, simulated NOE peak signals were taken from the curves at the specific time point (1.47 s), which corresponds to the point of acquisition of the spectra. These signals were further multiplied by a factor describing the frequency offset dependence of the pulse sequence containing the selective ^{13}C pulse. As a result, expected signal integrals for the 28 peaks in the 12 spectra were obtained.

For each trial pose, the simulated signals were compared with experimental signal integrals by calculating an NOE Score function, defined in arbitrary units as

$$\text{NOE Score} = \sqrt{\sum_{i=1}^{28} (S_{\text{Exp},i} - S_{\text{Sim},i})^2} \qquad (4.12)$$

Here, the sum extends over all 28 observed NOE signals. The calculated NOE scores range from 635 to 1966 for the 250 poses from the docking calculations. A histogram of the obtained scores is shown in Figure 4.6, which appears as a nearly normal distribution with a mean value of 1244 and without outliers. Of all the ligand poses, only seven poses lie in the first bin with the lowest NOE Score values of 600–730, where the best trial pose should be located.

The use of the observed transfer of hyperpolarization from ligand to protein for the purpose of selecting poses as described above can be validated by the correlation of the NOE Scores with structural RMSD values between the respective ligand poses and a reference structure. In Figure 4.7a, Spearman's correlation coefficient (ρ_S) to a known
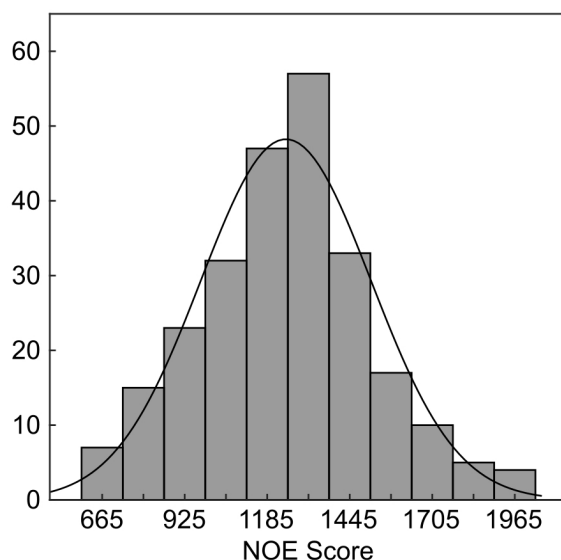
Figure 4.6: Histogram of the calculated NOE scores for 250 trial ligand poses. The data were also fitted with a normal distribution ($\mu = 1244$; $\sigma = 269$). Reprinted with permission from [184].

crystal structure of the protein with the same ligand is shown. The correlation coefficient between the NOE Score values and ligand pose RMSD is 0.88 when considering only the five protons inverted in the D-DNP experiments. Upon calculation of the whole ligand structure, the correlation coefficient drops to 0.59. With no protons inverted for the glutamate moiety, there is no experimental selectivity for this part of the ligand. To exclude the influence from this part of the molecule in the correlation, the analysis was also performed considering the ligand structure without the glutamate part, giving a Spearman's correlation coefficient of 0.77. In the crystal structure of the folic acid–DHFR complex, there are also no short distances of ≤ 5 Å between the protons in the glutamate moiety (Hα, Hβ, and Hγ) and the six observed protein methyl protons. The correlations including the experimental data are better than the same correlation that is obtained from the AutoDock Score (binding energy) (Figure 4.7b), even when including the parts of the structure in which no experimental selectivity exists. This comparison demonstrates that the NOE

78
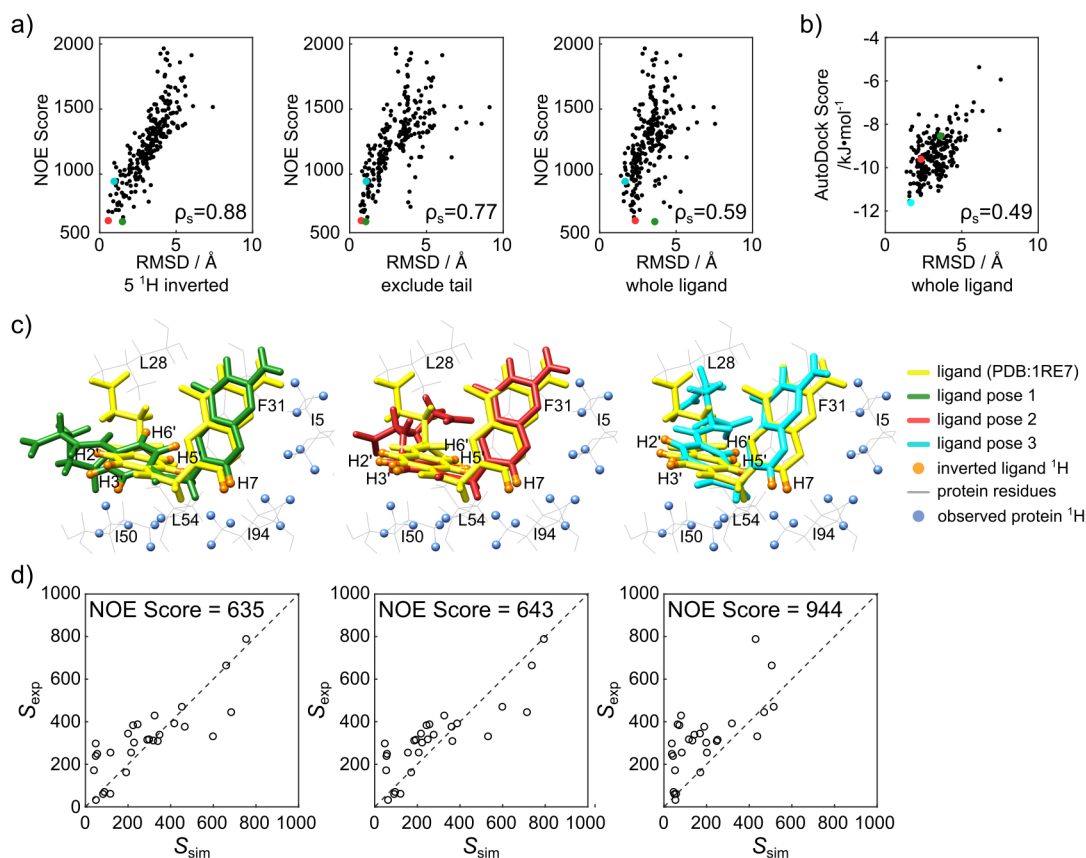
Figure 4.7: a) Correlation plots of NOE Score vs RMSD. RMSD values in three panels were calculated, including only five protons inverted in the D-DNP NMR experiments (left), the whole ligand structure excluding the glutamate moiety (middle), and the entire ligand structure (right). b) Correlation plot of binding energy (AutoDock Score) vs RMSD (whole ligand structure). c) Comparison of three ligand poses with the crystal structure of folic acid complexed with DHFR (PDB entry 1RE7).[177] The three poses are also highlighted in panels a and b. d) Signal integrals of 28 NOE peaks (Figure 4.4b) obtained from 12 D-DNP NMR experiments vs signal of simulated peaks plotted for the three ligand poses. Reprinted with permission from [184].

Score ranks the trial poses that are close to the experimentally determined crystal structure highly and rejects false poses generated by the docking program.

Structures of three selected ligand poses are shown in Figure 4.7c, and for each of the three poses, experimental signal integrals for the observed methyl group signals from all experiments are plotted against the simulated signals in Figure 4.7d. Compared to poses 1

and 2, a larger deviation from the diagonal of the graph is observed for pose 3, suggesting an expanding disagreement between the experimental and simulated results, which is consistent with its larger NOE Score value. Pose 1 is the structure that gives the lowest NOE Score of 635. The RMSD value comparing the selected five protons in this pose to those in the crystal structure is 1.49 Å, ranking 24th among all poses. Compared with the ligand pose in the crystal structure, the pteroyl moiety overall shows a good agreement, except for the not perfectly aligned benzene ring. Its low NOE Score might be explained by the indistinguishable aromatic protons in the free ligand, H2′ and H6′, and H3′ and H5′. Pose 2, which has the second lowest NOE Score of 643, also has the lowest RMSD value of 0.56 Å when calculating only the five selected ligand protons, and the second lowest RMSD value of 0.75 Å when considering the whole pteroyl group. In Figure 4.7c, pose 2 shows an excellent agreement with the crystal structure in particular for the pteroyl part of the ligand, in which all five protons selectively inverted belong. The glutamate tails for both poses 1 and 2 agree less closely with the crystal structure, for the reason stated above. The structure of the pose giving the lowest Autodock Score (Figure 4.7b) is also shown in Figure 4.7c as pose 3. Its pteroyl part does not align well with the crystal structure, giving a relatively higher NOE Score of 944. This comparison demonstrates that the NOE Score function can refine the docked poses. Specifically, the NOE Score function is effective for the part of the ligand with selectively addressed protons, while calculation of the binding energy by AutoDock is based on the entire ligand structure.

The previously defined leakage rate, $\rho*$, representing relaxation mechanisms other than proton–proton dipolar relaxation, has a variety of possible sources, and hence, it is impossible to measure or determine it directly. Here, it was adjusted until the minimum average NOE Score value for all poses was reached (Figure 4.3). The lowest average score was obtained for a leakage factor of 2.7 s^{-1}. This observed value is somewhat larger than typical values of 0.1–2 s^{-1} used in non-hyperpolarized samples.[207–209] The larger leakage

τ_{c} (ns)	5	6	7	8	9	10	11
$\rho^*(\mathrm{s}^{-1})$	2.35	2.50	2.60	2.70	2.75	2.85	2.90
$\rho_{\mathrm{S}}(5\ {}^{1}\mathrm{H}$ inverted)	0.86	0.87	0.88	0.88	0.88	0.88	0.88
ρ_{S}(exclude tail)	0.75	0.76	0.76	0.77	0.77	0.77	0.77
ρ_{S}(whole ligand)	0.57	0.58	0.59	0.59	0.60	0.60	0.60
Ranking (pose with lowest RMSD)	3	2	2	2	2	2	4

Table 4.2: Effects of the protein correlation time on the calculation results, including the optimized leakage rate (ρ^*); Spearman's correlation coefficients (ρ_{S}); as well as the ranking of the pose with the lowest RMSD value between the ligand pose and the ligand in the reference crystal structure, among 250 trial poses. Reprinted with permission from [184].

factor can be explained by factors including additional paramagnetic relaxation due to radicals and an increased relaxation rate of amide protons and other exchangeable protons with solvent suppression pulses applied in the experiments. The effect of the leakage factor on calculated Spearman's correlation factor, between the NOE Score and the RMSD value of the five inverted protons, is also plotted in Figure 4.3. The best correlation is reached with a leakage factor of 2.9 s^{-1}, which is close to the value of 2.7 s^{-1} determined by lowest average NOE Scores. The calculation of the expected NOE requires the correlation time of the protein as a parameter. For the data shown in the figures, a rotational correlation time τ_{c} of 8 ns was assumed. Because the NOE depends on the correlation time, the effects of calculating with different τ_{c} values between 5 and 11 ns are evaluated in Table 4.2. Differences in resulting Spearman's correlation coefficients are <0.04. The ranking of the ligand pose with the lowest RMSD to the crystal structure in all calculations is between 2 and 4 of 250, indicating that the value of the correlation time over this range does not play a major role.

Here, the receptor structure was chosen as DHFR in complex with another ligand, MTX, and kept rigid during the entire docking process. However, in many systems, conformational changes would be expected upon binding of different ligand molecules,

and even small motions of the side chains would have non-negligible effects on the docking results. Although the ligands MTX and folic acid are structurally similar, it is still reasonable to consider differences in receptor conformation with the binding of the two ligands. An alternative approach that may further refine the results by avoiding any bias caused by energy-based calculations would be the use of experimentally obtained structural restraints to guide the docking process. Recently, a method based on a purely data-driven conformational search with intraligand and intermolecular NOE restraints measured by traditional 3D NOESY was developed by Orts et al.[210,211] This approach may be viable for use with hyperpolarization when a larger number of structural constraints become available. To achieve a two-sided selection on both the ligand ^{1}H and protein ^{13}C dimensions, multiple 1D D-DNP NMR measurements were performed here. The efficiency of this process may be improved, for example, by using multichannel detection,[212] Hadamard spectroscopy,[136] or ultrafast two-dimensional NMR,[127] each of which is compatible with D-DNP NMR.

4.4   Conclusions

In summary, the combination of computational docking and hyperpolarized NMR measurements provides a method for obtaining structural information about a protein–ligand complex. Using a two-sided coherence selection, the required distance information can be obtained from a series of D-DNP NMR experiments. Because of the use of DNP, fast acquisition within a few seconds can be achieved, which make this method useful for some samples that are not easily purified or stable enough for conventional multidimensional NMR experiments.

