2019

# Nuclear magnetic resonance methodologies for the study of nanoparticle surface adsorbed ligands

Timothy Karl Egner
*Iowa State University*

**Nuclear magnetic resonance methodologies for the study of nanoparticle surface adsorbed ligands**

by

**Timothy Karl Egner**

A dissertation submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Chemistry

Program of Study Committee:
Vincenzo Venditti, Major Professor
Julien Roche
Aaron Rossini
Igor Slowing
Brett VanVeller

The student author, whose presentation of the scholarship herein was approved by the program of study committee, is solely responsible for the content of this dissertation. The Graduate College will ensure this dissertation is globally accessible and will not permit alterations after a degree is conferred.

Iowa State University

Ames, Iowa

2019

**TABLE OF CONTENTS**

Page

LIST OF FIGURES

Page

LIST OF TABLES

Page

## ACKNOWLEDGMENTS

I would like to thank my committee chair, Vincenzo Venditti, and my committee members, Igor Slowing, Aaron Rossini, Julien Roche, and Brett VanVeller, for their guidance and support throughout the course of this research.

In addition, I would also like to thank my friends, family, colleagues, for their support during my time as a student. I would especially like to thank the NMR facility staff Sarah Cady, Shu Xu, and Bruce Fulton.

Finally, I'd like to thank Adeline Boettcher for her constant support and assistance in my learning, research, problem solving… just about everything. Her help has been invaluable in preparing this dissertation.

## ABSTRACT

Nanomaterials have become increasingly important to study as their applications have expanded from heterogeneous catalysis, medicinal therapeutics, optoelectronics, and more. Understanding how these materials function is of utmost important to developing the next generation of functional nanoparticles. While there are various methods of characterizing nanoparticles using spectroscopic techniques such as infrared and Raman, solution state nuclear magnetic resonance has not seen the same type of methodological advancement.

Large molecular systems are difficult to study by NMR. Signal relaxation occurs faster as the molecule of interest has a slower solution tumbling rate. The result for the relatively large nanoparticle system are spectral lines that are broadened to the point of invisibility. Direct observation of spins is not helpful in these cases. Instead, surface contrast NMR experiments are proposed for the indirect detection of ligands that adsorb and desorb from a nanoparticle surface. The effect of signal broadening is used as an amplification factor to measure lowly populated surface bound ligand states.

An important aspect of performing NMR experiments on a sample that contains nanoparticles is that the nanoparticles must remain in suspension for the duration of the experiment. Nanoparticles typically are not soluble and tend to aggregate and sediment quickly. The use of the hydrogel, agarose, is proposed to trap nanoparticles within the pores of the gel to keep these materials homogeneously distributed for extended periods of time. We show that agarose is an inert matrix that does not participate in the ligand exchange process and does not interfere with data quality. The exchange processes of cholic acid and phenol onto the surface of ceria nanocubes were used as a model system. $^{1}$H dark state exchange saturation transfer (DEST)

and relaxation dispersion experiments were performed and data analyzed using the Bloch-McConnell equations. Cholic acid was found to be involved in a two site exchange process comprised of a surface bound and unbound state. Phenol was found to participate in a three site exchange; unbound, a weakly associated state characterized by a high degree of rotational freedom, and a tightly associated surface bound state characterized by dramatically decreased rotational freedom. Accessibility to the tightly associated state was found to be only allowed by first passing through the weakly associated intermediate state.

The hydrogenation of phenol can be achieved using $H_2$ and Pd supported on ceria under mild conditions with high conversion rates. However, the role of the ceria support, required in this catalytic system, is not fully understood. Ceria possesses exceptional redox properties but a full description of the interactions between phenol and ceria remains unknown. Here, we further our surface contrast NMR methodologies with the addition of $^{13}C$ relaxation ($R_1$ and $R_2$) and DEST measurements which allows for ligand dynamics modelling. We found that phenol interacts with the ceria surface in two binding modes consistent with hydrogen bonding and a covalent interaction with an oxygen defect site in the crystal lattice. As alluded to in our previous study, the tighter covalent interaction is only accessible by passing through the weak hydrogen bonding intermediate first. Additionally, phenol can bind to Pd in a flat conformation again passing through a hydrogen bonding to ceria intermediate. Adding phosphate into mixture perturbs the interactions between phenol hydrogen bonding to ceria but not to Pd. The corresponding decrease to the catalytic conversion suggests that the ability for phenol to hydrogen bond to the ceria surface plays an important role in this catalytic reaction.

The use of agarose gels in the previous studies described here have been extremely important for the successful, practical application of NMR experimentation to ligand-nanoparticle systems. However, agarose will only gel with an aqueous solvent which severely limits the scope of this methodology to systems that use a non-aqueous medium. Several gels that are compatible with organic solvents were characterized for their use in trapping nanoparticles from sedimentation. These included agarose, polystyrene, polydimethylsilicone, ((4,6-O-Benzylidene)-Methyl-$\alpha$-D-Glucopyranoside), methylcellulose, and polystearylacrylate. Six basic criteria were used to qualify important characteristics of the gels, namely low residual NMR signal, low internal viscosity, large pore size, macroscopic integrity, solvent compatibility, and ideal preparatory conditions. Overall, polystearylacrylate appears to be the most suitable for application to organic solvent based surface contrast NMR studies as it adequately satisfies all diagnostic criteria. Methylcellulose is another strong candidate, especially for its ease of preparation, but its application is limited as it will only gel using DMF as a solvent.

In addition to the project described above, I have collaborated on a number of other projects applying NMR to solve research questions. I have developed a software tool for fitting alignment tensors of proteins members in an ensemble to residual dipolar coupling data. This was used for defining a structural ensemble for AlkBH5 to better understand the origin of its catalytic activity and selectivity. I have characterized sugar composition of batches of alginate to aid in determining the chemical basis for its optical properties. This investigation was used to solve a reliability issue in the production of transparent soil. I have also contributed to characterizing a new ionic liquid formed from hydrogen sulfide and oleylamine. This ionic liquid was used as a 'green' sulfur precursor in the synthesis of metal sulfide nanoparticles.

# CHAPTER 1

## INTRODUCTION AND THESIS STRUCTURE

This thesis is composed of seven chapters based on manuscripts prepared by myself and collaborators, or collaborations to which I contributed. Chapter 2 is an introduction to the work described in the subsequent chapters. This chapter includes the theoretical basis for my work as well as the general plan of my investigations.

Chapter 3 is a review article published in Chemical Reviews that I coauthored with Vincenzo Venditti and Marius Clore on the use of NMR and X-ray scattering techniques for protein structural ensemble refinement.

Chapter 4 is a manuscript published in Angewante entitled "Mechanistic insights into nanoparticle surface adsorption by solution NMR spectroscopy". This study describes the use of an agarose gel to enable the study of fast sedimenting nanoparticle by NMR. Here we use $^{1}H$ DEST and relaxation dispersion measurements to characterize ligand adsorption/desorption onto nanoparticles.

Chapter 5 is a manuscript that is currently under review entitled " 'Surface contrast' NMR reveals non-innocent role of support in $Pd/CeO_2$ catalyzed phenol hydrogenation". This study expands our previous methodology for detecting ligand-nanoparticle interaction to also use $^{13}C$ NMR experiments. More sophisticated exchange and motional dynamics models were used to elucidate greater detail on the ligand surface interactions.

Chapter 6 is my contributions to an ongoing project focused on developing gel systems that are compatible with organic solvents.

Chapter 7 describes my contribution to several projects that I collaborated on and have been published.

Chapter 8 describes conclusions to the projects and manuscripts discussed throughout this thesis and some potential future directions related to these works.

# CHAPTER 2

## LITERATURE REVIEW

Scientific studies have substantially been advanced with the advent and use of spectroscopic techniques. One such technique, nuclear magnetic resonance (NMR), has provided deep chemical, biological, physical and medical insight towards investigating and understanding different scientific systems. NMR utilizes radio wave electromagnetic radiation to perturb nuclear spin contained within a strong magnetic field. This technology was conceived in the late 1940s and has been advancing steadily since accruing numerous Nobel prizes over time. NMR is implicated in a variety of different types of research, such as the analysis of chemical mixtures, understanding dynamic motions of proteins, and non-invasive medical imaging [1–3].

Indeed, even some recent noble prizes in chemistry and physics have gone towards the development of new NMR technology and experiments. As an example, Kurt Wüthrich won the Nobel prize in chemistry in 2002 for his contributions to 3D structure determination of proteins using NMR [4]. Even now there are many areas of research that may gain valuable insight from the application of NMR studies to these fields. In this dissertation, I will describe my application of NMR studies to nanoparticle systems especially in terms of new experiments and analysis, as well as furthering previous methods. Additionally, I will discuss important practical considerations and limitations of studying nanoparticle systems by NMR.

*Studying nanomaterials by NMR*

The study of nanomaterials is a relatively new field that is quickly gaining popularity for its variety of applications. Importantly, nanoparticles have been useful in heterogeneous catalysis[5], optoelectronics [6], and look to have a promising future as medicinal

nanotherapeutics [7]. Nanomaterials owe many of their attractive properties to their high surface area to volume ratio. Chemical reactions can be efficiently catalyzed on their surfaces, however there are different sets of properties accessible to these small materials, such as quantum effects [8].

Many catalytic processes are improved by using NPs. As an example, vanadia supported on mesoporus silica has been shown to selectively convert propane to propene without significant oxidation side products that are more prevalent in tradition methods [9]. Of course, only the surface participates in the reaction and the bulk material which is physically inaccessible contributes very little. By reducing the particle size, more surface per unit mass is exposed to process reactants which increases efficiency over bulk material. Also, NP catalysts are still easily separable from product mixtures by centrifugation just as larger bulk material. This is very unlike homogenous catalysts systems which may require complicated chemical separation and purification to recycle the catalyst.

NPs used in medical applications take advantage of the same type of properties, mainly surface to volume ratio, but have other considerations as well. Iron oxide nanoparticles are being considered for use as an iron supplement that is safe for human use and provides a long-term source of iron for patients with anemia [10]. Iron from the NP is slowly released into solution for use by cells in the body, predominately red blood cells. An important aspect of producing this material and making it safe for use in humans is developing a surface ligand coating that is compatible with cells and the immune system as well as keeping the material soluble. In this study, folic acid and chitosan based surface ligands are employed and dramatically improve the effectivity of the treatment in anemic rats [10].

In general, ligands transiently or permanently attached to NP surfaces play important roles in the materials function or application. Snelders et al has shown that by applying phosphine based ligands to a Rh nanoparticle catalyst, improved reductive selectively can be achieved [10]. Here reduction of an aromatic ring is preferred over carbonyl reduction. While these results are impressive, understanding the underlying mechanism would allow this type of process to be applied to other catalytic systems.

Studying mechanisms of interaction on the surface of nanoparticles has been approached by techniques such as IR, RAMAN, MD simulations, and ssNMR. IR and RAMAN can provide important information on ligand binding modes and vibrations and are relatively easy to implement in the case of high surface coverage [11]. Special experimental setup with these techniques allow for enhanced signal and therefore better detection of surface adsorbed ligands. However, there are limitations to the type of information acquired by these methods; particularly, kinetic and thermodynamic information is difficult to acquire. In addition, weak ligand interaction may be entirely undetectable because of the dilute quantity of bound ligands.

Molecular dynamics simulation and density function theory (DFT) can provide useful information as well. A computational study on ligand structure on Au nanoparticle surfaces has shown the dynamic behavior of ligand chains differ as a function of temperature and length [12]. Care must be taken to interpret computational results in the absence of experimental validation. MD simulations can provide atomic level resolution of these systems, however, the quality of these results may depend immensely on the force field and/or approximations used.

Solid state NMR, which has access to significantly larger molecular system sizes than traditional solution state NMR, is an option for studying NPs. Typically, both the ligand and the

NP itself can be observed by ssNMR which gives an extra level detail using this technique. Davidowski and Holland used ssNMR to characterize binding interactions of various phophonic acids covalently attached onto silica [13]. They were able to obtain some differentiation between mono-, bi-, and tridentate ligand binding as well as the average packing density of the surface ligands.

Another chemical technique that has recently been employed to study NP systems is NMR. NMR can yield very detailed information about chemical systems and proves to be useful in some NP systems. The Mancin group has shown in a few publications that nanoparticle receptors can be used with NMR to reveal he presence of specific chemicals in solution [14]. Here, they take advantage of spin diffusion via saturation transfer difference experiment to detect if a ligand with high affinity to the receptor is present. In 2016, Salorinne et al performed an extensive and complete assignment of 4-MBA ligands on $Au_{108}$ NPs [15]. This system is quite unique in the fact that there seems to be only a single constitutional isomer of this ligand capped NP. It has previously been crystallized and its structure determined by x-ray diffraction. Through the combined use of NMR measurements and MD simulation, the authors were able to characterize this ligands coating structure and motional properties. Similarly, a study from 1998 characterized the structural and dynamic aspects of thiophenol on CdS NPs [16]. Here, relaxation experiments were employed to differentiate potential binding modes by differences in their allowed motions.

An important limiting factor between these and many other NMR NP studies is the size of the nanoparticle system. Solution NMR, specifically relaxation of spins, relies on relatively fast solution state tumbling for signals to be observable. For example, bulk water at 25°C has a

correlation time of approximately 1.7ps and the observed line width is narrow [17]. As the tumbling rate slows, the NMR signal becomes broader due to an increased relaxation rate (transverse relaxation). As an example, the tumbling rate could be slowed by cooling the solution or adding a thickening agent such as glycerol. The result is a coherent NMR signal that dephases more quickly and the Fourier transform of which is a broader signal. Large molecular systems also exhibit slow solution tumbling. Ligands attached to small nanoparticles have NMR signals that are significantly broadened out and if the NP is large enough, which are completely unobservable. The above studies use nanoparticles that are on the order of 1-3nm in diameter which is small enough to still observe NMR signals. However, the methods utilized in the above studies do not generalize to larger NP systems as they rely on the bound ligand having observable signals.

*Surface Contrast NMR*

Here I describe a new method of measuring ligand-NP interactions that generalize to significantly larger systems. In fact, many aspects of these experiments become more powerful as the size of the system increases. I will now describe what we are calling the Surface Contrast NMR (scNMR) experiments that take advantage of the increased relaxation rates of a ligand bound to a nanoparticle surface.

There are two basic requirements of the scNMR experiments: [1] there must be chemical exchange between a free and surface bound form of the ligand of interest and [2] there must be a difference in rotational correlation time between these distinct chemical states. While the ligand is in the free state, its correlation time is relatively fast, commensurate with free solution tumbling. When ligands are attached to a NP surface, the correlation time increases due to the

slow tumbling of the NP. In the cases of rigid or flexible surface attachment, the tumbling rate may be modulated by local motion. The effect of the slow tumbling in the surface bound state will affect the measured data as this "information" is transferred back to the free (detectable) state. The exact effect on the measured spectrum is dependent on a multitude of exchange parameters and intrinsic properties specific to each state of the exchange.

In characterizing an exchange or equilibrium process there are several important parameters that describe the system in terms of kinetic and thermodynamic properties. For kinetics, the adsorption and desorption rate constants describe the ensemble average forward and reverse rates of exchange. For a molecule interacting with a nanoparticle, the on rate, $k_{on}$, is a first order process dependent on the concentration of ligand and the off rate, $k_{off}$, is a pseudo zero order process. The thermodynamics describe where the balance in the equilibrium lies, namely the bound and unbound fractional population of the ligands. Here, the bound population $p_B$ is the percentage of ligand interacting with the nanoparticle surface and the unbound population $p_A$, is $1$-$p_B$. These four parameters are related by the equation:

$$p_A * k_{on}^{App} = p_B * k_{off}$$

and determining $p_B$ and $k_{off}$ is sufficient to solve for the remaining parameters. It is important to note that only the $k_{on}^{App}$ can be determined through the experimentation and analysis described below. Information about the concentration and stoichiometry of the nanoparticle binding sites would be required to determine $k_{on}$. This can be difficult, to near impossible, when considering nanoparticle with an indeterminate number of binding sites and distribution of particle dimensions.

Additionally, the techniques described below can yield information on the structural and dynamic properties of the system under investigation. The relaxation enhancement effect of the ligand in the nanoparticle bound state is determined by the overall effective motion of the ligand. As mentioned previously, this includes the tumbling rate of the nanoparticle itself and the local motion of the ligand. These relaxation rates can be quantitatively compared to theoretical simulations of relaxation to calculate correlation times and amplitude of motion for both the nanoparticle and the surface attached ligand.

Structural information may also be present in NMR measurements. As the bound state ligand signals cannot be directly observed, indirect information about chemicals shift can be used to characterize the binding mode. Changes in electron density around a given nuclei will result in a change in the chemical shift which is reported on in several experiments that are used below. These data can be interpreted qualitatively based on the sign of the change to infer structural changes. Additionally, DFT calculation could be used to calculate chemical shift to compare with experimental data, however this method is not used here. Finally, the allowed dynamic motions of a ligand may be dependent on its binding mode. Using the various models to characterize the bound state dynamics may also allow for hypothesis testing for binding mode determination.

The analysis of quantitative NMR data can be done through iterative optimization of a simulated NMR experiment. All the above parameters contribute to various aspects of the appearance of the spectrum and a convenient mode of analysis is simulation using the Bloch equation:

$$\frac{d\boldsymbol{M_{xy}}}{dt} = \gamma\left(\boldsymbol{M_{xy}}(t)\times\boldsymbol{B_{xy}}(t)\right) - \boldsymbol{R_2}\,\boldsymbol{M_{xy}}(t)$$

$$\frac{d\boldsymbol{M_z}}{dt} = \gamma\left(\boldsymbol{M_z}(t)\times\boldsymbol{B_z}(t)\right) - \boldsymbol{R_1}\,(\boldsymbol{M_z}(t) - \boldsymbol{M_z}(0))$$

where **M**(t) is the magnetization of a set of spins in the x, y, and z directions, **B**(t) is the external field in the x,y,z directions, $\gamma$ is the gyromagnetic proportionality constant that is nuclei specific, and $R_{1,2}$ are the relaxation matrices. Performing numerical integration to solve this differential equation gives the time domain free induction decay, characteristic to NMR experiments. This can be Fourier transformed to result in the simulated spectrum of the spin system of interest. The above equations can be modified and/or simplified depending on the requirements of the desired experiment.

Additionally, these expressions can be modified to include chemical exchange as shown here:

$$\frac{d\boldsymbol{M_{xy}}}{dt} = \gamma\left(\boldsymbol{M_{xy}}(t)\times\boldsymbol{B_{xy}}(t)\right) - \boldsymbol{R_2}\,\boldsymbol{M_{xy}}(t) + \boldsymbol{k}\,\boldsymbol{M_{xy}}(t)$$

$$\frac{d\boldsymbol{M_z}}{dt} = \gamma\left(\boldsymbol{M_z}(t)\times\boldsymbol{B_z}(t)\right) - \boldsymbol{R_1}\left(\boldsymbol{M_z}(t) - \boldsymbol{M_z}(0)\right) + \boldsymbol{k}\,(\boldsymbol{M_z}(t) - \boldsymbol{M_z}(0))$$

Here, **k** is the exchange matrix that contains the rate constants, $k_{ij}$, that describe the rate of exchange between each pair of states *ij*. The above equation shows the n-site exchange formulation which can be made specific for the exchange model of choice.

Relaxation modeling is very important in simulating data for a ligand-NP system. Here, the relaxation rates will be different for each state of the exchange and depend on multiple factors. Briefly, I will discuss relaxation mechanisms especially in terms of dynamic motion to illustrate how these parameters will be modelled.

In solution NMR, there are two primary mechanisms of relaxation that need to be considered: dipolar interactions and chemical shift anisotropy. Relaxation due to dipolar

interactions is caused by variation in local magnetic field strength due to nearby magnetic dipoles. This interaction is a second rank interaction and therefore dependent on the angle between the interacting dipoles, relative to the applied field, and the distance between the dipoles. For small molecules, this is the dominant relaxation mechanism. Chemical shift anisotropy (CSA) causes relaxation by variation in local magnetic field due to an anisotropic distribution of electron density around the spin of interest. This relaxation mechanism contributes significantly less to the overall relaxation process. For a $^{13}C$ spin with a directly attached $^{1}H$ with a CSA of approximately 200ppm, CSA contributes roughly 5% to the overall transverse relaxation process. While this is a small contribution, it is large enough that it must be taken into account. For spins with significantly smaller CSA (<20ppm) this term could be neglected.

Relaxation rates can be calculated as a function of rotational correlation time and field strength as follows:

$$R_1 = \frac{d^2}{4}\left(J(\omega_H - \omega_C) + 3J(\omega_C) + 6J(\omega_C + \omega_H)\right) + c^2 J(\omega_C)$$

$$R_2 = \frac{d^2}{8}\left(4J(0) + J(\omega_H - \omega_C) + 3J(\omega_C) + 6J(\omega_H) + 6J(\omega_C + \omega_H)\right) + \frac{c^2}{6}\left(4J(0) + 3J(\omega_C)\right)$$

Where:

$$d = \frac{\mu_0}{4\pi}\hbar\gamma_C\gamma_H r^{-3}$$

$$c = \frac{\gamma_C B_0 \Delta\sigma}{\sqrt{3}}$$

$$J(\omega) = \frac{2}{5}\frac{\tau_C}{1 + \tau_C^2\omega^2}$$

Where, $\mu_0$ is the magnetic permittivity, $\hbar$ is the reduced Planck's constant, $\gamma_u$ is the gyromagnetic ratio of the specified nuclei, *r* is the internuclear bond distance (taken to be 1.09 Å here), *J(ω)* is the spectral density function, $\tau_C$ is the rotational correlation time, and Δσ is the chemical shift anisotropy. The above formulation for *J(ω)* is the case for isotropic global motion. The first term corresponds to relaxation due to dipolar interactions and the second term is for CSA. *J(ω)* is known as the spectral density function and is the link between rotational motion and relaxation rates. For simple isotropic tumbling, *J(ω)* increases as $\tau_C$ increases until $(\tau_C\omega)^2 > 1$ at which point it decreases again for non-zero *ω*.

This expression for *J(ω)* can be adjusted to include global and local motions if the local motion is much faster than the global motion [18]. This is the model free formulation of the spectral density function as it does not make any assumption on what motion is present.

$$J(\omega) = \frac{2}{5}S^2\left[\frac{\tau_{NP}}{1+\tau_{NP}^2\omega^2}\right] + \frac{2}{5}(1-S^2)\left[\frac{\tau_e}{1+\tau_e^2\omega^2}\right]$$

Here, the parameter $S^2$, referred to as the generalized order parameter, is included to describe a weighed sum of local verses global motion. This is loosely interpreted as the amplitude of the local motion where an order parameter of 1 equates to a perfectly rigid local bond vector and $S^2$=0 would be completely flexible (180° reorientation).

As the technique is advanced (Chapter 5), anisotropic motion will be considered for specific ligand binding modes. Here, it is necessary to make assumptions on how the rotational motion occurs. In this case, instead of using $S^2$ as a general indication of motion, it can be substituted for other expressions specific to certain dynamics. Here is an example of where there is a specific axis of rotation present in a molecule and the principle axis system of the relaxation vectors are at an angle $\theta_{M\text{-}PAS}$ relative to said axis:

$$J(\omega) = \frac{2}{5}[Y_2^0(\theta)]^2 \left[\frac{\tau_{NP}}{1+\tau_{NP}^2\omega^2}\right] + \frac{2}{5}[Y_2^1(\theta)]^2 \left[\frac{\tau_e}{1+\tau_e^2\omega^2}\right]$$

The derivation of this expression can be found in Sachleben et al (1998) [15].

The above theory can only be utilized properly if enough data is acquired to satisfy all independent parameters. The number of experiments and type of data needed may depend on attributes of the system under study, and thusly would be safer and more general to acquire more data wherever possible. There are a few experiments that will give the type of needed information for analysis using the previously described theory and have been used in a similar way for studies in proteins. Here, the DEST, RD, and $R_1$ NMR experiments will be used extensively to provide the necessary experimental data for comparison against simulations. Each experiment will be explained in detail.

*Dark State Exchange Saturation Transfer*

The DEST (Dark state Exchange Saturation Transfer) experiment derives itself from a related experiment called CEST (Chemical Exchange Saturation Transfer). In either of these experiments, continuous wave irradiation is applied to a region of the NMR spectrum to saturate any spins that resonate at that frequency [19–21]. A hard excitation pulse is applied and data is acquired. Any spins that are saturated will not give signal, or will be attenuated, and unaffected signals remain at a constant intensity. The experiment is repeated such that saturation is scanned through the spectral regions of interest. In the absence of chemical exchange, the information provided by this experiment is trivial. However, in the presence of exchange, spins that participate in two or more NMR distinct environments can transfer saturation from one state to another under the proper exchange conditions. In this way, irradiation may be applied at a given signal and both that signal and at least one other will decrease in intensity.