5.   DETERMINATION OF PROTEIN-LIGAND BINDING MODES USING FAST

MULTI-DIMENSIONAL NMR WITH HYPERPOLARIZATION

5.1   Introduction

Interactions of small molecules and proteins are the key components involved in many essential biological processes, including enzymatic catalysis, cellular signaling pathways, and regulation of protein functions. A comprehensive understanding of the structure and dynamics pertaining to these interactions is required to not only reveal the molecular mechanisms in biological processes, but also to provide the knowledge to develop synthetic small-molecule drugs that can alter the functions of the target proteins. Nuclear magnetic resonance (NMR) spectroscopy enables the non-invasive investigation of the intermolecular interactions in aqueous solution,[142,213] and recent advances in live cell NMR[214] has allowed to monitor molecular interactions directly in cellular environments.[54,215–217]

Observing magnetization transfer through the nuclear Overhauser effect (NOE) is a powerful NMR-based method to probe molecular structure, and has been widely applied for studying the water-protein[218,219] or ligand-protein interactions.[32] Dissolution dynamic nuclear polarization (D-DNP) is a hyperpolarization technique, which can boost the sensitivity of NMR signals by several orders of magnitude.[113] Using the D-DNP to generate hyperpolarized spins serving as the NOE source, the efficiency of the NOESY measurement can also be significantly improved. As a versatile technique, DNP is capable of hyperpolarizing a wide range of small molecules, including water and typical ligand molecules. In addition to accelerating an NMR experiment owing to the sensitivity gain, hyperpolarization of small molecules also provides a natural contrast for exclusively observing signals originating from the small molecule of interest.

Several applications have been demonstrated, where ^{1}H spins of water or other small molecules are hyperpolarized by dissolution DNP, and subsequently participate in polarization transfer to the protein or other small molecules through intermolecular NOEs. Interactions between water and protein can be directly studied by observing polarization transfer from hyperpolarized water to the protein. Hyperpolarization can be transferred to amide protons on the protein through proton exchange, and further spread within the protein through intramolecular NOE.[220] By the same mechanisms, water molecules can serve as an agent to introduce hyperpolarization to the protein spins for subsequent use. This method has been applied to obtain high-resolution NMR spectra of intrinsically disordered proteins (IDPs),[131] as well as folded proteins, using fast two-dimensional (2D) measurements.[132] Further, our group has recently demonstrated the application of this method to measure protein signals during the folding process, which can provide insight into the structure and dynamics in protein folding.[133]

Interactions between water and protein can be directly studied by observing polarization transfer from hyperpolarized water molecules to the protein. From the quantitative modeling of the time-dependent transferred signals on the protein amide and aliphatic protons, it has been revealed that water hyperpolarization can be transferred to amide protons on the protein through proton exchanges and further spread within the protein through intramolecular NOE to nearby protons including aliphatic protons.[220] Owing to the rapid polarization transfer from water to protein through these mechanisms, the water molecules can also serve as an agent to introduce hyperpolarization to the protein spins. This method has been applied to obtain high-resolution NMR spectra of intrinsically disordered proteins (IDPs)[131] and the folded protein using fast 2D measurements.[166] Further, our group has recently demonstrated the application of this method to measure the protein signals during the folding process. The difference of residue-specific enhancements gained on the protein between under the conditions of

folded and refolding provides an insight into the structure and dynamics during protein folding.[133]

When a ligand is present in the protein solution, hyperpolarization from water can be transferred to the ligand, and used for the detection of binding.[124] In addition, D-DNP can directly generate hyperpolarization on ligand spins. Protein mediated transfer of polarization from one hyperpolarized ligand to another, competing ligand, has been detected in the presence of the protein target, providing structural information about the ligand binding epitope.[137] Polarization transferred from a hyperpolarized ligand can also be observed directly as selectively enhanced protein signals, revealing structural information related to ligand-protein interaction.[123,148]

Knowledge of the protein-ligand complex structure provides essential information to guide the ligand optimization process in structure-based drug design (SBDD). Computational docking is a rapid and inexpensive method for predicting the orientation and conformation of the ligand when binding to the target protein structure. The integration of NMR restraints into docking calculations can further improve the prediction reliability and has emerged as a popular means in drug discovery.[20,21] A common strategy consists of the generation of a set of possible ligand binding modes *in silico*, which are then ranked using the experimental NMR data. Even sparse contact information, based on intermolecular NOEs between ligand and protein,[193,194] intermolecular NOEs between competitively binding ligands,[103,104] or intra-ligand NOEs,[98] are sufficient for this purpose. In a further refinement of this method, data-driven docking directly integrates experimental data into the docking algorithm. One such approach is the high ambiguity driven biomolecular docking method,[85,221] where experimental data can be converted into ambiguous interaction restraints, which are used to guide the docking processes.[94,222] Recently, a highly automated approach using intra-ligand NOEs and ambiguous intermolecular NOEs without protein chemical shift assignments, was shown

to enable the determination of the ligand binding mode in a receptor binding pocket.[210,211]

Recently, we developed a method for the determination of ligand binding epitope structures by a combination of molecular docking and intermolecular NOEs obtained from a set of 1D hyperpolarized ^{1}H NMR spectra.[184] Here, we demonstrate that the efficiency of the NOE measurement is significantly improved by multi-dimensional NMR spectroscopy, which includes a 2D correlation of protein spins combined with Hadamard encoding of ligand signals in a third dimension. These spectra contain a similar level of information as would be available from conventional NMR spectra with longer acquisition time. The spectra are used to calculate the binding epitope structure of a ligand for the dihydrofolate reductase (DHFR) protein.

5.2 Experimental Section

Uniformly ^{13}C enriched DHFR was produced by expression from the pET-Duet-1 plasmid in *E. Coli* BL21(DE3) cells, cultured in M9 minimum medium containing 3 g/L [U-^{13}C]glucose. The His-tagged protein was purified by a Ni^{2+}-NTA column followed by gel filtration chromatography as previously described.[148] Purified DHFR was prepared as a 5 mM stock solution in 50 mM phosphate buffer (pH 6.8) containing 1 mM EDTA, 1 mM DTT, 50 mM KCl. The concentration was determined by measuring UV-Vis absorption at 280 nm.

For DNP hyperpolarization, samples of 350 mM folic acid (Spectrum Chemicals, Gardena, CA) were prepared in D$_2$O/DMSO-d$_6$ (v/v 7:3) containing 15 mM 4-hydroxy-2,2,6,6-tetramethylpiperidine-1-oxyl (TEMPOL; Sigma-Aldrich, St. Louis, MO). 30 µL aliquots of the mixture were hyperpolarized on ^{1}H spins in a HyperSense DNP polarizer (Oxford Instruments, Abingdon, UK) under 100 mW microwave irradiation at a frequency of 94.005 GHz, and a temperature of 1.4 K. After 20 min, a frozen DNP sample was dissolved by 4 mL phosphate buffer (50 mM, pH 6.8) preheated until reaching a vapor
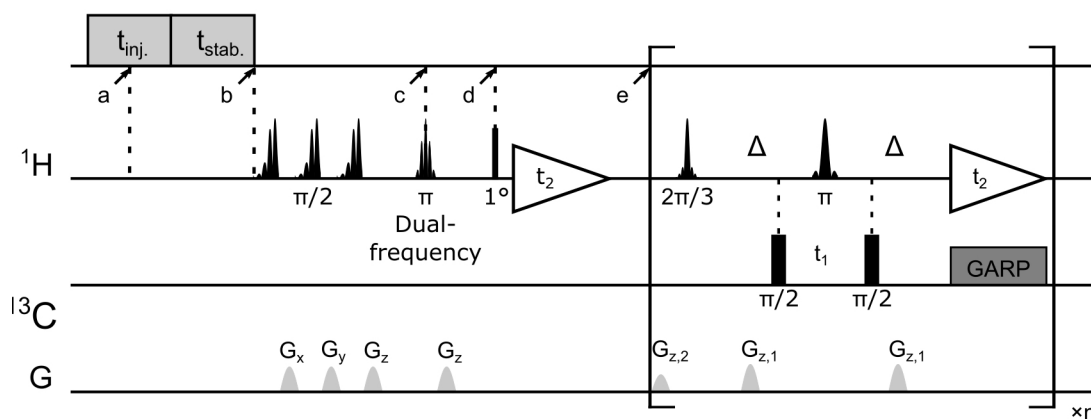
Figure 5.1: Pulse sequence for the [^{1}H-^{13}C]-SOFAST-HMQC experiment with ligand resonances encoded by a Hadamard scheme. After the injection ($t_{inj.}$ = 375 ms) and sample stabilization ($t_{stab.}$ = 385 ms), the NMR experiment was triggered. Repeated 20 ms EBURP $\pi/2$ pulses followed by pulsed-filed gradients G_x (44.8 G·cm^{-1}), G_y (38.6 G·cm^{-1}), G_z (33.5 G·cm^{-1}) were applied for water suppression. The two $\pi/2$ pulses on the ^{13}C channel were applied with $\gamma B_1/2\pi$ = 11.4 kHz. A 12.7 ms dual Gaussian shaped pulse with flip angle of π was applied simultaneously on two ligand ^{1}H resonances followed by a pulsed filed gradient G_z (47.7 G·cm^{-1}). The first ^{1}H scan was acquired after a hard pulse with a small flip angle (1°) for enhancement determination of hyperpolarized ligand signals. In the following [^{1}H-^{13}C]-SOFAST-HMQC pulse sequence, a 2.8 ms PC9 shaped pulse (flip angle $2\pi/3$, ±2 ppm bandwidth) and a 1.9 ms RSNOB shaped π-pulse were centered at 0 ppm for selective methyl proton excitation and refocusing. The coherence transfer delay was set to $1/(2J_{CH})$ as Δ = 3.5 ms. A ^{13}C GARP decoupling sequence ($\gamma B_1/2\pi$ = 3.1 kHz) was applied during the acquisition. Pulsed field gradients were applied with $G_{z,1}$ (7.5 G·cm^{-1}) and $G_{z,2}$ (4.8 G·cm^{-1}). A total of 40×1200 complex points were acquired for the ^{13}C and ^{1}H dimensions, with $t_{1,max}$ = 8.2 ms and $t_{2,max}$ = 100 ms, respectively.

pressure of 6 bar and then injected into a 5 mm NMR tube. The NMR tube was pre-installed in the 400 MHz NMR magnet (Bruker Biospin, Billerica, MA). In the NMR tube, a total of 25 μL of 5 mM protein solution was preloaded, which mixed with the hyperpolarized ligand during injection. The injection was achieved by applying N$_2$ gas at a forward pressure of 1.81 MPa against a back pressure of 1.03 Mpa. The middle of this injection period was considered as the point when the mixing started (a in Figure5.1 , $t = 0$ s). NMR measurements were programmed to automatically start after an additional period

for sample stabilization (b in Figure 5.1, $t = 0.57$ s). The final sample concentrations were determined individually for all the samples, resulting in average values of 4.9±0.4 mM and 0.34±0.02 mM for the ligand and protein, respectively. All hyperpolarized NMR spectra were acquired in a 5 mm triple-resonance inverse detection (TXI) probe (Bruker Biospin, Billerica, MA) at a temperature of 303 K. The NMR pulse sequence consists of a selective dual-frequency inversion pulse for ligand ^{1}H encoding (c in Figure 5.1, $t = 0.65$ s), a small flip-angle excitation (point d, $t = 0.68$ s) and the subsequent ^{1}H acquisition for determination of ligand ^{1}H enhancement, and a [^{1}H-^{13}C]-SOFAST-HMQC[130] sequence for detection of ^{1}H and ^{13}C correlation for the protein methyl groups (point e, $t = 0.78$ s). Two ligand resonances were chosen to be inverted simultaneously in each experiment according to the Hadamard matrix (Equation 5.1). Four consecutive DNP NMR experiments were conducted for a complete encoding. The enhancement factors for ligand protons in each experiment were determined by comparing the peak integrals measured in the 1D DNP NMR spectra with those obtained at thermal polarization. The final concentration of folic acid was measured using absorbance at 350 nm by UV-Vis spectroscopy after each DNP experiment, while the protein concentration was determined by comparing the ^{1}H NMR signals recorded under thermal polarization to a known standard. Backbone and side-chain chemical shifts assignments of the DHFR complexed with folic acid were obtained previously,[148] by mapping reported values to the experimental conditions used.[179]

The four hyperpolarized SOFAST-HMQC spectra were first normalized by a factor, which is the product of the ligand signal enhancement from the corresponding 1D NMR spectra and the final protein concentration in each sample (Table 5.1). Spectra were then further processed with the Hadamard transform, to generate reconstructed spectra consisting of methyl NOE peaks corresponding to each of the three encoded ligand resonances.