The primary difference between CEST and DEST is the power of the applied field for saturation (typically 1-50Hz for CEST, 150-1000Hz for DEST). In the case of a CEST experiment, either state of the chemical exchange will likely appear as a sharp signal in the NMR spectrum. Higher saturation fields are suitable for nanoparticle studies as the bound state signal is expected to be broad and the higher field provides a wider bandwidth for saturation. A side effect of this is shorter experiment times as the number of saturation frequencies needed are lower for a given spectral width.

The BM equations can be modified to account for the application of a saturating field as in the DEST experiment:

$$\frac{d}{dt}\begin{bmatrix} E/2 \\ I_x^A \\ I_y^A \\ I_z^A \\ I_x^B \\ I_y^B \\ I_z^B \\ I_x^C \\ I_y^C \\ I_z^C \end{bmatrix} = -\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & R_2^{A*} & \Omega^A & -\omega_y & -k_{BA} & 0 & 0 & -k_{CA} & 0 & 0 \\ 0 & -\Omega^A & R_2^{A*} & \omega_x & 0 & -k_{BA} & 0 & 0 & -k_{CA} & 0 \\ -2\Theta^A & \omega_y & -\omega_x & R_1^{A*} & 0 & 0 & -k_{BA} & 0 & 0 & -k_{CA} \\ 0 & -k_{AB}^{app} & 0 & 0 & R_2^{B*} & \Omega^B & -\omega_y & -k_{CB} & 0 & 0 \\ 0 & 0 & -k_{AB}^{app} & 0 & -\Omega^B & R_2^{B*} & \omega_x & 0 & -k_{CB} & 0 \\ -2\Theta^B & 0 & 0 & -k_{AB}^{app} & \omega_y & -\omega_x & R_1^{B*} & 0 & 0 & -k_{CB} \\ 0 & -k_{AC}^{app} & 0 & 0 & -k_{BC} & 0 & 0 & R_2^{C*} & \Omega^C & -\omega_y \\ 0 & 0 & -k_{AC}^{app} & 0 & 0 & -k_{BC} & 0 & -\Omega^C & R_2^{C*} & \omega_x \\ -2\Theta^C & 0 & 0 & -k_{AC}^{app} & 0 & 0 & -k_{BC} & \omega_y & -\omega_x & R_1^{C*} \end{bmatrix}$$

Here, $\Theta^p = \mathrm{R}_1^p I_z^p$, $R_2^{p*} = R_2^p + k_{pq} + k_{pr}$, $R_1^{p*} = R_1^p + k_{pq} + k_{pr}$, $\Omega^p$is the difference between the resonance frequency of state *p* and the applied saturation field, and $\omega_{y,x}$is the continuous wave saturation field strength on the applied axis. The above equation shows the formulation for a three-site exchange process. This equation is numerically integrated and the first point of the simulated FID is used as a proxy for the amplitude of the signal of interest.

The DEST experiment provides both kinetic and thermodynamic information. However, only specific values for the kinetics become relevant if the exchange process occurs on the seconds to milliseconds timescale. Exchange processes, slower the s to ms range may not

transfer saturation between states as signal relaxation dominates. Exchange faster than this timescale typically is in the fast exchange regime and only shows a single signal in the NMR spectrum. Useful information can still be obtained in the case of vastly different relaxation rates, even at the same chemical shift as in the situation of a nanoparticle bound / free ligand exchange. The kinetic information is limited to a lower bound on the exchange rate constants.

*Relaxation Dispersion - $R_2$*

The transverse relaxation rate is a measure of the time it takes for coherent spins to dephase. There are a few different mechanisms that result in signal relaxation, but dipolar interactions and chemical shift anisotropy are the only significant contributors. $R_2$ can be measured, commonly, using a pulse sequenced called CPMG [22,23]. The spin of interest is prepared in the xy-plane and a train of π-pulses are applied to the spins for a set period. The signal is then measured, and its intensity compared to a reference experiment with the omission of the relaxation delay (which includes the pi-pulse train). The natural log of the decrease in signal intensity is related to the transverse relaxation rate. It is important to note that the repetition rate of the π-pulses influences the final result in the case of chemical exchange occurring at the μs-ms timescale. This variable component of $R_2$ is referred to as the exchange contribution to $R_2$, or $R_{ex}$.

As mentioned previously, $R_2$ is very sensitive to the tumbling rate of spin of interest and reports well on the thermodynamics of exchange in the case of two NMR distinct states with different correlation times. Additionally, relaxation dispersion measurements (where multiple repetition rate of the π-pulse train are used and $R_2$ is measured for each) report on kinetics and

chemical shift information about the system as well. The BM equation that describes this experiment is shown here for a two-site exchange:

$$M(t) = (AA^*A^*A)^n M(0)$$

$$M(0) = \begin{bmatrix} M^A \\ M^B \\ M^C \end{bmatrix}$$

$$A = e^{-R\tau_{CP}/2}$$

$$R = iR^{CS} + R^{rel} + R^{ex}$$

$$R^{CS} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -\Delta\omega^{AB} & 0 \\ 0 & 0 & -\Delta\omega^{AC} \end{bmatrix}$$

$$R^{rel} = \begin{bmatrix} R_2^A & 0 & 0 \\ 0 & R_2^B & 0 \\ 0 & 0 & R_2^C \end{bmatrix}$$

$$R^{ex} = \begin{bmatrix} k_{AB}^{app} + k_{AC}^{app} & -k_{BA} & -k_{CA} \\ -k_{AB}^{app} & k_{BA} + k_{BC} & -k_{CB} \\ -k_{AC}^{app} & -k_{BC} & k_{CB} + k_{CA} \end{bmatrix}$$

The definition of variables here are the same as for the DEST experiment with the addition of $\Delta\omega^{ij}$ which is the difference in frequency (rad $s^{-1}$) between the *i*th and *j*th state and $\tau_{CP}$ is the time delay between two π pulses during the CPMG pulse train. The calculated $R_2$ value can be fit from the simulated signal decay or simply:

$$R_2^{calc} = \frac{\ln(M(t_1)/M(t_2))}{t_2 - t_1}$$

Importantly, $R_2$ is dependent on the rate of exchange, the chemical shift difference between the two states, and the applied magnetic field. Kinetic exchange information on the µs-ms timescale is reported on with this experiment. To ensure that the final fits for this type of data are non-

degenerate, this experiment is usually performed at two (or more) different magnetic field strengths. We have used this approach in out studies.

*Longitudinal Relaxation - $R_1$*

The longitudinal relaxation rate is the rate constant that describes the transition of an excited spin in the xy-plane relaxing to its equilibrium position along the z-axis. This rate constant is most commonly measured in an inversion recovery experiment. Magnetization is prepared on the nuclei of interest in the z direction and then a π-pulse is applied to flip the magnetization towards –z. The magnetization relaxes towards +z during a delay period and then the remaining intensity is probed through excitation (to the xy-plane) with a π/2 pulse and detected. By varying the length of the delay period, the rate of longitudinal relaxation can be measured.

The information contained in $R_1$ for a system under exchange is primarily restricted to thermodynamic properties as kinetic parameters control relatively weak changes in $R_1$. Unlike $R_2$, $R_1$ is not affected by changes in chemical shift. This is quite useful because there are fewer parameters required to simulate this experiment. Including $R_1$ measurements decreases the degeneracy of the fitting process. Additionally, $R_1$'s trend as a function of correlation time is different to that of $R_2$, especially in the case of $\omega\tau_C > 1$. For a two-site exchange scheme, data may be fit uniquely using only $R_1$ and $R_2$ measurements.

The BM formulation for simulation of an $R_1$ measurement is shown below:

$$R_1^{calc} = \frac{\ln(M(t_1)/M(t_2))}{t_2 - t_1}$$

$$M(t) = (AA^*A^*A)^n M(0)$$

$$M(0) = \begin{bmatrix} M^A \\ M^B \\ M^C \end{bmatrix}$$

$$A = e^{-tR}$$

$$R = R^{rel} + R^{ex}$$

$$R^{rel} = \begin{bmatrix} R_1^A & 0 & 0 \\ 0 & R_1^B & 0 \\ 0 & 0 & R_1^C \end{bmatrix}$$

$$R^{ex} = \begin{bmatrix} k_{AB}^{app} + k_{AC}^{app} & -k_{BA} & -k_{CA} \\ -k_{AB}^{app} & k_{BA} + k_{BC} & -k_{CB} \\ -k_{AC}^{app} & -k_{BC} & k_{CB} + k_{CA} \end{bmatrix}$$

This expression is similar to the formulation for the relaxation dispersion experiment with the lack of chemical shift dependence.

*Combined Use of DEST, RD, $R_1$*

Each of the above experiments provide detail towards explaining chemical exchange processes. Frequently, when these experiments or similar are used to investigate other systems (often proteins), assumptions are made about relaxation rates. For example, a protein with a loop that exhibits dynamic motion, a common assumption is that the transverse relaxation rate is the same for all states of the exchange [24]. This is certainly not true in the case of nanoparticle binding. These same types of assumptions cannot be made here and so, more data on the system must be collected to accurately describe it. The use of these two or more of experiments in conjunction during data fitting serves to break the degeneracy that might otherwise be an issue. Additionally, because of the various timescales of exchange (µs-ms for RD, ms-s for DEST) that are surveyed in these measurements, the combined approach allows for more complete detection of exchange processes.

*Sample preparation*

Previously, NP studies that utilized NMR had focused on NP systems that are soluble in a solvent of interest. However, many NPs are not soluble, but only suspendable. Depending on the lifetime of the suspension it may become impossible to perform the proper NMR analysis on that system due to the concentration changing as the particle sediment. NPs can be functionalized to increase their solubility, but if that is not their natively catalytic form then it would be undesirable to do so. Ideally, nanoparticles should be able to be held in suspension for an indefinite period to allow for sufficient time to complete the prescribed experiments. One solution would be to trap NPs in the pores of gels to prevent them from settling. While gels can have a relatively solid macroscopic structure, the solvent in the pore spaces still resemble a liquid. Later chapters (chapter 6) will explore the utilities of gels and important considerations of potential sources of errors due to a gel being present during data measurement. Additionally, a few select gel types are reported on for their utility with common laboratory solvents to expand the repertoire of this methodology.

*Model system for study*

The methodology for studying ligand nanoparticle interaction by solution NMR as described above will be tested on a cerium dioxide nanoparticle model system that participates in the catalytic reduction of phenol. Phenol, primarily derived from crude oil, can be reduced to cyclohexanone and cyclohexanol using $H_2$ gas for use as a feedstock for various purposes ranging from drug synthesis to plastics manufacturing. This process is performed catalytically in industrial settings but typically requires high pressures of $H^2$ and high temperatures. It was recently found by Nelson et al (2015) [25] that this reduction reaction can be performed at room temperature

with ambient pressures of $H^2$ using Pd supported on ceria (Pd/ceria). Additionally, this catalytic system shows high conversion and selectivity for cyclohexanone. There are various hypotheses for what may cause this activity, however it is not fully understood.

By applying NMR techniques to investigate this reaction system, we aim to uncover important mechanistic details of this nanoparticle catalyst. Indeed, we characterized absorption and desorption exchange of phenol ligands on the ceria surface to surprising detail. This includes measuring amounts of phenol hydrogen bonding to the ceria surface verses amounts that interact with oxygen defect sites. We have been able to perturb reaction rates with the addition of inhibitors and rationalize these perturbations in terms of binding data.