250 ligand trial poses were generated by docking folic acid into the DHFR binding

Exp.	Enhancement			Factor				Concentration (mM)	
Number	H7	H8	H2′/H6′	H3′/H5′	Hα	Hγ	Hβ	Hβ'	Folic acid	DHFR
1	650	157	384	491	335	267	138	112	5.06	0.333
2	623	114	-255(0.69)*	-293(0.62)*	321	214	178	145	4.96	0.324
3	-403(0.65)*	131	369	-322(0.68)*	322	212	122	104	4.40	0.365
4	-375(0.60)*	139	-244(0.67)*	469	320	194	170	99	5.30	0.332

Table 5.1: Ligand signal enhancement factors and sample concentrations for four DNP-NMR experiments. The signal enhancement was determined by comparing the hyperpolarized ^{1}H signals measured with a 1° excitation with the ligand ^{1}H signals acquired at thermal polarization after the DNP-NMR experiment. * indicates that the peak was selectively inverted and the number in parentheses is the inversion factor, which was determined by comparing the relative enhancement factor for this peak in Exp.1 where no inversion was applied.

pocket using AutoDock 4.2[77] as previously described,[184] and complete relaxation and conformation exchange matrix analysis[97] were applied in the same way to simulate time-dependent polarization transfer for all involved ligand and protein ^{1}H spins. All equivalent protons, including the methyl protons, were treated as individual atoms.[97] For the ligand and protein concentrations, the enhancement factors for the hyperpolarized ligand and the additional factors to account for incomplete inversion of the ligand signals (Table 5.1), the averaged values from 4 experiments were used. Starting from the occurrence of mixing (point a in Figure 5.1, $t = 0$ s), the band-selective inversion was applied as inverting the corresponding two ligand signals at $t = 0.65$ s (point c). The simulated NMR signals in the four encoded SOFAST-HMQC spectra were obtained for all methyl protons involved in the calculation using the simulated magnetization at the start of the 2D measurement of $t = 0.78$ s (point e). The Hadamard transform was applied and the potential methyl peaks appearing at the position for the same NOE peaks were grouped together to generate the simulated NOE signals corresponding to each encoded ligand resonance.

A scoring function, NOE Score, was defined to represent the deviation of the simulated results from the experimental data, covering N methyl groups for observed NOE signals

and M methyl signals that are not observed in the 3 reconstructed spectra, as NOE Score $=$ $\sqrt{\sum_{i}^{N} \left(S_{\text{Exp},i} - S_{\text{Sim},i} \right)^2 + \sum_{i}^{M} S_{\text{Exp},i}^2}$. Here, S_i represents the relative peak intensity defined as the ratio of individual signal to the sum of intensities of all observed peaks in a single reconstructed spectrum for both simulated and experimental results. For the two indistinguishable Hδ methyls in leucine for L28, L36, L54 and L156 without stereospecific assignment, all possible combinations were calculated and the lowest NOE Score was selected. The calculated NOE Score was used to rank the 250 poses.

5.3 Results and Discussion

Rapid injection of the hyperpolarized ligand folic acid into the protein DHFR solution leads to selective polarization transfer from the ligand ^{1}H spins to their nearby ^{1}H spins on the protein. Signals resulting from this polarization transfer can be seen in the hyperpolarized 2D SOFAST-HMQC spectra shown in Figure 5.2a. A spectrum with hyperpolarized ligand (blue) is compared to a spectrum acquired with the same sample and the same SOFAST-HMQC pulse sequence after hyperpolarization has decayed (gray). The spectra are acquired with selective ^{1}H excitation of the side-chain methyl region only, to avoid consumption of the polarization of the ligand spins (2 ppm to 8.7 ppm) other than through the inevitable relaxation. In the spectrum with hyperpolarization, several strongly enhanced peaks are visible. This observation is confirmed by comparing the one-dimensional projections of the two spectra (Figure 5.2a, top). The extent of peak overlap in the projection however illustrates the necessity for acquiring the 2D data. In Figure 5.2b, slices of the 2D spectra measured with and without DNP hyperpolarization are compared, showing the signal enhancements through transfer of hyperpolarization for single peaks. These slices can also be compared with one-dimensional ^{1}H NMR traces acquired with a coherence filter to select the corresponding ^{13}C chemical shift. Such spectra were previously reported in Figure 1b of ref. [184]. They contain the same
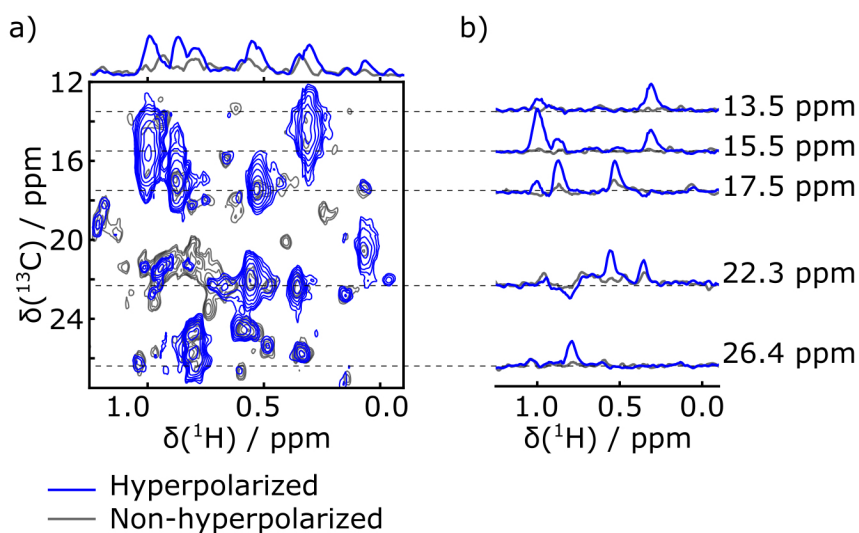
Figure 5.2: a) 2D SOFAST-HMQC spectra showing the methyl chemical shift region of 0.34 mM DHFR measured after admixing of 5.3 mM hyperpolarized folic acid (blue). A spectrum of the same sample after decay of the hyperpolarization is underlaid in gray. The spectra are recorded with 40 increments in the indirect dimension. The 1D traces at the top are positive sum projections of the 2D spectra. b) 1D slices extracted at several ^{13}C chemical shifts, as indicated by the dashed lines, from both the hyperpolarized and non-hyperpolarized 2D spectra.

correlations, albeit in narrow chemical shift ranges. The comparison illustrates that the hyperpolarized 2D measurement significantly outperforms the 1D experiment in terms of information content, with a single 2D spectrum covering a width of 24 ppm in the ^{13}C dimension.

Although polarization transfer through intermolecular NOE leads to selective enhancement of protein signals in the 2D DNP SOFAST-HMQC spectra, a correlation with the origin of polarization on individual ligand protons needs to be established in order to obtain atom-to-atom distance information. This correlation information was encoded in the spectra by applying selective inversion pulses to ligand signals before the NOE mixing time and the start of the 2D measurement.

The hyperpolarized ^{1}H NMR spectrum of the folic acid ligand shows that the largest

91
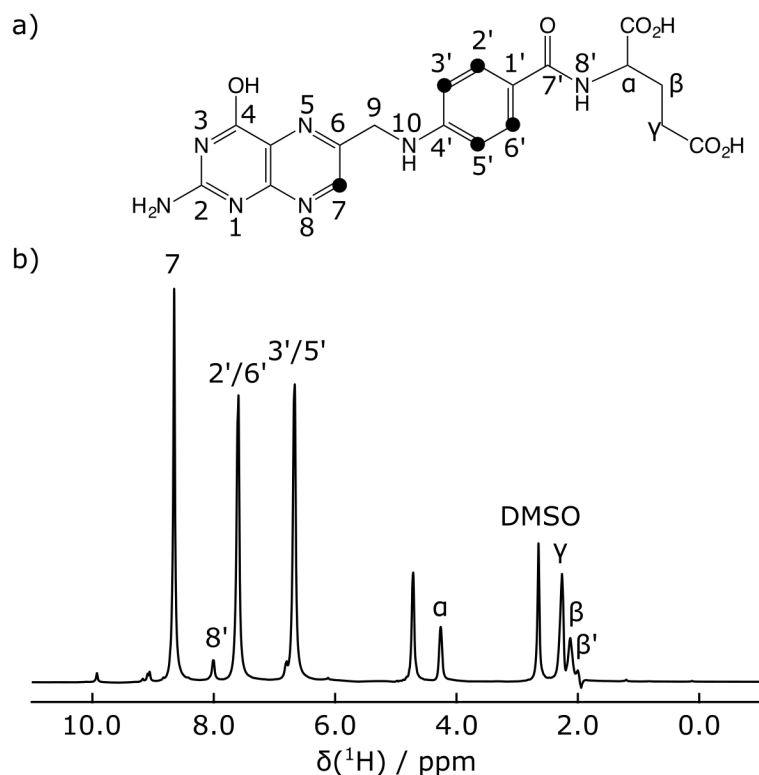
Figure 5.3: a) Structure of folic acid. b) Hyperpolarized ^{1}H spectrum of folic acid with peak assignments labeled. DMSO designates the dimethyl sulfoxide signal from the glassing matrix used for DNP hyperpolarization.

signals were observed for three ligand peaks, H_7, $H_{2'}/H_{6'}$ and $H_{3'}/H_{5'}$ (Figure 5.3). Signal enhancements compared to non-hyperpolarized NMR spectra at 400 MHz were 1160, 530, and 500, respectively. These signals are also well-resolved and outside of the spectral range containing the protein methyl resonances.

The chemical shift information for these three signals was encoded in the spectra (1)-

(4) using a Hadamard scheme, as defined by a 4×4 matrix:

$$
\mathbf{H}_4 =
\begin{array}{c c c}
\text{a} & \text{b} & \text{c} \\
\end{array}
\begin{bmatrix}
1 & 1 & 1 & 1 \\
1 & -1 & -1 & 1 \\
-1 & 1 & -1 & 1 \\
-1 & -1 & 1 & 1
\end{bmatrix}
\begin{array}{l}
(1) \\
(2) \\
(3) \\
(4)
\end{array}
\tag{5.1}
$$

Each row of this matrix corresponds to one of 4 NMR experiments. The first three columns represent three encoded ligand peaks; a: H_7, b: $H_{2'}/H_{6'}$ and c: $H_{3'}/H_{5'}$, respectively, while the fourth column represents the sum of all other ligand resonances. Entries of -1 indicate that in the respective experiment, the corresponding signal was inverted. According to this scheme, no inversion was applied in the first experiment, whereas two frequencies out of the three selected ligand resonances were selectively inverted using a dual-frequency pulse in the other three experiments.

In the first scan of each of the DNP experiments, the acquisition of a 1D NMR spectrum was integrated into the mixing time between the inversion pulse and the 2D acquisition. The excitation pulse for this spectrum consisted of a $1°$ hard pulse. These scans are shown in Figure 5.4a. They can be used to confirm the success of ligand ^{1}H encoding, and also to determine accurate enhancement factors which may vary between experiments. The results of the corresponding four Hadamard-encoded ^{1}H-^{13}C SOFAST-HMQC experiments are shown in Figure 5.4b. Compared to the spectrum with all (+1) encoding, a signal reduction for the methyl peaks can be seen in each of the other three spectra with selective ligand inversion. However, no peaks drop to the negative level. This behavior is expected because of positive polarization build-up during the mixing and stabilization period of about 0.65 s before the selective π pulse is applied, in addition to the possibility that other non-inverted ligand protons also contribute to the polarization
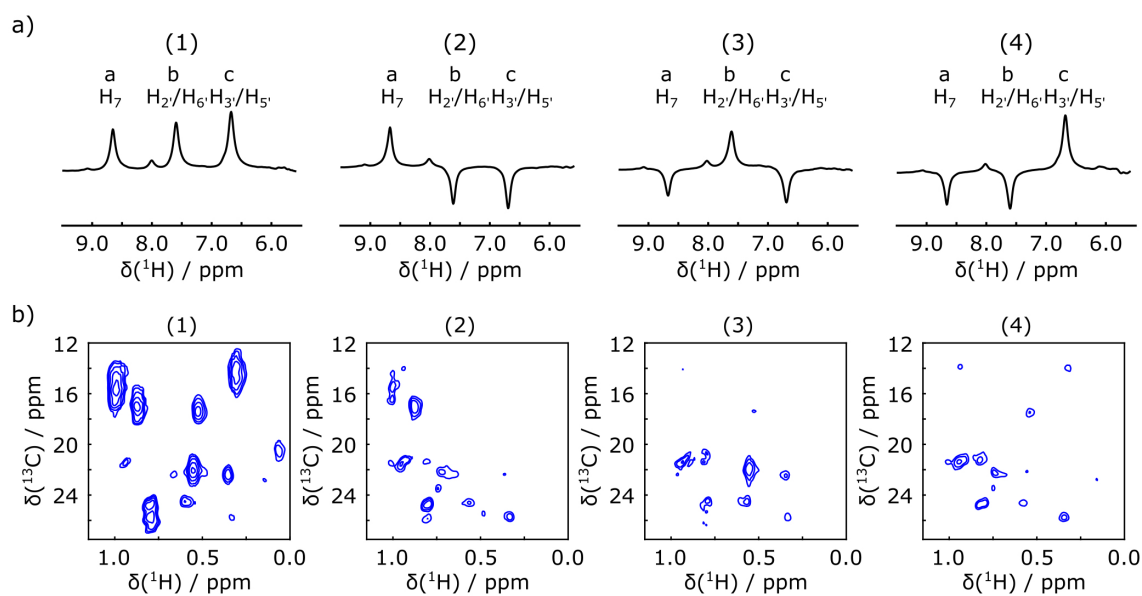
Figure 5.4: a) Hyperpolarized ^{1}H spectra of folic acid in the presence of preloaded DHFR, measured as the first scan in the DNP experiment with 1° excitation and encoded according to the Hadamard matrix in the text with selective inversion on resonances a-c. b) The following 2D SOFAST-HMQC spectra of enhanced protein signals in the methyl region.

transferred to the same methyl peak.