**References**


1. Fan TW-M, Lane AN. Applications of NMR spectroscopy to systems biochemistry. Prog Nucl Magn Reson Spectrosc. 2016/02/06. 2016;92–93: 18–53. doi:10.1016/j.pnmrs.2016.01.005

2. Forseth RR, Schroeder FC. NMR-spectroscopic analysis of mixtures: from structure to function. Curr Opin Chem Biol. 2010/11/09. 2011;15: 38–47. doi:10.1016/j.cbpa.2010.10.010

3. Karakoulia SA, Triantafyllidis KS, Tsilomelekis G, Boghosian S, Lemonidou AA. Propane oxidative dehydrogenation over vanadia catalysts supported on mesoporous silicas with varying pore structure and size. Catal Today. 2009;141: 245–253. doi:https://doi.org/10.1016/j.cattod.2008.05.024

4. Palmer AG, Patel DJ. Kurt Wuthrich and NMR of biological macromolecules. Structure. 2002;10: 1603–1604. doi:10.1016/s0969-2126(02)00915-2

5. Xia Y, Yang H, Campbell CT. Nanoparticles for Catalysis. Acc Chem Res. 2013;46: 1671–1672. doi:10.1021/ar400148q

6. Podborska A, Oszajca M, Gawęda S, Szaciłowski K. Nanoparticles with logic and numeracy : towards “computer-on-a-particle” optoelectronic devices. IET Circuits, Devices & Syst. 2011;5: 103–114. doi:10.1049/iet-cds.2010.0068

7. Murthy SK. Nanoparticles in modern medicine: state of the art and future challenges. Int J Nanomedicine. 2007;2: 129–141. Available: https://www.ncbi.nlm.nih.gov/pubmed/17722542

8. Bera D, Qian L, Tseng T-K, Holloway PH. Quantum Dots and Their Multimodal Applications: A Review. Materials (Basel). 2010;3: 2260–2345. doi:10.3390/ma3042260

9. Hashem F, Nasr M, Ahmed Y. PREPARATION AND EVALUATION OF IRON OXIDE NANOPARTICLES FOR TREATMENT OF IRON DEFICIENCY ANEMIA. Int J Pharm Pharm Sci. 2018;10: 142. doi:10.22159/ijpps.2018v10i1.22686

10. Snelders DJM, Yan N, Gan W, Laurenczy G, Dyson PJ. Tuning the Chemoselectivity of Rh Nanoparticle Catalysts by Site-Selective Poisoning with Phosphine Ligands: The Hydrogenation of Functionalized Aromatic Compounds. ACS Catal. 2012;2: 201–207. doi:10.1021/cs200575r

11. Zhang D, Ansar SM. Ratiometric Surface Enhanced Raman Quantification of Ligand Adsorption onto a Gold Nanoparticle. Anal Chem. 2010;82: 5910–5914. doi:10.1021/ac1010124

12. Ghorai PK, Glotzer SC. Molecular Dynamics Simulation Study of Self-Assembled Monolayers of Alkanethiol Surfactants on Spherical Gold Nanoparticles. J Phys Chem C. 2007;111: 15857–15862. doi:10.1021/jp0746289

13. S. K. Davidowski and G. P. Holland. Solid-State NMR Characterization of Mixed Phosphonic Acid Ligand Binding and Organization on Silica Nanoparticles. Langmuir. 2016;32: 3253-3261. doi:10.1021/acs.langmuir.5b03933

14. Gabrielli L, Rosa-Gastaldo D, Salvia M-V, Springhetti S, Rastrelli F, Mancin F. Detection and identification of designer drugs by nanoparticle-based NMR chemosensing. Chem Sci. 2018;9: 4777–4784. doi:10.1039/C8SC01283K

15. Salorinne K, Malola S, Wong OA, Rithner CD, Chen X, Ackerson CJ, et al. Conformation and dynamics of the ligand shell of a water-soluble Au102 nanoparticle. Nat Commun. 2016;7: 10401. doi:10.1038/ncomms10401

16. Sachleben JR, Colvin V, Emsley L, Wooten EW, Alivisatos AP. Solution-State NMR Studies of the Surface Structure and Dynamics of Semiconductor Nanocrystals. J Phys Chem B. 1998;102: 10117–10128. doi:10.1021/jp9820121

17. Lankhorst D, Schriever J, Leyte JC. Determination of the Rotational Correlation Time of Water by Proton NMR Relaxation in H217O and Some Related Results. Berichte der Bunsengesellschaft für Phys Chemie. 1982;86: 215–221. doi:10.1002/bbpc.19820860308

18. Lipari G, Szabo A. Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 1. Theory and range of validity. J Am Chem Soc. 1982;104: 4546–4559. doi:10.1021/ja00381a009

19. Ceccon A, Tugarinov V, Clore GM. TiO2 Nanoparticles Catalyze Oxidation of Huntingtin Exon 1-Derived Peptides Impeding Aggregation: A Quantitative NMR Study of Binding and Kinetics. J Am Chem Soc. 2019;141: 94–97. doi:10.1021/jacs.8b11441

20. Anthis NJ, Clore GM. Visualizing transient dark states by NMR spectroscopy. Q Rev Biophys. 2015;48: 35–116. doi:10.1017/S0033583514000122

21. Fawzi NL, Libich DS, Ying J, Tugarinov V, Clore GM. Characterizing methyl-bearing side chain contacts and dynamics mediating amyloid beta protofibril interactions using (1)(3)C(methyl)-DEST and lifetime line broadening. Angew Chem Int Ed Engl. 2014;53: 10345–10349. doi:10.1002/anie.201405180

22. Kleckner IR, Foster MP. An introduction to NMR-based approaches for measuring protein dynamics. Biochim Biophys Acta. 2010/11/06. 2011;1814: 942–968. doi:10.1016/j.bbapap.2010.10.012

23. Karamanos TK, Tugarinov V, Clore GM. Unraveling the structure and dynamics of the human DNAJB6b chaperone by NMR reveals insights into Hsp40-mediated proteostasis. Proc Natl Acad Sci. 2019;116: 21529 LP – 21538. doi:10.1073/pnas.1914999116

24. Venditti V, Tugarinov V, Schwieters CD, Grishaev A, Clore GM. Large interdomain rearrangement triggered by suppression of micro- to millisecond dynamics in bacterial Enzyme I. Nat Commun. 2015;6: 5960. doi:10.1038/ncomms6960

25. Nelson NC, Manzano JS, Sadow AD, Overbury SH, Slowing II. Selective Hydrogenation of Phenol Catalyzed by Palladium on High-Surface-Area Ceria at Room Temperature and Ambient Pressure. ACS Catal. 2015;5: 2051–2061. doi:10.1021/cs502000j

## CHAPTER 3

# HYBRID APPROACHES TO STRUCTURAL CHARACTERIZATION OF CONFORMATIONAL EMSEMBLES OF COMPLEX MACROMOLECULAR SYSTEMS COMBINING NMR RESIDUAL DIPOLAR COUPLINGS AND SOLUTION X-RAY SCATTERING

Published in *Chemical Reviews* in 2016.

Vincenzo Venditti[1,2] Timothy Egner[1], and G. Marius Clore[3]

[1]Department of Chemistry, Iowa State University
[2]Roy J. Carver Department of Biochemistry, Biophysics, and Molecular Biology, Iowa State University
[3]Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health

### Abstract

Solving structures or structural ensembles of large macromolecular systems in solution poses a challenging problem. While NMR provides structural information at atomic resolution, increased spectral complexity, chemical shift overlap, and short transverse relaxation times (associated with slow tumbling) render application of the usual techniques that have been so successful for medium sized systems (<50 kDa) difficult. Solution X-ray scattering, on the other hand, is not limited by molecular weight but only provides low resolution structural information related to the overall shape and size of the system under investigation. Here we review how combining atomic resolution structures of smaller domains with sparse experimental data afforded by NMR residual dipolar couplings (which yield both orientational and shape information) and solution X-ray scattering data in rigid-body simulated annealing calculations provides a powerful approach for investigating the structural aspects of conformational dynamics in large multidomain proteins. The application of this hybrid methodology is illustrated for the

128 kDa dimer of bacterial Enzyme I which exists in a variety of open and closed states that are sampled at various points in the catalytic cycles, and for the capsid protein of the human immunodeficiency virus.

## Introduction

Biological macromolecules sample a range of conformational states (or microstates).[1,2] The distribution of this ensemble of microstates is dynamic and highly sensitive to changes in external conditions such as binding state,[3] ligand concentration,4 pH,[5–7] ionic strength,[8] and post-translational modifications.9 Important biological processes, including enzymatic catalysis,[10–15] ligand binding,[16-17] allostery,[18] and signaling,[9,19] depend on the exact composition of the ensemble and on interconversion rates between microstates. Understanding the delicate balance between structure and dynamics that governs biological function represents a new frontier in modern structural biology and molecular biophysics and has driven several technical and conceptual advances in the field over the past few years.

Traditionally, high-resolution structures of conformational microstates have been obtained through the application of NMR and X-ray crystallographic studies in which the experimental conditions are adjusted to select one member, or a small subset of closely related members, of the conformational ensemble.[5,6] Although this approach allows determination of very precise structural models, trapping all different microstates accessed by a biopolymer can be a time-consuming and often impractical endeavor. More recently, computational strategies have been developed for interpreting structural data from highly heterogeneous systems, such as multidomain and intrinsically disordered proteins. Common aspects of these methods are the

use of an ensemble-based representation for the system of interest, and the interpretation of experimental observables as a property of the overall ensemble instead of the individual microstates. Currently, structural ensembles are generated using one of two major approaches. The first calculates the ensemble by simulated annealing driven by the experimental data. Experimental restraints can be applied in either an ensembleor time-average manner. The second involves first generating a large pool of possible structures and then selecting among these the most appropriate ensemble that fulfills the desired experimental observables. The computational strategies for generating dynamic structure ensembles have been extensively reviewed [20-21] and will not be discussed here.

There are a number of experimental methods that can provide structural and dynamical information to describe conformational ensembles. X-ray diffraction is affected by multiple sources of disorder (e.g., protein dynamics and crystal-lattice distortions) and has been used to model conformational heterogeneity in the crystal state.[10,22] Other methods, such as Förster resonance energy transfer (FRET)[23] and double electron−electron resonance (DEER) EPR spectroscopy,[24-25] have been used to investigate the structure and dynamics of biopolymers in solution. However, both FRET and DEER require covalent labeling of the molecule under investigation with fluorophores or paramagnetic centers, respectively, and, most importantly, only provide a single distance restraint per sample. These limitations have hampered the widespread use of FRET and DEER-derived distance restraints in calculations of conformational ensembles.

Undoubtedly, the most powerful experimental technique to investigate conformational ensembles is NMR spectroscopy. Indeed a large variety of NMR parameters can be measured

that report on the structure and dynamics of biopolymers over a wide range of time scales and amplitudes of motion, and there are a number of methods to treat several types of NMR-derived restraints in ensemble calculations, including chemical shifts,[26] nuclear Overhauser enhancement (NOE) distance restraints,[27–31] generalized order parameters ($S^2$),[30,32-33] scalar three-bond J couplings,[34-35] residual dipolar couplings (RDC),[36–40] chemical shift anisotropy (CSA),[37] paramagnetic relaxation enhancement (PRE),[41–43] and hydrogen/deuterium exchange protection factors.[44] The biggest advantage of NMR over other solution techniques is that it provides structural restraints for hundreds to thousands of individual atoms, or small groups of atoms, per sample per experiment.[45] However, when investigating high molecular weight systems, such as large complexes and multidomain proteins, line broadening (due to increases in rotational correlation time) and resonance overlap (due to the increased number of NMR-active nuclei) can drastically reduce the number of analyzable cross-peaks in multidimensional NMR spectra, resulting in sparse data sets. In such cases, hybrid methods that couple sparse NMR data with structural information from other lower-resolution techniques, such as cryo-electron microscopy[46] or small/wide-angle X-ray scattering (SAXS/WAXS),[47-48] have been shown to be very powerful in solving the structure of complex molecular systems. Here we review the use simulated annealing driven by experimental NMR and SAXS/WAXS data for the quantitative investigation of the structure and dynamics of complex macromolecules at atomic resolution.

### 1. Overview of the use of SAXS/WAXS in structural biology

In structural biology, the term low resolution often has negative implications. Most instrumental techniques strive for the highest resolution possible with the gold standard being atomic resolution. Techniques such as crystallography and NMR easily achieve this goal but often

come up short when looking at complex systems such as protein assemblies, nucleic acids, and transmembrane proteins. Additionally, analyses of highly flexible systems, such as intrinsically disordered proteins, suffer greatly or may even be impossible using only crystallography or NMR. Lower-resolution techniques such as atomic force microscopy, cryoelectron microscopy, and SAXS/WAXS are emerging as extremely useful tools for structural investigation of these difficult systems.[49–51]

SAXS/WAXS is a solution state technique where a highly collimated X-ray beam is scattered by an incident sample containing the analyte of interest. This is a contrast technique that is very sensitive to changes in electron density. Because the analyte is tumbling in solution, its three-dimensional (3D) structure is reduced to a one-dimensional (1D) set of spherical shells and is represented as a one-dimensional plot of scattering intensity I against the momentum transfer vector $q$, given by:

$$q = 4\pi \sin\theta/\lambda \quad (1)$$

where 2θ is the scattering angle and λ the wavelength of the incident X-ray beam.