Reconstruction of the encoded information was performed by adding or subtracting the spectra of Figure 5.4b according to a Hadamard transform.[135] Before the reconstruction process, each spectrum was scaled with a normalization factor due to the variations in the final sample concentration after injection and the hyperpolarization level gained for each DNP experiment. The effect of the Hadamard reconstruction can be illustrated by assuming that the ligand protons a, b and c and all other ligand protons (represented by d) contribute to polarization transfer to n protein protons. The NOE signal for one protein proton i can be expressed as $(s_{ai} + s_{bi} + s_{ci} + s_{di})$, where s_{ki} ($k = a$, b, c or d) represents the signal intensity corresponding to the relative polarization transfer from each ligand proton. Upon inversion of a, b or c, a reduction of signal intensities would occur accordingly, with the change from s_{ki} to s'_{ki} ($k = a$, b, or c). When there is no NOE correlation between

94

the ligand proton k and the protein proton i, $s_{ki} = s'_{ki} = 0$. The ligand frequencies are Hadamard encoded with the matrix in 5.4. The originally obtained protein signals corresponding to different ligand resonances can be arranged into the columns of a matrix as

$$[\mathbf{S}] = \sum_{i=1}^{n} \begin{bmatrix} s_{ai} & s_{bi} & s_{ci} & s_{di} \\ s_{ai} & s'_{bi} & s'_{ci} & s_{di} \\ s'_{ai} & s_{bi} & s'_{ci} & s_{di} \\ s'_{ai} & s'_{bi} & s_{ci} & s_{di} \end{bmatrix} \tag{5.2}$$

Protein signals in a set of reconstructed spectra can be obtained by decoding the original spectra with the Hadamard matrix:

$$[\mathbf{R}] = [\mathbf{H}]_4^{\mathrm{T}}[\mathbf{S}] = \begin{bmatrix} 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} [\mathbf{S}]$$

$$= \sum_{i=1}^{n} \begin{bmatrix} 2(s_{ai} - s'_{ai}) & 0 & 0 & 0 \\ 0 & 2(s_{bi} - s'_{bi}) & 0 & 0 \\ 0 & 0 & 2(s_{ci} - s'_{ci}) & 0 \\ 2(s_{ai} - s'_{ai}) & 2(s_{bi} - s'_{bi}) & 2(s_{ci} - s'_{ci}) & 4s_{di} \end{bmatrix} \tag{5.3}$$

Thereby, the signals correlated through NOE to the 3 ligand resonances are separated into the 3 reconstructed spectra represented by the first three rows in $[\mathbf{R}]$.Therefore, pure correlated signals are generated by the Hadamard transform.

The final reconstructed 2D spectra are shown in Figure 5.5. Each of these spectra contains the protein methyl group signals originating from polarization of one of the three ligand protons a, b or c. Simultaneous incorporation of 2D NMR and Hadamard encoding with the dissolution DNP techniques allows fast acquisition of intermolecular NOEs. The information is similar to that from conventional 3D filtered NOESY experiments shown in ref.[148], but is obtained in a fraction of the time with hyperpolarization.
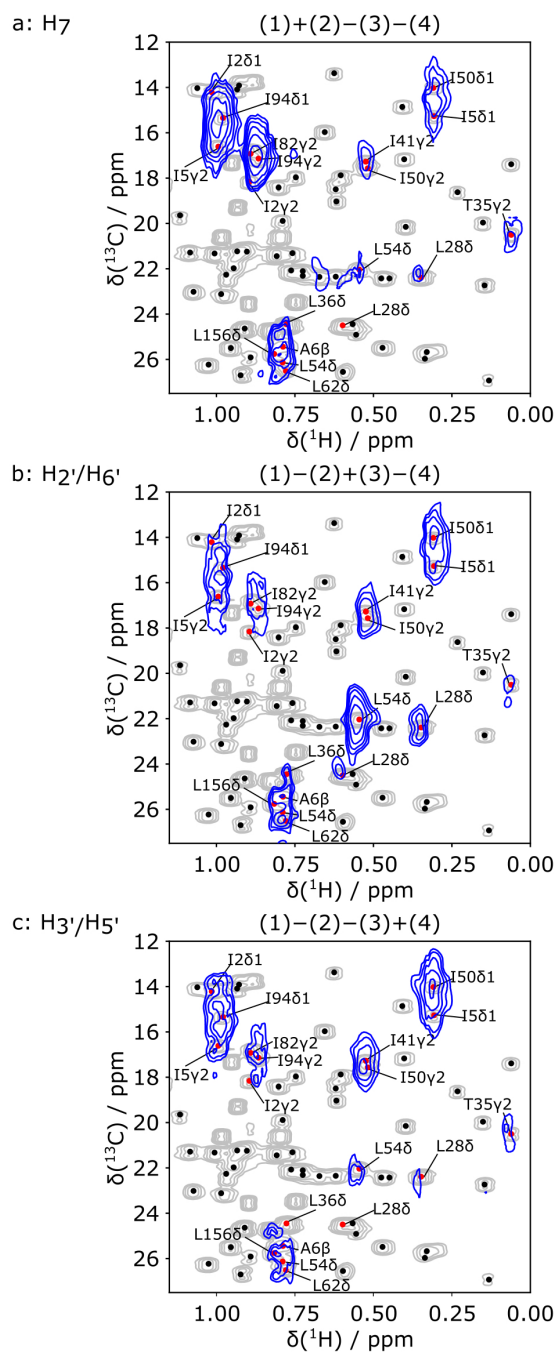
Figure 5.5: Hadamard reconstructed SOFAST-HMQC spectra (blue) of the methyl group chemical shift region of the protein. The ligand protons from which polarization originated are indicated above each spectrum. Underlaid in gray is the conventional HSQC spectrum of the protein. Red dots and methyl group labels indicate all assignments overlapped with the observed signals in the reconstructed DNP spectra. Additional assigned peaks in the conventional spectrum are represented with black dots.

96

In each of the three reconstructed spectra, distinct signal patterns are observed, allowing the determination of the source of polarization for each observable methyl group. With Hadamard encoding of the ligand side, a gain in signal-to-noise ratio is expected, since each reconstructed spectrum contains information from four individually acquired spectra. At the same time, the protein signals that do not result from transfer of encoded polarization are eliminated after the reconstruction. As a result, all observable signals in the reconstructed 2D spectra correspond to protein spins, in contact with the ligand methyl groups.

In Figure 5.5, each of the three reconstructed spectra is superimposed on a conventional ^{1}H-^{13}C HSQC spectrum. The conventional spectrum shows the chemical shift assignments of each methyl group. The same spectrum with the detailed assignments of all methyl groups is further displayed in Figure 5.6. Based on these assignments, the signals in the hyperpolarized 2D spectra can be identified for use in a scoring function for ranking the docked poses. The resolution of the hyperpolarized spectra however does not allow an unambiguous assignment in all cases. Therefore, all candidates that show overlap with the peaks in the reconstructed DNP spectra were considered as potentially part of the binding site and included in the further calculation. In total, 16 methyl groups were identified as potential assignments for the 9 observed NOE peaks in Figure 5.5.

A scoring function (NOE Score) for quantifying the difference between the simulated and experimental NOE signals was defined to rank the 250 poses. For each of the 250 folic acid poses generated by the docking program, the polarization transfer process that occurred during the D-DNP experiment covering the whole process starting from the mixing of the sample to the start of 2D acquisition for all ligand protons and protein protons located within 6 Å of the ligand were simulated based on the complete relaxation and conformational exchange matrix analysis (CORCEMA).[97] This strategy is similar to that described in ref.[184], but results in a more accurate scoring function because of the
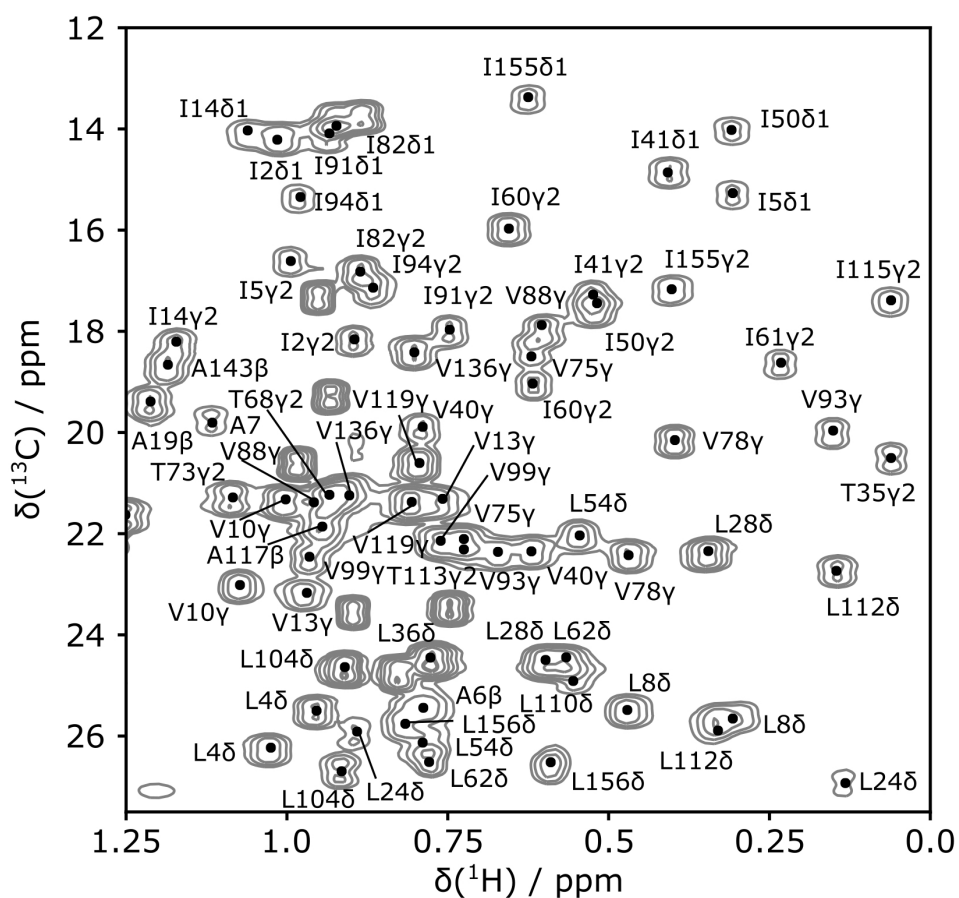
Figure 5.6: 2D HSQC spectrum of 1.5 mM DHFR in the presence of 15 mM folic acid acquired on a 500 MHz NMR spectrometer equipped with a TCI cryoprobe. The chemical shift assignments were previously determined,[148] adapted from published values[179] to current experimental conditions.

increased number of constraints from the 2D spectra.

The Hadamard transform was applied to the calculated results for all methyl groups within the chemical shift range in Figure 5.5. The sum of the calculated signals for each of the methyl groups with assignments overlapping with the observed peaks in the reconstructed spectra represents the expected signal for the observed NOE peak. For the remaining methyl groups, no NOE signals are expected. As a result, the scoring function includes information both for protons with and without NOE signals.