3D structure envelopes can be calculated through ab initio modeling, but the resulting solution is not necessarily the correct one as the information content contained within a 1D data set is necessarily limited (see “Computational Modeling” section). Nevertheless, there are several powerful verification techniques that can be used to evaluate structural hypotheses.[52–54]

In this section of the review we focus on the basics of sample preparation and requirements, instrumental design and operation, the theoretical basis of electromagnetic radiation scattering, data analysis, and finally 3D modeling of molecular structure. While this

paper covers only SAXS/WAXS, small angle neutron scattering (SANS) is very closely related and useful but will not be discussed here. There are many informative reviews on SANS in the literature.[55,56]

*2.1 Sample preparation for SAXS/WAXS*

For a SAXS experiment to yield high quality data, a correspondingly high quality sample is required. The analyte of interest must be monodisperse and of the highest possible purity (>95%). Including size exclusion chromatography as a final step in purification is required to ensure purity and check for aggregation.[57] In addition, the use of techniques such as multiangle light scattering combined with size exclusion chromatography can be useful for checking sample quality. Even a small degree of aggregation present in the sample can skew the results of analysis. This necessitates collecting data at multiple concentrations to determine the ideal conditions for each individual analyte. For proteins, the ideal concentration is generally between 1 and 10 mg/mL. A more in-depth discussion of this will be given in the Data Analysis section.

As with many other popular spectroscopic methods, SAXS/WAXS is a contrast technique. The data resulting from these experiments ($I_{total}$) are defined as the sum of the intensities of scattering due to the buffer ($I_b$) and the analyte ($I_a$):

$$I_{total} = I_b + I_a \quad (2)$$

In most cases this is not an issue as it is fairly straight forward to produce a second sample with an identical buffer system, just excluding the protein to be studied. These two samples are run under identical conditions and their corresponding data subtracted from one another:

$$I_a = I_{total} - I_b \quad (3)$$

There are, however, some other important considerations that need to be taken into account. The contrast between the analyte and bulk solvent is dependent upon differences in electron density.[58] Water, with low concentrations of salts, has an electron density of ~0.33 e/Å whereas proteins have a slightly higher electron density of ~0.44 e/Å.[59] Thus, salt concentration should be kept below 1 M and ideally less than 100 mM. In addition to signal from bulk solvent and analyte, scattering by ordered layers of solvent around the analyte must be taken into account. This implies that the presence of detergents would be detrimental to data quality, in terms of adherence to proteins and in formation of bicelles/micelles, and must be used at very low to null concentrations.[60] Overall the signal collected in these experiments (Itotal) is the sum of scattering intensities due to the buffer ($I_b$), protein ($I_a$), and ordered water layer around the protein ($I_w$).

$$I_{total} = I_b + I_a + I_w \quad (4)$$

Finally, because the sample is subjected to high energy X-rays, radiation damage can occur. Radical species scatter light very strongly and so it is particularly important to minimize radiation damage.[61] This is usually achieved in one of two ways. The X-ray source can be pulsed so as to let the sample recover from each exposure or a unidirectional constant flow cell can be utilized.[62] Flowing the solution through the exposure area will limit the amount of radiation any given portion of the sample receives but require much more material. A static cell may require as little as 1 µL whereas a flow cell may require up to 100 µL.

*2.2 Data Collection: Instrument Setup*

The main components of a SAXS instrument are the X-ray source, the collimator, the sample holder, and the detector (**Figure 3.1**).[63] X-rays from the source, whose energy is around the 1 Å range, are focused by the collimator. The spatially coherent beam interacts with the sample and is scattered by an angle 2θ, and then travels toward the detector.[64] The gathered data is sent to a computer for storage and analysis.

X-rays can be generated from a variety of sources. Historically, high voltage vacuum tubes were used but are not the best choice because of energy inefficiencies and the relatively low energy photons they produce.[65] The most common sources today are synchrotron beamlines that produce high energy photons by accelerating electrons in a magnetic field.[66] These are particularly useful because many different instruments can be set up to use the same radiation source simultaneously.

Focusing X-rays using lenses is near impossible and reflective focusing is only a little better.[67] Instead, collimation is used as a subtractive means of colinearizing photons in the beam.[68] This is ideal because even with a focused photon source there will be changes in the diameter of the beam, as a function of distance from the lens, which would interfere with the resulting data. A perfectly collimated photon source is one with no divergence; that is all photons travel a parallel path with respect to the others. In this way, the only signal on the detector outside of the beam's radius is due to scattering from the sample and not due to imperfections in the photon source.

The beam interacts with the sample (whose general criteria has been covered above) and is scattered (as described below) by an angle 2θ. The sample holder must either interact minimally with the incident photons both in terms of scattering and absorbance,[69] or must do so in a very predictable and reproducible way so that its effect can be subtracted out of the data.

*2.3 Data Analysis*

The data from a SAXS/WAXS experiment consist of a plot of scattering intensity over reciprocal space. As discussed above, the scattering intensity of the analyte, $I(q)$, is the difference between the scattering of the analyte in the sample matrix and the matrix itself. The independent variable q is defined by eq 1. Occasionally in the literature, the variables *s* and *h* are used for *q* or $q/2\pi$, respectively, so one must be careful to note the definition of the momentum transfer variable. q has units of $Å^{-1}$ which roughly translates to 2π times the inverse resolution, being that $I(q)$ is the Fourier transform of the electron density function ρ(r).

The curve $I(q)$ can generally thought to be broken into three regions; low *q*, medium *q*, and high *q* (**Figure 3.2a**). Low and medium *q* data are collected at a small angle (hence SAXS), while high q data are collected at wide angle (hence WAXS).

Low *q* refers to the range $q < 0.1\ Å^{-1}$ and reflects the size of the particle of interest. This metric is called the radius of gyration, $R_{gyr}$ (defined as the root-mean-square distance of all atoms from their relative location to the particle's center of mass). $R_{gyr}$ is a good indicator of overall size and can be calculated by regression to $I(0)$ using either the Guinier[70] or Debye[71-72] approximations:

Guinier: $$\ln I(q) = \ln I(0) - \frac{q^2 R_{gyr}^2}{3} \quad (5)$$

Debye: $$I(q) = \frac{2I(0)}{q^4 R_{gyr}^4}(q^2 R_{gyr}^2 - 1 + e^{-q^2 R_{gyr}^2})$$

$R_{gyr}$ can also be calculated by other methods that will be covered below.

The medium *q* range ($0.1 < q < 0.5$ Å$^{-1}$) is important for defining the shape of the particle. In **Figure 3.2b** we show a few examples of particle shapes and their corresponding *I*(*q*) curve. It is relatively difficult to determine the shape of a particle just based on visual inspection, but luckily there are computational procedures to aid in this respect (see below).

The high *q* range of the curve ($q > 0.5$ Å–1) is dominated by the atomic resolution structure of the particle of interest, and is usually characterized by a lower signal-to-noise ratio.

While it is difficult to access the quality of SAXS data by a cursory glance at I(q), there is one important thing to consider. *I*(*q*) is the product of a form factor, *F*(*q*)[73], that describes the shape of the scattering particle, and a structure factor, *S*(*q*), that describes the interaction between different particles in the sample.[74]

$$I(q) = F(q) \cdot S(q) \quad (7)$$

Ideally *S*(*q*) equals unity for every value of *q*, and *I*(*q*) only reflects the particle shape. At low *q*, *S*(*q*) is very sensitive to oligomerization state and the presence of aggregation, and therefore it is usually good practice to check the early points of the SAXS profile to assess the aggregation state of the sample. In particular an increase in *I*(*q*) is diagnostic of an aggregated sample[75], while, in contrast, a dip in the *I*(*q*) curve indicates repulsion between molecules.[76] A more quantitative way of evaluating the aggregation state is by using the Guinier plot. For *S*(*q*) =

1, the Guinier transformation (eq 5) yields a straight line for $0.65/R_{gyr} < q < 1.3/R_{gyr}$.[77] Curvature in this plot can be attributed to sample imperfections and a residual plot can help identify such errors (**Figure 3.3**).

SAXS/WAXS experiments are limited in their information richness relative to similar instrumental techniques such as X-ray crystallography because of solution state tumbling which averages out the three-dimensional electron density map to a one-dimensional set of spherical shells. In crystallography, the Patterson function is an autocorrelation function that shows all atom pairs (including self-pairs) and provides the best solution to crystal structures without solving the phase problem.[78] The SAXS analog is the pair-distribution function,[79] $P(r)$, the interpretation of which is a histogram of all atom–atom correlations plotted against the distance between the atoms (**Figure 3.4**). $P(r)$ is calculated by an indirect Fourier transform of $I(q)$ over $0 < r < D_{max}$:[80]

$$P(r) = \frac{r}{2\pi^2}\int_0^\infty qI(q)\sin qr\,\mathrm{d}q \qquad (8)$$

where $r$ is the distance between scattering elements (atoms in the case of a protein) and $D_{max}$ is the maximum diameter of the particle. Because the value of $D_{max}$ may not be known prior to calculation of $P(r)$, iterative optimization can be used to define the ideal bounds for the integral. The pair-distribution function is useful for determining the shape of the analyte as well as its size. Spherical proteins will give a symmetrical, bell-shaped $P(r)$ distribution while increasing deviations will result in tailing of the peak shape to higher values of $r$. When $P(r) = 0$, $r = Dmax$ as this is the point where there are no larger atom–atom pairwise distances. Once $D_{max}$ has been determined, another method of approximating $R_{gyr}$ can be calculated using the equation:

$$R_{gyr}^2 = \int_0^{D_{max}} r^2 P(r)\mathrm{d}r \Big/ \int_0^{D_{max}} P(r)\mathrm{d}r \quad (9)$$

*2.4 Computational Modeling*

Noncomputational SAXS analysis is useful for basic characterization of macromolecules, but the real value lies in computational modeling. It is possible to define molecular envelopes that describe the low-resolution shape of a protein or complex as well as carry out hypothesis testing to differentiate between discrepant models.[81-82] However, one needs to take care when considering possible envelopes calculated from SAXS data. 1D SAXS data cannot be used to directly calculate a 3D envelope. Extrapolating into higher dimensional space requires some level of "guess and check" and it is possible (and likely) to come up with many different envelopes that fit the SAXS curve quite well. No matter the quality of the SAXS data, an envelope can still be calculated! This is not to say that computational modeling of SAXS data is unreliable; on the contrary, it has been successfully used to resolve differences between solid and solution state structures, and to model missing residues in crystal structures.[83-84] In addition, the chance of accidentally misinterpreting data is almost eliminated when SAXS is used in combination with other techniques such as NMR or crystallography (see below). But it is important to keep in mind the limitations and dangers of overanalyzing results derived from SAXS.

The SAXS curve is back-calculated from a structural model using the equation

$$I(q) = \langle |A(q)|^2 \rangle_\Omega \quad (10)$$

where $\langle\ \rangle\Omega$ denotes the average over solid angle Ω (the average due to molecular tumbling in solution) and $A(q)$ is given by

$$A(q) = \sum_j f_j^{eff}(q)e^{iqx_j} + \sum_k f_k^{sph}(q)e^{iqx_k} \quad (11)$$

where q is the reciprocal space scattering vector with amplitude $q$; $j$ sums over all atoms; $f_j^{eff}(q)$ is the effective atomic scattering form factor; $x_j$ is the position of atom $j$; $k$ sums over points representing boundary-associated solvent; and $f_k^{sph}(q)$ and $x_k$ are, respectively, the positions and scattering form factors of these points. In eq 11 the first sum describes the scattering from each solute atom, while the second sum describes scattering from ordered solvent molecules bound to the surface of the macromolecule.[85] Details about back-calculation of SAXS profiles, as well as of the mathematical tricks implemented for speeding up computation, are reviewed elsewhere.[86-88]

In the context of analyzing conformational ensembles, the experimental SAXS curve is a population-weighted average of the SAXS profile of each conformer.[47] Therefore, given known structures for distinct conformational states, relative populations can be determined by simple linear combination of back-calculated scattering curves to maximize agreement with the experimental $I(q)$.

### 3. Overview of the Use of Residual Dipolar Couplings in Structural Biology

Of the various types of data that can be obtained by solution state NMR, residual dipolar couplings (RDCs) are very useful for structure refinement because they yield long-range orientational information.[89-90] RDCs describe the orientation of bond vectors relative to the static magnetic field.[91] To observe RDCs, the molecule under investigation must tumble anisotropically in solution so that the RDCs are not averaged out to zero. Weak alignment (of the order of $10^{-3}$) can be readily achieved using various dilute alignment media.[92] This section of the paper will

focus on the theoretical basis for RDCs, alignment media, alignment tensor prediction, and use of RDC restraints in structure calculation.

*3.1 Theoretical Background*

Dipolar couplings are through space interactions between magnetic nuclei. The dipolar interaction between two spins ($D_{AB}$) of gyromagnetic ratios $\gamma_A$ and $\gamma_B$ at a given distance, r, is given by

$$D_{AB}(\theta) = D_{max} \langle \frac{3\cos^2\theta - 1}{2} \rangle \quad (12)$$

where ϑ is the angle between the internuclear bond vector and the external magnetic field (**Figure 3.5**) and $D_{max}$ is the maximum value of the dipolar coupling, given by

$$D_{max} = -\frac{\mu_0 \gamma_A \gamma_B h}{8\pi^3 r^3} \quad (13)$$

where μ0 is the magnetic permeability in a vacuum and *h* is Planck's constant.[93] ⟨⟩ in eq 12 denotes the average over all possible orientations of the internuclear bond vector relative to the external magnetic field. In isotropic solution all orientations are possible, and $D_{AB}$ averages to zero. Addition of an alignment medium breaks the orientational symmetry and results in nonzero values of $D_{AB}$.

Eq 12 can be recast in terms of a molecular coordinate frame, called the alignment tensor, that describes the relative orientation of the molecule with respect to the alignment medium:

$$D_{AB}(\theta, \phi) = D_a^{AB}[3\cos^2\theta - 1 + \frac{3}{2}\sin^2\theta\cos 2\phi] \quad (14)$$

where θ is the angle between the internuclear bond vector and the z axis of the alignment tensor, φ the angle between the xy plane projection of the internuclear bond vector and the x axis (**Figure 3.5**), $D_a^{AB}$ the magnitude of the axial component of the alignment tensor, and η the rhombicity.[94]

*3.2 Alignment Media*

The maximum strength of a dipolar coupling for completely aligned samples is of the order of 20 kHz for $^{1}H–^{15}N$ and 40 kHz for $^{1}H–^{13}C$ dipolar interactions. This would cause the NMR spectrum to be so complex and broad that it would be uninterpretable. To circumvent this problem, dilute alignment media are used resulting in an ordering of $\sim 10^{-3}$ with maximum residual dipolar couplings (RDC) of about 20 Hz.[95-96] There are many different alignment media and these have been extensively reviewed.[97-103] Here we give a brief outline of the most popular media employed in structural biology applications.

Bicelles were the first media to be used for RDC measurements.[95] They are relatively easy to prepare and alignment is somewhat tunable based on the concentration of lipids. Most commonly, dimyristoylphosphatidylcholine (DMPC) and dihexanoylphosphatidylcholine (DHPC) are used to make bicelles. At low temperatures, bicelles are isotropically oriented but at temperatures greater than 30 °C they transition to a nematic liquid crystalline phase. The normal vector of the bicelles is oriented perpendicular to the applied magnetic field. Using DMPC and DHPC results in purely steric alignment but trace amounts of charged lipids can be incorporated to produce varying degrees of electrostatic alignment.[97] This makes bicelles quite useful because multiple RDC data sets can be acquired under similar solution conditions.

The second alignment medium to be described comprised rod-shaped, filamentous bacteriophages and viruses.[96, 104] A commonly used phage is pf1 that is approximately 20 000 Å long and 60 Å in diameter.[98] These virus particles, which are commercially available, spontaneously align in a magnetic field making them especially convenient for NMR. Their surface is covered in negative charges so proteins are aligned through both steric and electrostatic interactions. This can cause problems with positively charged proteins because they may align too much. The extent of alignment can be tuned by varying the concentration of phage and salt. Typically, the concentration of phage employed is 10–20 mg/mL with 100 mM or less NaCl at pH 6.5 to 8.

Polyacrylamide gel can be used for alignment of a protein when the gel is mechanically stressed.[105, 106] The pores in the gel become elongated when compressed which allows for steric alignment. First the acrylamide is polymerized using 0.1% w/v ammonium persulfate and 0.5% w/v tetramethylethylenediamine in a tube, generally, with an inner diameter between 3.5 mm to 8 mm. The gel is then washed and dehydrated. The gel is rehydrated in an NMR tube using the protein solution and then compressed using the plunger of a Shigemi NMR tube. Alternatively, radial compression can be achieved by pushing a gel with a diameter larger than the NMR tube through a funnel into the NMR tube. The alignment tensors of radially and axially compressed samples are of opposite sign and do not provide new data. Electrostatic alignment using negative or positive charges is also possible if 2-acrylamido-2-methyl-1-propanesulfonic acid or diallyldimethylammonium chloride, respectively is substituted for acrylamide during polymerization.[107] Either dialysis or dissolving the gel and centrifuging readily accomplish sample recovery.

*3.3 Measuring RDCs*

Measurement of RDCs is based on various common J scalar coupling experiments as well as some more sophisticated pulse schemes. RDCs appear in NMR spectra much in the same way that J-couplings do (**Figure 3.6**). The RDC between two nuclei is additive with that of the corresponding J coupling (**Figure 3.6**). To determine the magnitude of an RDC, one must know the sum of the dipolar and J coupling as well as the J coupling by itself, as shown in the equation below:

$$T = J + D \tag{15}$$

where *J* and *D* are the size, in Hz, of the J coupling and RDC, respectively, and *T* is the observed splitting.[108] This constitutes the basis of measuring RDCs in all of the available types of experiments. To determine the value of *J*, the experiment is carried out on a sample of the macromolecule under isotropic conditions. Using the same conditions, save for the inclusion of an alignment medium to produce anisotropy, the experiment is repeated to observe the new splitting that is equal to *T*. Subtraction of the splitting *T* from *J* yields the value of *D*.

The most commonly collected RDC for proteins is the $^{1}H–^{15}N$ backbone amide set because of its sensitivity and spectral resolution even for large, slow tumbling proteins. However, many other nuclei pairs provide useful structural information including $^{1}HN$-$^{1}H\alpha$, $^{1}HN$-$^{13}C\alpha$, $^{1}H\alpha$-$^{13}C\alpha$, $^{15}N–^{13}C\alpha$, $^{15}N–^{13}C'$, $^{13}C\alpha$-$^{13}C'$, $^{13}C$methyl-$^{13}C$, $^{1}H$methyl-$^{13}C$methyl, and $1H–^{1}H$.[104, 109-114] It is generally more straightforward to collect and analyze data from nuclei pairs that have a relatively large J-coupling, such as $^{1}H–^{15}N$ and $^{1}H–^{13}C$, because there is almost no possibility of an unresolved doublet. However, there is a greater chance of spectral overlap with other signals due

to the larger J-coupling. With smaller J-coupled systems, such as $^{13}C–^{13}C$ and $^{15}N–^{13}C$, there is less of a problem with signals interfering with one another and more of a problem with unresolved doublets.[102]

There are two main techniques for collecting RDC data: IPAP and quantitative J-modulation.[115, 116] Of these, there are important extensions that apply to various challenges. IPAP stands for in-phase/antiphase, and as the name implies, this experiment allows for the collection of the in-phase doublet spectrum, and then subsequently in an interleaved manner, the antiphase doublet spectrum.[109] The most commonly used example of this strategy is in the $^{15}N$-IPAP heteronuclear single quantum coherence (HSQC) experiment, where the spectrum yields data just as a regular $^{1}H–^{15}N$ HSQC but includes the $^{1}J_{HN}$ splitting in the nitrogen dimension. This data collection scheme is unique in that it allows for the differentiation between the upper and lower components of a doublet. When the two spectra are summed together, the low field component of the doublet is added together while the high field component is canceled out. Likewise, the difference of the spectra results in canceling the lower field component, while adding the higher component. This technique is particularly useful when there is concern of spectral overlap due to doubling the number of peaks in the coupled spectrum.

The quantitative *J*-correlation experiment measures the *J* coupling between two nuclei by exploiting differences in signal intensities due to evolution of the scalar coupling.[117] Two experiments are acquired in an interleaved fashion. The first is a decoupled reference spectrum. The second spectrum allows the coupling to evolve during a spin–echo period. The ratio of the peak intensities is proportional to the *J* coupling between the nuclei. Alternatively, the time delay during the spin–echo period can be varied through a series of experiments so as to collect

multiple peaks of varying intensities.[118] These intensities are fit to a trigonometric function to obtain the value of the coupling constant. The second method is generally more accurate, mostly by virtue of using more data points. The results of either method are resilient to signal loss effects, such as relaxation, that might skew peak integrations because each spectrum is normalized to a reference experiment that is identical to the attenuated experiment. More sophisticated versions of the basic quantitative *J* experiment have been developed and are primarily focused on expanding the applicability to larger systems. ARTSY (amide RDC by TROSY spectroscopy) is one such experiment based on a two-dimensional $^{1}H–^{15}N$ transverse relaxation optimized (TROSY) HSQC.[119] Here, the normal interleaved reference/attenuated data collection scheme is utilized. The main difference is that in the reference experiment the $^{1}H$ signal is allowed to dephase for half of the spin–echo period and in the attenuated experiment it dephases for the entire spin–echo period. As was the case before, taking the ratio of the attenuated to reference peak intensities will allow for the calculation of $J$ or $J + D$ in isotropic and aligned media, respectively.

When dealing with large systems, transverse ($R_2$) relaxation and spectral crowding becomes a considerable problem when measuring RDCs. One method to surmount both of these issues simultaneously is selective isotope labeling of methyl groups. During the production of the protein of interest, precursors are used to selectively label methyl groups of the side chains of isoleucine (I), leucine (L), valine (V), and alanine residues.[120-123] Spectra of these selectively labeled proteins are much less crowded than their uniformly labeled counterparts and, additionally, the relaxation characteristics of $^{13}CH_3$ methyl groups are particularly ideal.[124-125] This enables high resolution measurement of RDC data in large systems that might otherwise suffer

from very poor spectral quality. For ILV-selectively labeled proteins it is possible to collect $^{1}$Hmethyl-$^{13}$Cmethyl and $^{13}$Cmethyl-$^{13}$C RDCs in spectra that are extended to a third dimension ($^{13}C^{\beta/\gamma}$) to further increase resolution of crowded spectra.[126] This experiment is again based on the quantitative *J* methodology and yields best results using multiple delay values during *J* evolution and curve fitting to obtain the coupling constant.

*3.4 Data Analysis*

To utilize RDCs as restraints in structure refinement, the alignment tensor must be determined. The alignment tensor comprises 5 terms: the magnitude of the axial component of the tensor ($D_a$), the rhombicity η of the tensor, and three parameters describing the orientation of the tensor. In simulated annealing refinement, the three orientation parameters can be represented by an orthogonal axis system that is treated as a rigid body and allowed to rotate during the course of the calculation. $D_a$ and $\eta$ can be determined directly from a histogram of the measured RDCs by noting that the maximum of the distribution (corresponding to $D_{zz}$) for a fixed distance interaction is given by $2D_a^A$, the minimum (corresponding to $D_{yy}$) by $-D_a(1 + 1.5\eta)$, and the mode (corresponding to $D_{xx}$) by $-Da(1-1.5\eta)$, with the additional constraint that the sum of $D_{xx} + D_{yy} + D_{zz} = 0$ (**Figure 3.7**).[127] The robustness of this simple approach can be enhanced by application of maximum likelihood methods.[30, 128] Alternatively, the values of $D_a$ and $\eta$ can also be optimized during simulated annealing.[30]

If a structure has already been determined, its quality can be assessed using singular value decomposition (SVD) to obtain the best-fit alignment tensor that minimizes the difference between the observed RDCs and those back-calculated from the structure.[93]

The quality of the RDC fit to the coordinates is assessed using an R-factor (spanning from 0 indicating no agreement to 1 for perfect agreement) that can be expressed as

$$R = \{\langle (D_{obs} - D_{calc})^2 \rangle / (2\langle D_{obs}^2 \rangle)\}^{1/2} \quad (16)$$

where $D_{obs}$ and $D_{calc}$ are the observed and calculated RDCs, respectively.[129] $\langle D_{obs}^2 \rangle$ can either be the experimental value or can be calculated exactly by $2(D_a)^2(4 + 3\eta^2)/5$.[129] The latter, however, is not suitable for an ensemble of structures since each ensemble member will have different values of $D_a$ and $\eta$.