The five poses with the best NOEScore are shown as blue structures in Figure 5.7a , with all methyl protons within 5 Å of these poses also displayed. These methyl groups cover most of the NOE signals observed in the reconstructed spectra in Figure 5.5, except
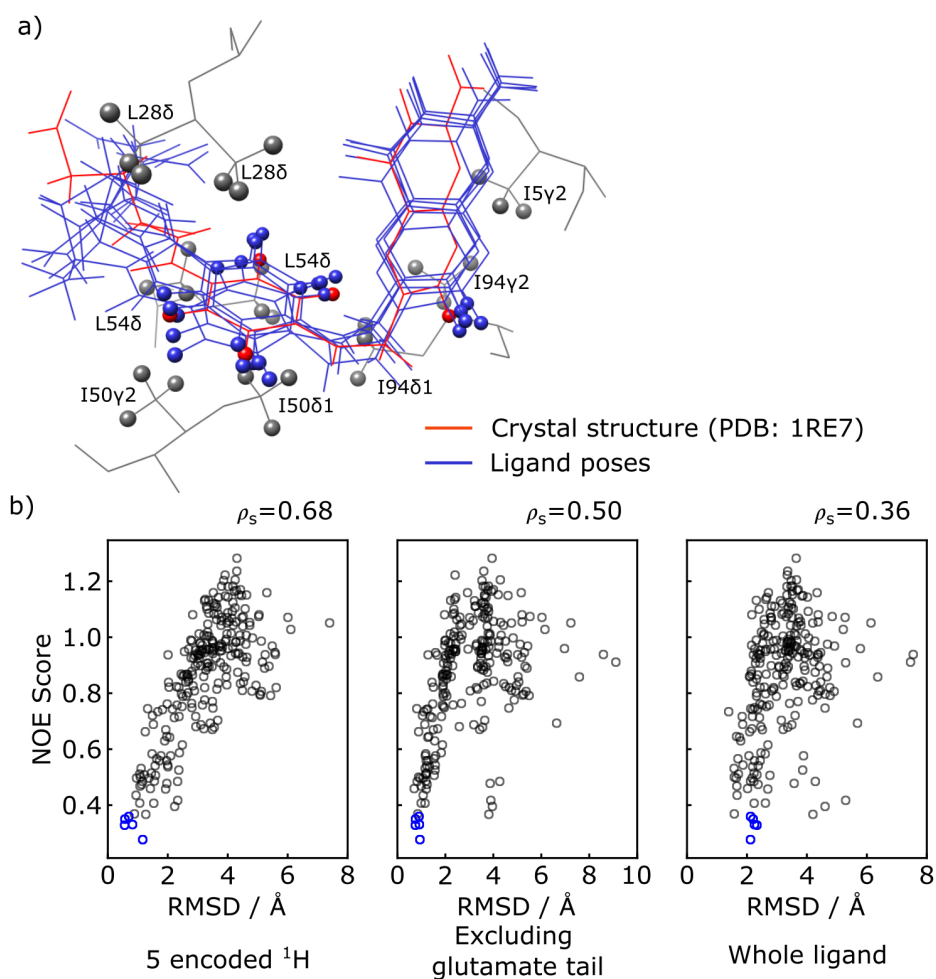
Figure 5.7: Evaluation of ligand trial poses ranked by NOEScore. a) Overlay of the five docked poses with the best NOEScore (blue) and the ligand in the crystal structure (red). The five encoded protons of the ligand are shown as spheres in all of the poses. Protein methyl groups from the crystal structure that are within 5 Å of the five selected poses are represented with gray spheres. b) Correlation plot of NOEScore vs. RMSD between the trial pose and the crystal structure. The blue circles represent the five poses displayed in (a). The RMSD values in the three panels are calculated considering the five ligand protons encoded with selective inversion (left), the ligand structure excluding the glutamate portion (middle), and the whole ligand structure (right).

for a weak NOE signal assigned to T35γ2. For comparison, a reference ligand structure from the crystal structure of DHFR-folic acid complex (PDB:1RE7[177]) is underlaid in red. It can be seen that the agreement among the selected poses, as well as with the reference structure, is high in particular for the 5 encoded protons with the averaged structural root-mean-square deviation (RMSD) values of 0.76 Å. When considering the entire pteroyl moiety, where the encoded protons are located, the average RMSD values are 0.85 Å, while the glutamate tail in the rest of the ligand remains relatively disordered. Similar conclusions can be drawn from the correlation plot of the NOE Score vs. structural RMSD values between the ligand poses and the reference structure (Figure 5.7b). The highest correlation of these quantities is observed when considering only the 5 ligand protons that were encoded (left panel). This is followed by considering exclusively the pteroyl moiety (central panel). The entire ligand structure shows the lowest correlation (right panel). This result is reasonable, considering that atoms in the glutamate tail, without encoding, have no direct NOE correlation information. The five poses selected by the NOE Score are compared with the five poses with lowest calculated binding energy in Figure 5.8. The latter structures, without experimental input, give average RMSD values to the reference structure of 3.63 Å for the 5 encoded protons and 3.16 Å for the pteroyl moiety. The inclusion of the experimental information therefore leads to a clear improvement in the accuracy of ligand pose selection.

In this method, methyl group chemical shift assignment information is required so that the experimental signals can be compared to the back-calculated NOEs. As described previously, the simulation of the intermolecular NOEs includes all methyl groups sharing similar ^{1}H and ^{13}C chemical shifts. Methyl groups that are not located in the binding site, although included in the calculation, do not significantly contribute to the simulated NOE intensities for the correct ligand poses if located at a large distance from the ligand spins. This feature allows some ambiguity in the NOE assignment and might

L28δ

L28δ

I5γ2

L54δ

I94γ2

L54δ

I94δ1

I50γ2

I50δ1

—— Crystal structure (PDB: 1RE7)
—— Ligand poses
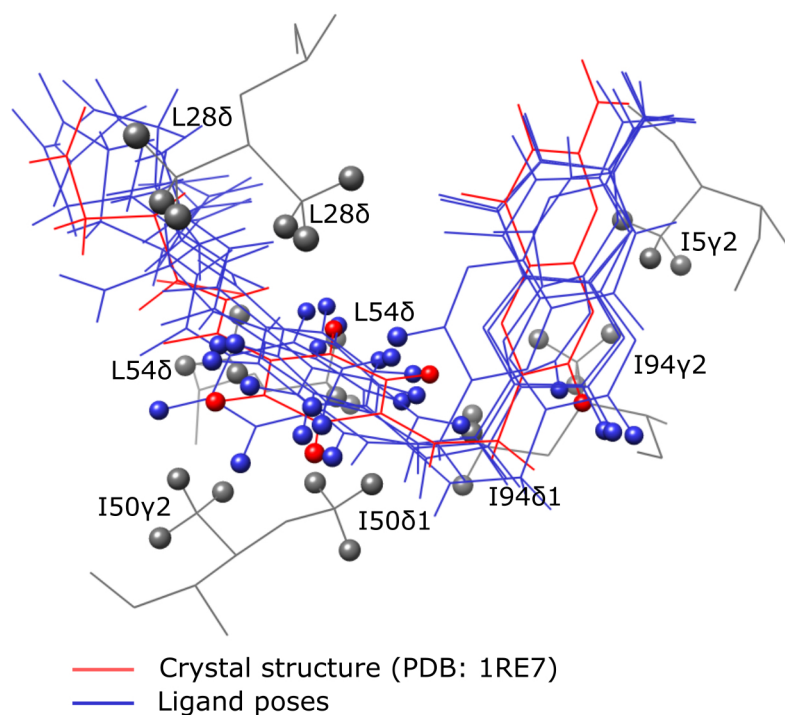
Figure 5.8: Overlay of the five docked poses with the lowest binding energy calculated by the AutoDock program (blue) and the ligand in the crystal structure (red). The average RMSD value of the 5 selected poses against the reference structure is 3.63 Å when calculating for the 5 encoded protons, and 3.16 Å for the pteroyl moiety.

provide the potential for extending the current method to an assignment-free approach. Several methods have already been described for determination of the structure of ligand-protein interaction site based on NOE distance restraints, where no protein resonance assignments are required. Constantine *et al.* proposed to rescore the pre-docked ligand based on matching the observed and predicted patterns of intermolecular NOEs.[193] In a distance restraints-driven method by Orts *et al.*, all possible assignment combinations are screened, with filtering steps included to reduce the total number of possibilities to be calculated.[223] To apply these strategies with DNP-assisted intermolecular NOE experiments, the resolution of the 2D measurement should be improved to avoid peak overlapping. In addition, highly efficient experiments with more ligand frequencies

encoded can be developed to collect a larger number of NOEs.

In the 2D SOFAST experiment, spin polarization of the measured protein spins is consumed in each scan, so that the success of multi-dimensional spectroscopy depends on polarization continuously transferring from the ligand to the protein. As a consequence, hyperpolarization loss due to T_1 relaxation of ligand ^{1}H spins limits the total experimental time and the number of indirect points that can be measured. Techniques that can slow down the ligand T_1 relaxation, including separating the radicals after dissolution, increasing the temperature, or using deuterated solvent, provide possible ways to increase both the resolution and sensitivity of the experiment.

5.4 Conclusions

In summary, the combination of hyperpolarized 2D NMR spectroscopy for protein detection and Hadamard encoding of ligand resonances results in pseudo 3D measurements of specific intermolecular NOEs between protein methyl groups and ligand proton spins. In DHFR, a total of four DNP-NMR experiments identified 9 different NOE contacts, which resulted in the selection of a ligand structure that shows a good agreement with the reference structure for the part of ligand containing the encoded protons, giving an average root-mean-square value of 0.76 Å for the five encoded protons. The described method of using multi-dimensional NOE data from hyperpolarized spins to score pre-docked ligand poses can correctly solve the structure of protein-ligand complexes in the binding site. Future applications may include acceleration of ligand optimization in structure-based rational drug design.

6. SUMMARY

Monitoring polarization transfer between protein and ligand or between ligand protons based on the nuclear Overhauser effect (NOE) is a robust means for structural characterization of ligand-protein binding. However, NOE detection is subject to significant sensitivity limitations, since only a small fraction of signal transfers from the source spin to the target spin. The amount of polarization transfer is further reduced in ligand-protein binding observation with both the ligand and protein only partially remaining in their bound states at any time. Dissolution dynamic nuclear polarization (D-DNP) is a hyperpolarization technique, which offers a solution for the low sensitivity of traditional NOE experiments. The generation of hyperpolarized NOE source spins using D-DNP can provide a signal enhancement of several thousand fold.

The sensitivity gain provided by D-DNP can be combined with ligand-observed NMR approaches for characterizing ligand-protein binding. A hyperpolarized NMR method of measuring intra-ligand NOEs was developed for investigating ligand-protein binding. Without the need for observing protein signals, there are no limitations on the target size, and the concentration of macromolecules can be substantially reduced. Time-dependent polarization transfer to a pre-suppressed ligand spin from its nearby hyperpolarized ligand spins was monitored, during which the hyperpolarization was partially consumed in each scan by a small-flip angle pulse. An NOE buildup curve was obtained from a single hyperpolarization, compared to multiple experiments with different mixing times required for the non-hyperpolarized NOE measurement. A change in the sign of the NOE with and without protein confirmed the binding. In addition, cross-relaxation rates between ligand protons in the bound form were determined, which can provide structural information about the bound ligand conformation. This method was demonstrated with proteins

immobilized on micrometer-sized polystyrene beads. A ten-fold enhancement in the cross-relaxation rate was detected compared to with protein in solution, due to a longer rotational correlation time associated with the ligand when binding to the immobilized protein. With hyperpolarization, single-scan detection of NOE was achieved with the ligand at low micromolar concentrations and the protein at high nanomolar concentrations. The compatibility of the D-DNP technique with NMR experiments in heterogeneous environments was demonstrated in this experiment, revealing its potential application to *in vivo* characterization of ligand-protein interactions directly on the cell surface. The short acquisition time of a few seconds can ensure stable cell suspension and high cell viability during the entire NMR measurement. In addition, the signal enhancement provided by D-DNP effectively excludes background signals from any non-hyperpolarized species.

Detailed structural information on protein-ligand complexes can be obtained by detecting NOE based polarization transfer directly from ligand to protein in protein-observed NMR. In spite of the substantial signal enhancement by D-DNP, the application of this approach for determination of molecular structure remains challenging due to its incompatibility with traditional multi-dimensional NMR. In order to obtain spin-to-spin correlation information, which is essential for structure elucidation, we first coupled 1D proton NMR experiments with a selective inversion on a ligand proton and a ^{13}C coherence selection on the protein side. The NMR measurement of enhanced protein signals was further integrated with the rapid 2D NMR for correlated protein ^{1}H and ^{13}C chemical shifts and a pseudo-third dimension for ligand ^{1}H encoding through Hadamard spectroscopy. Compared to the 1D DNP NMR, a larger number of NOE contacts were obtained with improved resolution in a smaller number of hyperpolarized experiments. This method of using multi-dimensional hyperpolarized NMR spectroscopy for fast measurement of intermolecular interactions can be potentially applied to other small molecule-macromolecule systems, such as ligands bound to DNA or polymers.

104

A protocol for combining hyperpolarized NOE restraints with molecular docking to determine the structure of the protein-ligand binding interface was developed. Hyperpolarized intermolecular NOEs were simulated by calculating polarization transfer in networks of nuclear spins for each pre-docked ligand pose generated by a docking program. The experimental NOE intensities were compared to the calculated data for ranking of the computationally generated poses. The structure of the ligand folic acid in the binding site of the protein dihydrofolate reductase was obtained, which was in good agreement with the reference structure.

In conclusion, hyperpolarized NMR methods were developed in this dissertation for structural characterization of protein-ligand complexes. Dissolution DNP has gained considerable attention as a powerful and versatile technique to enhance the sensitivity of NMR measurements. However, applications of D-DNP to solving molecular structures still are rare, possibly due to the difficulty in obtaining correlation information within a short acquisition period. To circumvent this limitation, different fast multi-dimensional NMR approaches were integrated with D-DNP. The measurement of intermolecular NOEs with a pseudo-3D D-DNP NMR method was demonstrated for the first time. The obtained hyperpolarized NOE data, in combination with molecular docking, led to the determination of a protein-ligand complex structure in its binding site. The improved efficiency in acquiring structural information on the ligand binding epitope by D-DNP NMR can potentially facilitate rational ligand optimization in structure-based drug design. Applications of these methods may further extended to gaining insights on a wide range of small molecule-macromolecule interactions in biological systems.

REFERENCES

(1) Berman, H. M. *Nucleic Acids Res.* **2000**, *28*, 235–242.

(2) Kalyaanamoorthy, S.; Chen, Y.-P. P. *Drug Discov. Today* **2011**, *16*, 831–839.

(3) Lounnas, V.; Ritschel, T.; Kelder, J.; McGuire, R.; Bywater, R. P.; Foloppe, N. *Comput. Struct. Biotechnol. J.* **2013**, *5*, e201302011.

(4) Greer, J.; Erickson, J. W.; Baldwin, J. J.; Varney, M. D. *J. Med. Chem.* **1994**, *37*, 1035–1054.

(5) Anderson, A. C. *Chem. Biol.* **2003**, *10*, 787–797.

(6) Batool, M.; Ahmad, B.; Choi, S. *Int. J. Mol. Sci.* **2019**, *20*, 2783.

(7) Jorgensen, W. L. *Science* **2004**, *303*, 1813–1818.

(8) Carvalho, A. L.; Trincão, J.; Romão, M. J. In *Ligand-Macromolecular Interactions in Drug Discovery: Methods and Protocols*, Roque, A. C. A., Ed.; Methods in Molecular Biology; Humana Press: Totowa, NJ, 2010, pp 31–56.

(9) Cooper, D. R.; Porebski, P. J.; Chruszcz, M.; Minor, W. *Expert Opin. Drug Discov.* **2011**, *6*, 771–782.

(10) Blundell, T. L.; Jhoti, H.; Abell, C. *Nat. Rev. Drug Discov.* **2002**, *1*, 45–54.

(11) Zheng, H.; Handing, K. B.; Zimmerman, M. D.; Shabalin, I. G.; Almo, S. C.; Minor, W. *Expert Opin. Drug Discov.* **2015**, *10*, 975–989.

(12) Davis, A. M.; Teague, S. J.; Kleywegt, G. J. *Angew. Chem. Int. Ed.* **2003**, *42*, 2718–2736.

(13) Zheng, H.; Hou, J.; Zimmerman, M. D.; Wlodawer, A.; Minor, W. *Expert Opin. Drug Discov.* **2014**, *9*, 125–137.

(14) Sugiki, T.; Furuita, K.; Fujiwara, T.; Kojima, C. *Molecules* **2018**, *23*, 148.

(15) Ziarek, J. J.; Peterson, F. C.; Lytle, B. L.; Volkman, B. F. *Methods Enzymol.* **2011**, *493*, 241–275.

(16) Orts, J.; Gossert, A. D. *Methods* **2018**, *138-139*, 3–25.

(17) Güntert, P.; Buchner, L. *J. Biomol. NMR* **2015**, *62*, 453–471.

(18) Śledź, P.; Caflisch, A. *Curr. Opin. Struct. Biol.* **2018**, *48*, 93–102.

(19) Ferreira, L.; dos Santos, R.; Oliva, G.; Andricopulo, A. *Molecules* **2015**, *20*, 13384–13421.

(20) Sturlese, M.; Bellanda, M.; Moro, S. *Mol. Inf.* **2015**, *34*, 513–525.

(21) Stark, J. L.; Powers, R. In *NMR of Proteins and Small Biomolecules*; Springer Berlin Heidelberg: Berlin, Heidelberg, 2011; Vol. 326, pp 1–34.

(22) Congreve, M.; Carr, R.; Murray, C.; Jhoti, H. *Drug Discov. Today* **2003**, *8*, 876–877.

(23) Gossert, A. D.; Jahnke, W. *Prog. Nucl. Magn. Reson. Spectrosc.* **2016**, *97*, 82–125.

(24) *Protein NMR spectroscopy: principles and practice*, 2nd ed; Cavanagh, J., Ed., OCLC: ocm70218044; Academic Press: Amsterdam ; Boston, 2007.

(25) Kumar, A. *Prog. Nucl. Magn. Reson. Spectrosc.* **2000**, *37*, 191–319.

(26) Williamson, M. P. *Prog. Nucl. Magn. Reson. Spectrosc.* **2013**, *73*, 1–16.

(27) Stevens, S. Y.; Sanker, S.; Kent, C.; Zuiderweg, E. R. *Nat. Struct. Biol.* **2001**, *8*, 947–952.

(28) Foster, M. P.; Wuttke, D. S.; Clemens, K. R.; Jahnke, W.; Radhakrishnan, I.; Tennant, L.; Reymond, M.; Chung, J.; Wright, P. E. *J. Biomol. NMR* **1998**, *12*, 51–71.

(29) Medek, A.; Hajduk, P. J.; Mack, J.; Fesik, S. W. *J. Am. Chem. Soc.* **2000**, *122*, 1241–1242.

(30) Lugovskoy, A. A.; Degterev, A. I.; Fahmy, A. F.; Zhou, P.; Gross, J. D.; Yuan, J.; Wagner, G. *J. Am. Chem. Soc.* **2002**, *124*, 1234–1240.

(31) Riedinger, C.; Endicott, J. A.; Kemp, S. J.; Smyth, L. A.; Watson, A.; Valeur, E.; Golding, B. T.; Griffin, R. J.; Hardcastle, I. R.; Noble, M. E.; McDonnell, J. M. *J. Am. Chem. Soc.* **2008**, *130*, 16038–16044.

(32) Breeze, A. L. *Prog. Nucl. Magn. Reson. Spectrosc.* **2000**, *36*, 323–372.

(33) Post, C. *Curr. Opin. Struct. Biol.* **2003**, *13*, 581–588.

(34) Inooka, H.; Ohtaki, T.; Kitahara, O.; Ikegami, T.; Endo, S.; Kitada, C.; Ogi, K.; Onda, H.; Fujino, M.; Shirakawa, M. *Nat. Struct. Biol.* **2001**, *8*, 161–165.

(35) Johnson, M. A.; Pinto, B. *Carbohydr. Res.* **2004**, *339*, 907–928.

(36) Sánchez-Pedregal, V. M.; Reese, M.; Meiler, J.; Blommers, M. J. J.; Griesinger, C.; Carlomagno, T. *Angew. Chem. Int. Ed.* **2005**, *44*, 4172–4175.

(37) Li, D.; DeRose, E. F.; London, R. E. *J. Biomol. NMR* **1999**, *15*, 71–76.

(38) London, R. E. *J. Magn. Reson.* **1999**, *141*, 301–311.

(39) Shuker, S. B.; Hajduk, P. J.; Meadows, R. P.; Fesik, S. W. *Science* **1996**, *274*, 1531–1534.

(40) Becattini, B.; Culmsee, C.; Leone, M.; Zhai, D.; Zhang, X.; Crowell, K. J.; Rega, M. F.; Landshamer, S.; Reed, J. C.; Plesnila, N.; Pellecchia, M. *Proc. Natl. Acad. Sci.* **2006**, *103*, 12602–12606.

(41) Becattini, B.; Pellecchia, M. *Chem. Eur. J.* **2006**, *12*, 2658–2662.

(42) Orts, J.; Griesinger, C.; Carlomagno, T. *J. Magn. Reson.* **2009**, *200*, 64–73.

(43) Mayer, M.; Meyer, B. *Angew. Chem. Int. Ed.* **1999**, *38*, 1784–1788.

(44) Mayer, M.; Meyer, B. *J. Am. Chem. Soc.* **2001**, *123*, 6108–6117.

(45) Kemper, S.; Patel, M. K.; Errey, J. C.; Davis, B. G.; Jones, J. A.; Claridge, T. D. *J. Magn. Reson.* **2010**, *203*, 1–10.

(46) Yan, J.; Kline, A. D.; Mo, H.; Shapiro, M. J.; Zartler, E. R. *J. Magn. Reson.* **2003**, *163*, 270–276.

(47) Dalvit, C.; Fogliatto, G.; Stewart, A.; Veronesi, M.; Stockman, B. *J. Biomol. NMR* **2001**, *21*, 349–359.

(48) Dalvit, C.; Pevarello, P.; Tatò, M.; Veronesi, M.; Vulpetti, A.; Sundström, M. *J. Biomol. NMR* **2000**, *18*, 65–68.

(49) Ludwig, C.; Michiels, P. J. A.; Lodi, A.; Ride, J.; Bunce, C.; Günther, U. L. *ChemMedChem* **2008**, *3*, 1371–1376.

(50) Ludwig, C.; Michiels, P. J. A.; Wu, X.; Kavanagh, K. L.; Pilka, E.; Jansson, A.; Oppermann, U.; Günther, U. L. *J. Med. Chem.* **2008**, *51*, 1–3.

(51) Cala, O.; Guillière, F.; Krimm, I. *Anal. Bioanal. Chem.* **2014**, *406*, 943–956.

(52) Bartoschek, S.; Klabunde, T.; Defossa, E.; Dietrich, V.; Stengelin, S.; Griesinger, C.; Carlomagno, T.; Focken, I.; Wendt, K. U. *Angew. Chem. Int. Ed.* **2010**, *49*, 1426–1429.

(53) Shimada, I.; Ueda, T.; Kofuku, Y.; Eddy, M. T.; Wüthrich, K. *Nat. Rev. Drug Discov.* **2019**, *18*, 59–82.

(54) Mari, S.; Invernizzi, C.; Spitaleri, A.; Alberici, L.; Ghitti, M.; Bordignon, C.; Traversari, C.; Rizzardi, G.-P.; Musco, G. *Angew. Chem. Int. Ed.* **2010**, *49*, 1071–1074.

(55) Madge, P. D.; Maggioni, A.; Pascolutti, M.; Amin, M.; Waespy, M.; Bellette, B.; Thomson, R. J.; Kelm, S.; von Itzstein, M.; Haselhorst, T. *Sci. Rep.* **2016**, *6*, 36012.

(56) Brancaccio, D.; Diana, D.; Di Maro, S.; Di Leva, F. S.; Tomassi, S.; Fattorusso, R.; Russo, L.; Scala, S.; Trotta, A. M.; Portella, L.; Novellino, E.; Marinelli, L.; Carotenuto, A. *J. Med. Chem.* **2018**, *61*, 2910–2923.

(57) Saio, T.; Ogura, K.; Shimizu, K.; Yokochi, M.; Burke, T. R.; Inagaki, F. *J. Biomol. NMR* **2011**, *51*, 395–408.

(58) Jahnke, W.; Rüdisser, S.; Zurini, M. *J. Am. Chem. Soc.* **2001**, *123*, 3149–3150.

(59) Pinto, L. F.; Correa, J.; Zhao, L.; Riguera, R.; Fernandez-Megia, E. *ACS Omega* **2018**, *3*, 2974–2983.

(60) Pintacuda, G.; John, M.; Su, X.-C.; Otting, G. *Acc. Chem. Res.* **2007**, *40*, 206–212.

(61) Brath, U.; Swamy, S. I.; Veiga, A. X.; Tung, C.-C.; Van Petegem, F.; Erdélyi, M. *J. Am. Chem. Soc.* **2015**, *137*, 11391–11398.

(62) Häussinger, D.; Huang, J.-r.; Grzesiek, S. *J. Am. Chem. Soc.* **2009**, *131*, 14761–14767.

(63) Peters, F.; Maestre-Martinez, M.; Leonov, A.; Kovačič, L.; Becker, S.; Boelens, R.; Griesinger, C. *J. Biomol. NMR* **2011**, *51*, 329–337.

(64) John, M.; Pintacuda, G.; Park, A. Y.; Dixon, N. E.; Otting, G. *J. Am. Chem. Soc.* **2006**, *128*, 12910–12916.

(65) Nitsche, C.; Otting, G. *Curr. Opin. Struct. Biol.* **2018**, *48*, 16–22.

(66) Guan, J.-Y.; Keizers, P. H. J.; Liu, W.-M.; Löhr, F.; Skinner, S. P.; Heeneman, E. A.; Schwalbe, H.; Ubbink, M.; Siegal, G. *J. Am. Chem. Soc.* **2013**, *135*, 5859–5868.

(67) Chen, W.-N.; Nitsche, C.; Pilla, K. B.; Graham, B.; Huber, T.; Klein, C. D.; Otting, G. *J. Am. Chem. Soc.* **2016**, *138*, 4539–4546.

(68) Kitchen, D. B.; Decornez, H.; Furr, J. R.; Bajorath, J. *Nat. Rev. Drug Discov.* **2004**, *3*, 935–949.

(69) Meng, X.-Y.; Zhang, H.-X.; Mezei, M.; Cui, M. *Curr. Comput. Aided-Drug Des.* **2011**, *7*, 146–157.

(70) Kuntz, I. D.; Blaney, J. M.; Oatley, S. J.; Langridge, R.; Ferrin, T. E. *J. Mol. Biol.* **1982**, *161*, 269–288.