The manner in which the protein transiently interacts with the alignment medium is determined by steric and electrostatic effects (see above). The degree to which these two factors affect the orientation of the macromolecule under investigation varies from one alignment medium to the next, and also depends on the shape and charge distribution of the protein. For primarily steric alignment, the orientation can be approximated very well by obstruction theory.[130] The probability to find the molecule in a certain orientation Ω = (α, β, γ) is given by

$$P(\Omega) = \frac{L - l(\Omega)}{4\pi L - \int l(\Omega)} \cong \frac{L - l(\Omega)}{4\pi L} \quad (17)$$

where L is the distance between two parallel planes, $l(\Omega)$ is the length of the molecule in the direction orthogonal to the planes, and α, β, and γ are Euler angles that describe the orientation of the alignment tensor.[131] From this equation, it is immediately obvious that molecules will preferentially align with their long axis parallel to the planes. There are common programs to carry out this calculation such as PALES/SSIA[132] and Xplor-NIH.[30] Xplor-NIH can also readily carry

out this calculation during simulated annealing which is extremely useful when refining a structure ensemble.

It is much more difficult to predict the orientation of a protein in an alignment media that is dominated by electrostatic interactions because the surface charge distribution of the molecule plays a big role in alignment. The alignment tensor of macromolecules that have a very uniform charge distribution, such as DNA, can be accurately described in electrostatic aligning media while most proteins suffer due to their nonuniform surface charge.[133]

Recently, it has been shown that it is also possible to perform structure calculations incorporating RDCs without using an alignment tensor.[134-135] The technique, dubbed the "ϑ method", gets around the necessity of the alignment tensor by using eq 12 directly in the refinement protocol (i.e., without recasting into eq 14). When using the ϑ method, the orientation of the molecule relative to the external magnetic field is optimized to maximize the linear correlation between experimental RDCs and the dipolar couplings back-calculated from the structure using eq 12. Therefore, even though the ϑ method eliminates the need for calculating the alignment tensor, it still requires optimization of four variable parameters (three rotational degrees of freedom and an RDC scaling factor) to describe the alignment of the molecule in the external magnetic field.

### 4. Combined Use of SAXS and RDC Data for Structure Determination of Complex Molecular Systems

Traditionally, NMR structure determination of macromolecules requires close to complete resonance assignments and acquisition of extensive data sets of NMR-derived

structural restraints. In particular, short (up to ~6 Å) distance restraints, which are obtained by the analysis of NOE data, and restraints on the backbone torsion angles from three-bond scalar couplings and backbone chemical shifts, have been extensively employed to solve the 3D structures of globular proteins.[136] More recently, introduction of orientational restraints from RDC data measured in a dilute liquid crystalline media has enormously expanded the complexity of protein folds and assemblies amenable to structure determination by NMR.[95, 137, 138] Once local geometries are defined by NOE and dihedral angle restraints, RDC data can be used to orient rigid groups of atoms (such as secondary structure elements and domains) relative to one another. The beauty of this approach is that, if the structural domains can be treated as rigid bodies, only a small number of RDCs is required to orient the domains relative to one another.[139]

The main limitation of using RDCs to derive orientational information is the 4-fold degeneracy for orienting the alignment tensor that results in multiple solutions.[140] Ideally, NOE distance restraints[139] or a second noncolinear alignment tensor [141] can be used to resolve this ambiguity. However, these methods require measurements of interdomain NOEs or acquisition of an orthogonal set of RDC data, and are usually not applicable to larger systems for which only a very restricted set of NOEs may be measurable, and that may only be compatible with a single alignment medium. In such cases, an alternative strategy using hybrid methodology is required.

The most generally applicable hybrid method combines RDC data with SAXS/WAXS and treats structural domains as rigid bodies. RDCs provide orientational restraints and, in purely steric alignment media, shape information as well (see above), while SAXS provides complementary information on size and shape.[47] Conjoined rigid body/torsion angle simulated annealing driven by the RDC and SAXS data is then used to obtain structural solutions that are

consistent with the experimental data. To use such a hybrid approach, the structures of the individual domains must be known. This can be achieved either by experimental techniques (i.e., X-ray crystallography or NMR) or, under limited circumstances, by homology modeling when the degree of sequence similarity is very high (>60%) and there are no gaps or insertions between the experimental structure and the domain being modeled. In any case, before their use in the simulated annealing protocol, the quality of the structures must be carefully assessed by comparison of the RDCs measured on the multidomain molecule with those back-calculated from the high-resolution structures of the building blocks. It is worth mentioning that inconsistencies between experimental and back-calculated RDCs do not necessarily mean that the structures of the isolated domains are of poor quality. Indeed they can simply be the result of a conformational change induced by interdomain interactions that are absent under the experimental conditions used to determine the 3D structures of the building blocks. Rigid-body simulated annealing driven by SAXS and sparse RDC data has been used to solve the structures of large multidomain proteins,[47, 48, 142] complex nucleic acids,[143-145] and protein oligomers.[146] Recently, the technology has been expanded to the refinement of conformational ensembles.[147]

*4.1 Combined Use of SAXS and RDCs for Generation of Conformational Ensembles*

A crucial step in setting up a structure calculation protocol is to decide whether to represent the molecule under investigation using a single structure or a conformational ensemble. As a general rule, an ensemble representation can only be invoked if the experimental data cannot be explained by refinement of a single structure. In addition, in the case of multidomain proteins, the presence of large amplitude interdomain motion can be confirmed by

NMR relaxation measurements (sensitive to ps-ns time scale motions)[148, 149] or by the analysis of experimental RDC data (sensitive to motions up to the ms time scale).[150, 151]

When using a conformational ensemble the experimental data are considered as a global average representation of the system. Thus, at each step of the refinement protocol the RDC and SAXS data are back-calculated from the conformational ensemble as population-weighted averages over the ensemble members. This means that the relative populations of the ensemble members (also referred to as ensemble weights) must be optimized during the refinement. Broadly speaking, there are two ways to optimize ensemble weights. In several applications all ensemble members are given equal populations throughout the simulations.[20-21] These protocols rely on the fact that if a particular conformation is prevalent in the experimental sample, it will also be prevalent in the computational ensemble. The main limitation of such approach is that to properly describe relative populations between clusters of conformations, large ensemble sizes must be used. An alternative approach is to use a limited number of ensemble members and optimize their relative populations directly in the simulated annealing calculation.[142, 147]

If interconversion between conformational states is in the fast-to-intermediate exchange regime (nano- to millisecond range), each ensemble member has its own alignment tensor.[142, 147] Letting the alignment tensors float during refinement would add five variable parameters to the fit per ensemble member (see above), and would result in ill-defined and unstable fits. This fundamental problem can be circumvented by using a purely steric alignment medium so that the alignment tensor for each ensemble member can be directly calculated from its molecular coordinates at each step of the simulated annealing protocol.[131, 142, 147] We should also note that the tensor-free $\vartheta$ method[134] only reduces the number of parameters required to describe

alignment from five to four per ensemble member (see above) and therefore does not provide a suitable alternative.

The optimal ensemble size ($N_e$) is usually determined empirically by searching for the smallest ensemble size that satisfies the experimental data.[20] To avoid data overfitting (for example by using an unreasonably large ensemble size), parameters for evaluating agreement between experimental and back-calculated data must be carefully chosen, and, where possible, additional data sets should be collected and reserved for cross-validation of the calculated conformational ensemble. In the context of ensemble refinement against RDC and SAXS data, the agreement between the structural ensemble and experimental RDCs is conveniently evaluated in terms of an R-factor (see above).[129] The target global R-factor value is given by the weighted average of the R-factors of the individual structural domains:

$$\text{global } R_{target} = \sum_i \frac{R_i N_i}{N} \qquad (18)$$

where $R_i$ is the R-factor determined from the known structure of the *i*th structural domain using SVD, $N_i$ is the number of RDCs measured for the *i*th structural domain, and N is the total number of experimental RDCs measured for the multidomain protein. Agreement between experimental and back-calculated SAXS data, is evaluated in terms of $\chi^2$ with a target value of ~1.[152] Final R-factor and $\chi^2$ values that are smaller than the corresponding target value are indicative of overfitting the data.

In the following sections the successful application of rigid-body ensemble refinement driven by RDC and SAXS data to the challenging cases of the 128 kDa dimer of bacterial Enzyme I and the HIV-1 capsid protein will be discussed.

*4.2 Solution Structure and Dynamics of Bacterial Enzyme I*

Enzyme I (EI) is the first protein in the phosphoenolpyruvate (PEP):sugar phosphotransferase system (PTS)[153], a key signal transduction pathway involved in the regulation of central carbon metabolism in bacteria. EI is responsible for both activation and regulation of the overall PTS.[19, 154, 155] EI is a large, dynamic protein that presented a real challenge to structural biologists. In this section we will analyze how only the integrated analysis of data from multiple techniques allowed for characterization of the structure and dynamics of EI.

From the structural point of view, EI is a 128 kDa dimer of identical subunits (**Figure 3.8**). Each subunit comprises two structural domains.[156] The N-terminal domain (EIN) is further divided in two subdomains, named $EIN^{\alpha}$ and $EIN^{\alpha/\beta}$, respectively. $EIN^{\alpha/\beta}$ contains the active site residue (His189) that is autophosphorylated by phosphoenolpyruvate (PEP). $EIN^{\alpha}$ provides the binding site for HPr (the second protein in the PTS pathway).[157, 158] The C-terminal domain (EIC) is responsible for protein dimerization and contains the binding site for PEP.[17, 159]

Several X-ray structures of EI have been solved showing that the enzyme can adopt different conformations (**Figure 3.8**). Indeed, the free EI from Staphylococcus aureus[160] and Staphylococcus carnosus[161] display two distinct open states in which the His189 is positioned more than 20 Å away from the PEP binding site on EIC, and the EIN domain adopts a conformation, named the A-state, in which the $EIN^{\alpha}$ and $EIN^{\alpha/\beta}$ subdomains form extensive contacts with one another. On the other hand, the X-ray structure of a trapped phosphorylated intermediate of Escherichia coli EI,[162] solved by crystallizing EI from a solution containing PEP and $Mg^{2+}$ and quenching the autophosphorylation reaction using oxalate, shows the enzyme in a

closed conformation. In the closed state His189 is inserted into the PEP binding pocket on EIC and is positioned for in-line phosphoryl transfer from PEP to EIN. The open-to-closed transition is coupled to a conformational change in the EIN domain that involves a ~90° reorientation of $EIN^{\alpha}$ relative to $EIN^{\alpha/\beta}$ (**Figure 3.8**). The EIN conformation observed in closed EI is referred to as B-state. Interestingly these rearrangements do not affect the local fold of the structural domains ($EIN^{\alpha}$, $EIN^{\alpha/\beta}$ and EIC) that display the same structure in all the crystal structures. However, an analysis of the EI structures on the basis of experimental SAXS data acquired for free EI indicates that none of these crystal structures corresponds to the solution structure of the enzyme ($\chi^2$ for the fits to the SAXS data is >30 for all three crystal structures).[47, 163]

To solve the solution structure of free EI, a conjoined rigid body/torsion angle/Cartesian coordinate simulated annealing refinement protocol driven by the experimental RDC and SAXS was employed in which the structural domains were treated as rigid bodies, the backbone of the linkers were given Cartesian degrees of freedom, and side chains were allowed torsion degrees of freedom.[47] Due to the large size of the enzyme, assignment of the NMR spectra was achieved by transferring the assignments of the isolated EIN domain onto the spectra of the full-length protein. At the time this work was performed, no assignments were available for EIC. Therefore, the structure calculation used only 58 backbone $^{1}H–^{15}N$ RDCs from the EIN domain (29 for $EIN^{\alpha}$ and 29 for $EIN^{\alpha/\beta}$). These RDCs are fully consistent with the NMR and X-ray structures of isolated EIN in the A conformation but incompatible with the B conformation found in the structure of the closed state of EI; hence the EIN domain was held fixed to the NMR structure of free EIN. As mentioned earlier, the structure of the EIC dimer is the same in all the X-ray structures of EI; therefore, the EIC portion of the enzyme was kept fixed to the X-ray coordinates throughout the

calculation. Given that one of the principal axes of the alignment tensor must coincide with the $C_2$ symmetry axis of the dimer [164-165], the orientation of the EIN domains relative to the EIC dimer can be determined from sparse RDCs located only in the EIN domain. The results of the conjoined RDC/SAXS refinement revealed a new open structure for EI that is consistent with the experimental data (**Figure 3.9**). Interestingly, the data were fully satisfied by a single conformation, suggesting that the closed state, if at all present, is populated to a very minor extent (<5%).