(71) Halperin, I.; Ma, B.; Wolfson, H.; Nussinov, R. *Proteins: Struct., Funct., Genet.* **2002**, *47*, 409–443.

(72) Davis, I. W.; Baker, D. *J. Mol. Biol.* **2009**, *385*, 381–392.

(73) Ravindranath, P. A.; Forli, S.; Goodsell, D. S.; Olson, A. J.; Sanner, M. F. *PLOS Comput. Biol.* **2015**, *11*, ed. by Fetrow, J. S., e1004586.

(74) Gagnon, J. K.; Law, S. M.; Brooks, C. L. *J. Comput. Chem.* **2016**, *37*, 753–762.

(75) Lang, P. T.; Brozell, S. R.; Mukherjee, S.; Pettersen, E. F.; Meng, E. C.; Thomas, V.; Rizzo, R. C.; Case, D. A.; James, T. L.; Kuntz, I. D. *RNA* **2009**, *15*, 1219–1230.

(76) Goodsell, D. S.; Olson, A. J. *Proteins* **1990**, *8*, 195–202.

(77) Morris, G. M.; Huey, R.; Lindstrom, W.; Sanner, M. F.; Belew, R. K.; Goodsell, D. S.; Olson, A. J. *J. Comput. Chem.* **2009**, *30*, 2785–2791.

(78) Trott, O.; Olson, A. J. *J. Comput. Chem.* **2009**, NA–NA.

(79) Jones, G.; Willett, P.; Glen, R. C.; Leach, A. R.; Taylor, R. *J. Mol. Biol.* **1997**, *267*, 727–748.

(80) Wang, Z.; Sun, H.; Yao, X.; Li, D.; Xu, L.; Li, Y.; Tian, S.; Hou, T. *Phys. Chem. Chem. Phys.* **2016**, *18*, 12964–12975.

(81) Pagadala, N. S.; Syed, K.; Tuszynski, J. *Biophys. Rev.* **2017**, *9*, 91–102.

(82) Plewczynski, D.; Łaźniewski, M.; Augustyniak, R.; Ginalski, K. *J. Comput. Chem.* **2011**, *32*, 742–755.

(83) Dar, A. M.; Mir, S. *J. Anal. Bioanal. Tech.* **2017**, *08*.

(84) McCoy, M. A.; Wyss, D. F. *J. Am. Chem. Soc.* **2002**, *124*, 11758–11763.

(85) Dominguez, C.; Boelens, R.; Bonvin, A. M. J. J. *J. Am. Chem. Soc.* **2003**, *125*, 1731–1737.

(86) Schieborr, U.; Vogtherr, M.; Elshorst, B.; Betz, M.; Grimme, S.; Pescatore, B.; Langer, T.; Saxena, K.; Schwalbe, H. *ChemBioChem* **2005**, *6*, 1891–1898.

(87) Stark, J.; Powers, R. *J. Am. Chem. Soc.* **2008**, *130*, 535–545.

(88) McCoy, M. A.; Wyss, D. F. *J. Biomol. NMR* **2000**, *18*, 189–198.

(89) Cioffi, M.; Hunter, C. A.; Packer, M. J.; Spitaleri, A. *J. Med. Chem.* **2008**, *51*, 2512–2517.

(90) González-Ruiz, D.; Gohlke, H. *J. Chem. Inf. Model.* **2009**, *49*, 2260–2271.

(91) Aguirre, C.; ten Brink, T.; Cala, O.; Guichou, J.-F.; Krimm, I. *J. Biomol. NMR* **2014**, *60*, 147–156.

(92) Wang, B.; Raha, K.; Merz, K. M. *J. Am. Chem. Soc.* **2004**, *126*, 11430–11431.

(93) Wang, B.; Brothers, E. N.; van der Vaart, A.; Merz, K. M. *J. Chem. Phys.* **2004**, *120*, 11392–11400.

(94) Shah, D. M.; Ab, E.; Diercks, T.; Hass, M. A. S.; van Nuland, N. A. J.; Siegal, G. *J. Med. Chem.* **2012**, *55*, 10786–10790.

(95) Mohanty, B.; Williams, M. L.; Doak, B. C.; Vazirani, M.; Ilyichova, O.; Wang, G.; Bermel, W.; Simpson, J. S.; Chalmers, D. K.; King, G. F.; Mobli, M.; Scanlon, M. J. *J. Biomol. NMR* **2016**, *66*, 195–208.

112

(96)   Vögeli, B. *Prog. Nucl. Magn. Reson. Spectrosc.* **2014**, *78*, 1–46.

(97)   Moseley, H.; Curto, E.; Krishna, N. *J. Magn. Reson. B* **1995**, *108*, 243–261.

(98)   Curto, E. V.; Moseley, H. N. B.; Krishna, N. R. *J. Comput. Aided Mol. Des.* **1996**, *10*, 361–371.

(99)   Jayalakshmi, V.; Krishna, N. *J. Magn. Reson.* **2002**, *155*, 106–118.

(100)   Jayalakshmi, V.; Rama Krishna, N. *J. Magn. Reson.* **2004**, *168*, 36–45.

(101)   Reese, M.; Sánchez-Pedregal, V. M.; Kubicek, K.; Meiler, J.; Blommers, M. J. J.; Griesinger, C.; Carlomagno, T. *Angew. Chem. Int. Ed.* **2007**, *46*, 1864–1868.

(102)   Orts, J.; Grimm, S. K.; Griesinger, C.; Wendt, K. U.; Bartoschek, S.; Carlomagno, T. *Chem. – Eur. J* **2008**, *14*, 7517–7520.

(103)   Orts, J.; Tuma, J.; Reese, M.; Grimm, S. K.; Monecke, P.; Bartoschek, S.; Schiffer, A.; Wendt, K. U.; Griesinger, C.; Carlomagno, T. *Angew. Chem. Int. Ed.* **2008**, *47*, 7736–7740.

(104)   Orts, J.; Bartoschek, S.; Griesinger, C.; Monecke, P.; Carlomagno, T. *J. Biomol. NMR* **2012**, *52*, 23–30.

(105)   Carver, T. R.; Slichter, C. P. *Phys. Rev.* **1953**, *92*, 212–213.

(106)   Overhauser, A. W. *Phys. Rev.* **1953**, *92*, 411–415.

(107)   Abragam, A.; Goldman, M. *Rep. Prog. Phys.* **1978**, *41*, 395–467.

(108)   Bowers, C. R.; Weitekamp, D. P. *Phys. Rev. Lett.* **1986**, *57*, 2645–2648.

(109)   Eisenschmid, T. C.; Kirss, R. U.; Deutsch, P. P.; Hommeltoft, S. I.; Eisenberg, R.; Bargon, J.; Lawler, R. G.; Balch, A. L. *J. Am. Chem. Soc.* **1987**, *109*, 8089–8091.

(110)   Adams, R. W.; Aguilar, J. A.; Atkinson, K. D.; Cowley, M. J.; Elliott, P. I. P.; Duckett, S. B.; Green, G. G. R.; Khazal, I. G.; Lopez-Serrano, J.; Williamson, D. C. *Science* **2009**, *323*, 1708–1711.

(111)   Walker, T. G.; Happer, W. *Rev. Mod. Phys.* **1997**, *69*, 629–642.

(112)   Reynolds, S.; Patel, H. *Appl. Magn. Reson.* **2008**, *34*, 495–508.

(113)   Ardenkjaer-Larsen, J. H.; Fridlund, B.; Gram, A.; Hansson, G.; Hansson, L.; Lerche, M. H.; Servin, R.; Thaning, M.; Golman, K. *Proc. Natl. Acad. Sci.* **2003**, *100*, 10158–10163.

(114)   Nelson, S. J. et al. *Sci. Transl. Med.* **2013**, *5*, 198ra108–198ra108.

(115)   Golman, K.; Zandt, R. i.; Lerche, M.; Pehrson, R.; Ardenkjaer-Larsen, J. H. *Cancer Res.* **2006**, *66*, 10855–10860.

(116)   Bowen, S.; Hilty, C. *Angew. Chem. Int. Ed.* **2008**, *47*, 5235–5237.

(117)   Jensen, P. R.; Meier, S.; Ardenkjær-Larsen, J. H.; Duus, J. Ø.; Karlsson, M.; Lerche, M. H. *Chem. Commun.* **2009**, 5168.

(118)   Zeng, H.; Lee, Y.; Hilty, C. *Anal. Chem.* **2010**, *82*, 8897–8902.

(119)   Chen, C.-H.; Shih, W.-C.; Hilty, C. *J. Am. Chem. Soc.* **2015**, *137*, 6965–6971.

(120)   Day, I. J.; Mitchell, J. C.; Snowden, M. J.; Davis, A. L. *J. Magn. Reson.* **2007**, *187*, 216–224.

(121)   Lee, Y.; Zeng, H.; Ruedisser, S.; Gossert, A. D.; Hilty, C. *J. Am. Chem. Soc.* **2012**, *134*, 17448–17451.

(122)   Kim, J.; Liu, M.; Chen, H.-Y.; Hilty, C. *Anal. Chem.* **2015**, *87*, 10982–10987.

(123)   Min, H.; Sekar, G.; Hilty, C. *ChemMedChem* **2015**, *10*, 1559–1563.

(124) Chappuis, Q.; Milani, J.; Vuichoud, B.; Bornet, A.; Gossert, A. D.; Bodenhausen, G.; Jannin, S. *J. Phys. Chem. Lett.* **2015**, *6*, 1674–1678.

(125) Chen, C.-H.; Wang, Y.; Hilty, C. *Methods* **2018**, *138-139*, 69–75.

(126) Frydman, L.; Scherf, T.; Lupulescu, A. *Proc. Natl. Acad. Sci.* **2002**, *99*, 15858–15862.

(127) Frydman, L.; Blazina, D. *Nat. Phys.* **2007**, *3*, 415–419.

(128) Zeng, H.; Bowen, S.; Hilty, C. *J. Magn. Reson.* **2009**, *199*, 159–165.

(129) Schanda, P.; Brutscher, B. *J. Am. Chem. Soc.* **2005**, *127*, 8014–8015.

(130) Schanda, P.; Kupče, Ē.; Brutscher, B. *J. Biomol. NMR* **2005**, *33*, 199–211.

(131) Szekely, O.; Olsen, G. L.; Felli, I. C.; Frydman, L. *Anal. Chem.* **2018**, *90*, 6169–6177.

(132) Kadeřávek, P.; Ferrage, F.; Bodenhausen, G.; Kurzbach, D. *Chem. Eur. J.* **2018**, *24*, 13418–13423.

(133) Kim, J.; Mandal, R.; Hilty, C. *J. Phys. Chem. Lett.* **2019**, *10*, 5463–5467.

(134) Mobli, M.; Hoch, J. C. *Prog. Nucl. Magn. Reson. Spectrosc.* **2014**, *83*, 21–41.

(135) Kupče, E.; Nishida, T.; Freeman, R. *Prog. Nucl. Magn. Reson. Spectrosc.* **2003**, *42*, 95–122.

(136) Chen, H.-Y.; Hilty, C. *Anal. Chem.* **2013**, *85*, 7385–7390.

(137) Lee, Y.; Zeng, H.; Mazur, A.; Wegstroth, M.; Carlomagno, T.; Reese, M.; Lee, D.; Becker, S.; Griesinger, C.; Hilty, C. *Angew. Chem. Int. Ed.* **2012**, *51*, 5179–5182.

(138) Sadet, A.; Stavarache, C.; Bacalum, M.; Radu, M.; Bodenhausen, G.; Kurzbach, D.; Vasos, P. R. *J. Am. Chem. Soc.* **2019**, *141*, 12448–12452.

(139) Harner, M. J.; Frank, A. O.; Fesik, S. W. *J. Biomol. NMR* **2013**, *56*, 65–75.

(140) Pellecchia, M.; Bertini, I.; Cowburn, D.; Dalvit, C.; Giralt, E.; Jahnke, W.; James, T. L.; Homans, S. W.; Kessler, H.; Luchinat, C.; Meyer, B.; Oschkinat, H.; Peng, J.; Schwalbe, H.; Siegal, G. *Nat. Rev. Drug Discov.* **2008**, *7*, 738–745.

(141) Becker, W.; Bhattiprolu, K. C.; Gubensäk, N.; Zangger, K. *ChemPhysChem* **2018**, *19*, 895–906.

(142) Meyer, B.; Peters, T. *Angew. Chem. Int. Ed.* **2003**, *42*, 864–890.

(143) Vanwetswinkel, S.; Heetebrij, R. J.; van Duynhoven, J.; Hollander, J. G.; Filippov, D. V.; Hajduk, P. J.; Siegal, G. *Chem. Biol.* **2005**, *12*, 207–216.

(144) Chen, D.; Errey, J. C.; Heitman, L. H.; Marshall, F. H.; IJzerman, A. P.; Siegal, G. *ACS Chem. Biol.* **2012**, *7*, 2064–2073.

(145) Congreve, M.; Rich, R. L.; Myszka, D. G.; Figaroa, F.; Siegal, G.; Marshall, F. H. In *Methods in Enzymology*; Elsevier: 2011; Vol. 493, pp 115–136.

(146) Ni, F. *Prog. Nucl. Magn. Reson. Spectrosc.* **1994**, *26*, 517–606.

(147) O'Connor, C.; White, K. L.; Doncescu, N.; Didenko, T.; Roth, B. L.; Czaplicki, G.; Stevens, R. C.; Wüthrich, K.; Milon, A. *Proc Natl Acad Sci USA* **2015**, *112*, 11852–11857.

(148) Wang, Y.; Ragavan, M.; Hilty, C. *J. Biomol. NMR* **2016**, *65*, 41–48.

(149) Janolino, V. G.; Fontecha, J.; Swaisgood, H. E. *Appl Biochem Biotechnol* **1996**, *56*, 1–7.

(150) Green, N. In *Methods in Enzymology*; Elsevier: 1970; Vol. 18, pp 418–424.

(151) Bowen, S.; Hilty, C. *Phys. Chem. Chem. Phys.* **2010**, *12*, 5766–5770.

(152) Solomon, I. *Phys. Rev.* **1955**, *99*, 559–565.

(153) Kovacs, H.; Moskau, D.; Spraul, M. *Prog. Nucl. Magn. Reson. Spectrosc.* **2005**, *46*, 131–155.

(154) Green, N. M. In *Advances in Protein Chemistry*; Elsevier: 1975; Vol. 29, pp 85–133.

(155) Määttä, J. A. E.; Airenne, T. T.; Nordlund, H. R.; Jänis, J.; Paldanius, T. A.; Vainiotalo, P.; Johnson, M. S.; Kulomaa, M. S.; Hytönen, V. P. *ChemBioChem* **2008**, *9*, 1124–1135.

(156) Macur, S.; Farmer, B.; Brown, L. *J. Magn. Reson.* **1986**, *70*, 493–499.

(157) Murali, N.; Jarori, G. K.; Landy, S. B.; Rao, B. D. N. *Biochemistry* **1993**, *32*, 12941–12948.

(158) Strzelczyk, P.; Bujacz, G. *J. Mol. Struct.* **2016**, *1109*, 232–238.

(159) Livnah, O.; Bayer, E. A.; Wilchek, M.; Sussman, J. L. *FEBS Lett.* **1993**, *328*, 165–168.

(160) Garrett, D. S.; Seok, Y. J.; Peterkofsky, A.; Gronenborn, A. M.; Clore, G. M. *Nat. Struct. Biol.* **1999**, *6*, 166–173.

(161) Zuiderweg, E. R. P. *Biochemistry* **2002**, *41*, 1–7.

(162) Clore, G. M.; Gronenborn, A. M. *J. Magn. Reson.* **1983**, *53*, 423–442.

(163) Campbell, A. P.; Sykes, B. D. *Annu. Rev. Biophys. Biomol. Struct.* **1993**, *22*, 99–122.

(164) Landon, C.; Berthault, P.; Vovelle, F.; Desvaux, H. *Protein Sci.* **2001**, *10*, 762–770.

(165) Kaptein, R.; Dijkstra, K.; Nicolay, K. *Nature* **1978**, *274*, 293–294.

(166) Olsen, G.; Markhasin, E.; Szekely, O.; Bretschneider, C.; Frydman, L. *J. Magn. Reson.* **2016**, *264*, 49–58.

(167)   Schnell, J. R.; Dyson, H. J.; Wright, P. E. *Annu. Rev. Biophys. Biomol. Struct.* **2004**, *33*, 119–140.

(168)   Dauber-Osguthorpe, P.; Roberts, V. A.; Osguthorpe, D. J.; Wolff, J.; Genest, M.; Hagler, A. T. *Proteins: Struct., Funct., Genet.* **1988**, *4*, 31–47.

(169)   Jolivet, J.; Cowan, K. H.; Curt, G. A.; Clendeninn, N. J.; Chabner, B. A. *N. Engl. J. Med.* **1983**, *309*, 1094–1104.

(170)   Abali, E. E.; Skacel, N. E.; Celikkaya, H.; Hsieh, Y. In *Vitamins & Hormones*; Folic Acid and Folates, Vol. 79; Academic Press: 2008, pp 267–292.

(171)   Liu, C. T.; Hanoian, P.; French, J. B.; Pringle, T. H.; Hammes-Schiffer, S.; Benkovic, S. J. *Proc. Natl. Acad. Sci.* **2013**, *110*, 10159–10164.

(172)   Matthews, D. A.; Alden, R. A.; Bolin, J. T.; Freer, S. T.; Hamlin, R.; Xuong, N.; Kraut, J.; Poe, M.; Williams, M.; Hoogsteen, K. *Science* **1977**, *197*, 452–455.

(173)   Filman, D. J.; Bolin, J. T.; Matthews, D. A.; Kraut, J. *J. Biol. Chem.* **1982**, *257*, 13663–13672.

(174)   Bystroff, C.; Oatley, S. J.; Kraut, J. *Biochemistry* **1990**, *29*, 3263–3277.

(175)   Bystroff, C.; Kraut, J. *Biochemistry* **1991**, *30*, 2227–2239.

(176)   Reyes, V. M.; Sawaya, M. R.; Brown, K. A.; Kraut, J. *Biochemistry* **1995**, *34*, 2710–2723.

(177)   Sawaya, M. R.; Kraut, J. *Biochemistry* **1997**, *36*, 586–603.

(178)   Falzone, C. J.; Benkovic, S. J.; Wright, P. E. *Biochemistry* **1990**, *29*, 9667–9677.

(179)   Falzone, C. J.; Cavanagh, J.; Cowart, M.; Palmer, A. G.; Matthews, C. R.; Benkovic, S. J.; Wright, P. E. *J. Biomol. NMR* **1994**, *4*, 349–366.

(180)   Marley, J.; Lu, M.; Bracken, C. *J. Biomol. NMR* **2001**, *20*, 71–75.

(181) Wagner, R.; Berger, S. *Fresenius J. Anal. Chem.* **1997**, *357*, 470–472.

(182) Otting, G.; Wüthrich, K. *J. Magn. Reson.* **1989**, *85*, 586–594.

(183) Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. *J. Comput. Chem.* **2004**, *25*, 1605–1612.

(184) Wang, Y.; Hilty, C. *J. Med. Chem.* **2019**, *62*, 2419–2427.

(185) Huang, S.-Y.; Zou, X. *Int. J. Mol. Sci.* **2010**, *11*, 3016–3034.

(186) Wang, B.; Westerhoff, L. M.; Merz, K. M. *J. Med. Chem.* **2007**, *50*, 5128–5134.

(187) Cozzini, P.; Kellogg, G. E.; Spyrakis, F.; Abraham, D. J.; Costantino, G.; Emerson, A.; Fanelli, F.; Gohlke, H.; Kuhn, L. A.; Morris, G. M.; Orozco, M.; Pertinhez, T. A.; Rizzi, M.; Sotriffer, C. A. *J. Med. Chem.* **2008**, *51*, 6237–6255.

(188) Lexa, K. W.; Carlson, H. A. *Q. Rev. Biophys.* **2012**, *45*, 301–343.

(189) Forli, S.; Huey, R.; Pique, M. E.; Sanner, M. F.; Goodsell, D. S.; Olson, A. J. *Nat. Protoc.* **2016**, *11*, 905–919.

(190) De Vivo, M.; Masetti, M.; Bottegoni, G.; Cavalli, A. *J. Med. Chem.* **2016**, *59*, 4035–4061.

(191) Lin, J.-H.; Perryman, A. L.; Schames, J. R.; McCammon, J. A. *Biopolymers* **2003**, *68*, 47–62.

(192) Bertini, I.; Luchinat, C.; Parigi, G. *Prog. Nucl. Magn. Reson. Spectrosc.* **2002**, *40*, 249–273.

(193) Constantine, K. L.; Davis, M. E.; Metzler, W. J.; Mueller, L.; Claus, B. L. *J. Am. Chem. Soc.* **2006**, *128*, 7252–7263.

(194) Hajduk, P. J.; Mack, J. C.; Olejniczak, E. T.; Park, C.; Dandliker, P. J.; Beutel, B. A. *J. Am. Chem. Soc.* **2004**, *126*, 2390–2398.

(195) Wüthrich, K., *NMR of Proteins and Nucleic Acids*, Google-Books-ID: zfBqAAAAMAAJ; Wiley: 1986.

(196) Wang, R.; Gao, Y.; Lai, L. *J. Mol. Model.* **2000**, *6*, 498–516.

(197) Gschwend, D. A.; Sirawaraporn, W.; Santi, D. V.; Kuntz, I. D. *Proteins: Struct., Funct., Genet.* **1997**, *29*, 59–67.

(198) Davies, J. F.; Delcamp, T. J.; Prendergast, N. J.; Ashford, V. A.; Freisheim, J. H.; Kraut, J. *Biochemistry* **1990**, *29*, 9467–9479.

(199) Gargaro, A. R.; Soteriou, A.; Frenkiel, T. A.; Bauer, C. J.; Birdsall, B.; Polshakov, V. I.; Barsukov, I. L.; Roberts, G. C.; Feeney, J. *J. Mol. Biol.* **1998**, *277*, 119–134.

(200) Feeney, J.; Birdsall, B.; Kovalevskaya, N. V.; Smurnyy, Y. D.; Navarro Peran, E. M.; Polshakov, V. I. *Biochemistry* **2011**, *50*, 3609–3620.

(201) Kurdzesau, F.; van den Brandt, B.; Comment, A.; Hautle, P.; Jannin, S.; van der Klink, J. J.; Konter, J. A. *J. Phys. D: Appl. Phys.* **2008**, *41*, 155506.

(202) Lipari, G.; Szabo, A. *J. Am. Chem. Soc.* **1982**, *104*, 4546–4559.

(203) Baleja, J. D.; Pon, R. T.; Sykes, B. D. *Biochemistry* **1990**, *29*, 4828–4839.

(204) Moseley, H. N. B.; Lee, W.; Arrowsmith, C. H.; Krishna, N. R. *Biochemistry* **1997**, *36*, 5293–5299.

(205) Palmer, A. G.; Case, D. A. *J. Am. Chem. Soc.* **1992**, *114*, 9059–9067.

(206) Dellwo, M. J.; Wand, A. J. *J. Am. Chem. Soc.* **1993**, *115*, 1886–1893.

(207) Ni, F.; Zhu, Y.; Scheraga, H. A. *J. Mol. Biol.* **1995**, *252*, 656–671.

(208) Zuiderweg, E. R. P.; Gestwicki, J. E. *Biomol. NMR Assign.* **2017**, *11*, 11–15.

(209) Bhunia, A.; Jayalakshmi, V.; Benie, A. J.; Schuster, O.; Kelm, S.; Rama Krishna, N.; Peters, T. *Carbohydr. Res.* **2004**, *339*, 259–267.

(210) Orts, J.; Wälti, M. A.; Marsh, M.; Vera, L.; Gossert, A. D.; Güntert, P.; Riek, R. *J. Am. Chem. Soc.* **2016**, *138*, 4393–4400.

(211) Wälti, M. A.; Riek, R.; Orts, J. *Angew. Chem. Int. Ed.* **2017**, *56*, 5208–5211.

(212) Kim, Y.; Liu, M.; Hilty, C. *J. Magn. Reson.* **2018**, *295*, 80–86.

(213) Takeuchi, K.; Wagner, G. *Curr. Opin. Struct. Biol.* **2006**, *16*, 109–117.

(214) Freedberg, D. I.; Selenko, P. *Annu. Rev. Biophys.* **2014**, *43*, 171–192.

(215) Claasen, B.; Axmann, M.; Meinecke, R.; Meyer, B. *J. Am. Chem. Soc.* **2005**, *127*, 916–919.

(216) Xie, J.; Thapa, R.; Reverdatto, S.; Burz, D. S.; Shekhtman, A. *J. Med. Chem.* **2009**, *52*, 3516–3522.

(217) Kumar, A.; Kuhn, L.; Balbach, J. *Int. J. Mol. Sci.* **2019**, *20*, 378.

(218) Otting, G.; Liepinsh, E.; Wuthrich, K. *Science* **1991**, *254*, 974–980.

(219) Otting, G.; Wuethrich, K. *J. Am. Chem. Soc.* **1989**, *111*, 1871–1875.

(220) Kim, J.; Liu, M.; Hilty, C. *J. Phys. Chem. B* **2017**, *121*, 6492–6498.

(221) De Vries, S. J.; van Dijk, M.; Bonvin, A. M. J. J. *Nat Protoc* **2010**, *5*, 883–897.

(222) Proudfoot, A.; Bussiere, D. E.; Lingel, A. *J. Am. Chem. Soc.* **2017**, *139*, 17824–17833.

(223) Marielle Wälti; Julien Orts *Magnetochemistry* **2018**, *4*, 12.