More recently, a study combining NMR relaxation dispersion measurements and SAXS confirmed that the closed structure is sampled at detectable populations only in the presence of PEP.[163] Indeed, PEP-binding suppresses conformational exchange in two loops of EIC that are part of the EIN/EIC interface in closed EI. This structural stabilization of the EIC domain by PEP binding activates the open-to-closed transition that allows EI to access the catalytically competent closed state. In addition, systematic analysis of SAXS profiles acquired for wild type EI ($EI^{WT}$) and two active site mutants on the basis of the solution structure of open EI and of the crystallographic closed state, revealed that the closed state of the enzyme is largely prevalent (best fit population ~60%) in the complex between PEP and the H189A ($EI^{A}$) mutant of the enzyme,[163] opening the way to structural characterization of closed EI in solution.[142] 23, 20, and 25 backbone $^{1}H–^{15}N$ RDCs were measured for the EINα, EINα/β, and EIC, respectively, by aligning the EIA-PEP complex in neutral bicelles, and were used together with SAXS data to restrain a simulated annealing refinement calculation in which $EIN^{\alpha}$, $EIN^{\alpha/\beta}$ and EIC were treated as rigid bodies with the linkers given Cartesian degrees of freedom. Refinement against the RDC or SAXS data individually converges to a single structure representation of EIA-PEP that reproduces the experimental data

set used in the refinement but not the omitted data set (**Figure 3.10**). A two-member ensemble, however, is required to simultaneously satisfy the RDC and SAXS data (**Figures 3.10 and 3.11**). The two conformations are approximately equally populated and correspond to the crystallographic closed state, which is competent for phosphoryl transfer from PEP to the EIN domain, and a partially closed state that likely represents an intermediate between the fully closed state, which is only transiently sampled in the $EI^{WT}$-PEP complex, and the fully open apo state.[142] The partially closed state revealed by the hybrid RDC/SAXS approach is likely involved in substrate-binding and product-release steps, and eluded characterization by other crystallographic and solution techniques.

*4.3 Solution Structure and Dynamics of the HIV-1 Capsid Protein*

The human immunodeficiency virus (HIV-1) capsid protein assembles into a cone that encloses viral RNA.[166-168] The full-length capsid protein ($CA_{FL}$) consists of an arrow-shaped N-terminal domain (NTD) and a globular C-terminal domain (CTD) that undergoes a monomer/dimer equilibrium (dimer dissociation constant, $K_D$, ~40 μM at 25 °C) in solution (**Figure 3.12**). The N- and C-terminal domains are separated by a short, flexible linker (**Figure 3.12**). In the context of the mature HIV capsid, the N-terminal domains assemble into ~250 hexameric[169] and 12 pentameric[166] rings that form the exterior of the capsid. The N-terminal domain rings are connected to one another by symmetric C-terminal domain dimers (**Figure 3.12**). Although assembly of the HIV-1 capsid plays a crucial role in the virus lifecycle and infectivity, obtaining a comprehensive structural characterization of full-length capsid protein prior to assembly has been hampered by the fact that the system is highly dynamic and heterogeneous. Indeed, large amplitude motions between the N- and C-terminal domains, as

well as the presence of a dynamic monomer/dimer equilibrium, result in severe line-broadening of the NMR resonances of the linker and dimer-interface regions, making the study of full length capsid protein by conventional NMR techniques impossible.

By using the hybrid approach described here, Deshmukh et al. quantitatively determined the conformational space spanned by the N-terminal domain relative to the C-terminal domain in both monomeric and dimeric capsid protein.[147] The RDC and SAXS data were acquired for the full-length capsid protein at different concentrations (ranging from 50 μM to 260 μM) and used simultaneously to refine a conformational ensemble that include both monomeric and dimeric species. In the calculation, the relative populations of monomer and dimer were fixed based on the protein concentration and dimer $K_D$'s (determined independently by analytical ultracentrifugation at the same temperatures as those used in the RDC and SAXS measurements), and an equal number of ensemble members was used to describe the monomeric and dimeric species (i.e., $N_e^{monomer} = N_e^{dimer}$). The optimal ensemble size ($N_e = N_e^{monomer} = 2N_e^{dimer}$) was determined empirically to be 10. Indeed, further increases in ensemble size did not result in any significant improvements in agreement with the experimental RDC and SAXS data (**Figure 3.13**). The results show that the conformational space sampled by the N-terminal domain relative to the C-terminal domain is different in the monomer and dimer with a distinct pattern of transient interactions between the two domains (**Figure 3.14**). This is due to the fact that much of the conformational space sampled by the N-teminal domain in the monomer is no longer accessible in the dimer due to steric clash with the C-terminal domain dimer. The conformational ensembles derived for the dimer and monomer are characterized by six (**Figure 3.14A**) and three (**Figure 3.14B**) main structural clusters, respectively. Interestingly, one of the clusters obtained for the

protein dimer (cluster 6, which account for ~5% of the conformational ensemble of dimeric full length capsid protein) closely resembles the configuration sampled in both pentameric and hexameric capsid assemblies (**Figure 3.14C**).[147] These results suggest that the HIV-1 capsid assembles via conformational selection of a sparsely populated species and that stabilization of other clusters relative to cluster 6 may partially inhibit capsid assembly. In this regard the transient interactions between the N and C-terminal domains observed in cluster 2 of the dimer would be predicted to be enhanced by mutation of Pro38, Arg132 and Lys203 to a hydrophobic residue (alanine) consistent with the experimentally observed reduced capsid assembly rates for these three mutants.

## Concluding Remarks

An atomic level understanding of protein conformational dynamics is crucial to modem structural biology and holds the promise to address unanswered questions on the functioning of important biological molecules such as enzymes, molecular machines, and allosteric systems.[170] This goal, however, is challenging as the great majority of biologically relevant, dynamic systems are large and multimeric and, therefore, elude structural characterization by conventional techniques. In this context the use of integrative approaches that combine structural data from multiple techniques have been shown to be very useful in determining 3D structural models of large and complex molecular systems.[171] Here, we have reviewed the application of a hybrid method that uses conjoined rigid body/torsion angle/Cartesian simulated annealing refinement driven by SAXS and NMR-derived RDC data to model conformational states in large multidomain proteins. This hybrid approach is streamlined because, once the high-resolution structures of the rigid domains used as building-blocks are known from experimental or computational studies,

only sparse RDC data complemented by SAXS data are required to reliably calculate conformational ensembles. This hybrid strategy has been successfully employed to obtain a quantitative description of the magnitude and distribution of interdomain motions in the HIV-1 capsid protein[147] and the complex between bacterial Enzyme I and PEP.[142] The methodology described here is readily transferred to the study of many other challenging systems, especially those involving large multidomain proteins and their complexes. Moreover, not only can this approach characterize the conformational space sampled by one domain relative to another as in the case of the HIV-1 capsid protein,[147] but it can also detect the simultaneous existence of distinct conformations and characterize their structures as in the case of the 128 kDa complex of Enzyme I (H189A) with PEP.[142] Further, in the case of the HIV-1 capsid protein the simultaneous existence of monomer and dimer can readily be handled.[147] In this regard it is worth noting that it would be impossible to characterize a mixture of coexisting states by crystallography as the crystallization procedure would only allow one state to crystallize out, and the probability of obtaining crystals of the various states is likely to be very small. Similarly, the coexistence of multiple conformational states renders interpretation of cryoelectron microscopy images extremely difficult if not impossible. Of course, any structural approach involving solution NMR does place certain limits on molecular size but, with deuteration and appropriate methyl-specific labeling, systems up to 200–300 kDa can potentially be tackled.

## References

1. Frauenfelder, H.; Sligar, S. G.; Wolynes, P. G. The energy landscapes and motions of proteins Science 1991, 254, 1598– 1603DOI: 10.1126/science.1749933

2. Dill, K. A. Polymer principles and protein folding Protein Sci. 1999, 8, 1166–1180DOI: 10.1110/ps.8.6.1166

3. Tang, C.; Schwieters, C. D.; Clore, G. M. Open-to-closed transition in apo maltose-binding protein observed by paramagnetic NMR Nature 2007, 449, 1078–1082DOI: 10.1038/nature06232

4. Tsai, C. J.; Ma, B.; Nussinov, R. Folding and binding cascades: shifts in energy landscapes Proc. Natl. Acad. Sci. U. S. A. 1999, 96, 9970– 9972DOI: 10.1073/pnas.96.18.9970

5. Reiter, N. J.; Blad, H.; Abildgaard, F.; Butcher, S. E. Dynamics in the U6 RNA intramolecular stem-loop: a base flipping conformational change Biochemistry 2004, 43, 13739– 13747DOI: 10.1021/bi048815y

6. Venditti, V.; Clos, L., 2nd; Niccolai, N.; Butcher, S. E. Minimum-energy path for a u6 RNA conformational change involving protonation, base-pair rearrangement and base flipping J. Mol. Biol. 2009, 391, 894– 905DOI: 10.1016/j.jmb.2009.07.003

7. Venditti, V.; Niccolai, N.; Butcher, S. E. Measuring the dynamic surface accessibility of RNA with the small paramagnetic molecule TEMPOL Nucleic Acids Res. 2007, 36, e20DOI: 10.1093/nar/gkm1062

8. Suh, J. Y.; Tang, C.; Clore, G. M. Role of electrostatic interactions in transient encounter complexes in protein-protein association investigated by paramagnetic relaxation enhancement J. Am. Chem. Soc. 2007, 129, 12954– 12955DOI: 10.1021/ja0760978

9. Tsai, C. J.; Del Sol, A.; Nussinov, R. Protein allostery, signal transmission and dynamics: a classification scheme of allosteric mechanisms Mol. BioSyst. 2009, 5, 207– 216DOI: 10.1039/b819720b

10. Fraser, J. S.; Clarkson, M. W.; Degnan, S. C.; Erion, R.; Kern, D.; Alber, T. Hidden alternative structures of proline isomerase essential for catalysis Nature 2009, 462, 669– 673DOI: 10.1038/nature08615

11. Henzler-Wildman, K. A.; Lei, M.; Thai, V.; Kerns, S. J.; Karplus, M.; Kern, D. A hierarchy of timescales in protein dynamics is linked to enzyme catalysis Nature 2007, 450, 913– 916DOI: 10.1038/nature06407

12. Henzler-Wildman, K. A.; Thai, V.; Lei, M.; Ott, M.; Wolf-Watz, M.; Fenn, T.; Pozharski, E.; Wilson, M. A.; Petsko, G. A.; Karplus, M.; Hubner, C. G.; Kern, D. Intrinsic motions along an enzymatic reaction trajectory Nature 2007, 450, 838–844DOI: 10.1038/nature06410

13. Eisenmesser, E. Z.; Millet, O.; Labeikovsky, W.; Korzhnev, D. M.; Wolf-Watz, M.; Bosco, D. A.; Skalicky, J. J.; Kay, L. E.; Kern, D. Intrinsic dynamics of an enzyme underlies catalysis Nature 2005, 438, 117– 121DOI: 10.1038/nature04105

14. Tuttle, L. M.; Dyson, H. J.; Wright, P. E. Side-chain conformational heterogeneity of intermediates in the Escherichia coli dihydrofolate reductase catalytic cycle Biochemistry 2013, 52, 3464– 3477DOI: 10.1021/

15. Bhabha, G.; Lee, J.; Ekiert, D. C.; Gam, J.; Wilson, I. A.; Dyson, H. J.; Benkovic, S. J.; Wright, P. E. A dynamic knockout reveals that conformational fluctuations influence the chemical step of enzyme catalysis Science 2011, 332, 234– 238DOI: 10.1126/science.1198542

16. Clore, G. M. Interplay between conformational selection and induced fit in multidomain protein-ligand binding probed by paramagnetic relaxation enhancement Biophys. Chem. 2014, 186, 3– 12DOI: 10.1016/j.bpc.2013.08.006

17. Venditti, V.; Clore, G. M. Conformational selection and substrate binding regulate the monomer/dimer equilibrium of the C-terminal domain of Escherichia coli Enzyme I J. Biol. Chem. 2012, 287, 26989– 26998DOI: 10.1074/jbc.M112.382291

18. Motlagh, H. N.; Wrabl, J. O.; Li, J.; Hilser, V. J. The ensemble nature of allostery Nature 2014, 508, 331– 339DOI: 10.1038/nature13001

19. Clore, G. M.; Venditti, V. Structure, dynamics and biophysics of the cytoplasmic protein-protein complexes of the bacterial phosphoenolpyruvate: sugar phosphotransferase system Trends Biochem. Sci. 2013, 38, 515– 530DOI: 10.1016/j.tibs.2013.08.003

20. Anthis, N. J.; Clore, G. M. Visualizing transient dark states by NMR spectroscopy Q. Rev. Biophys. 2015, 48, 35– 116DOI: 10.1017/S0033583514000122

21. Angyan, A. F.; Gaspari, Z. Ensemble-based interpretations of NMR structural data to describe protein internal dynamics Molecules 2013, 18, 10548– 10567DOI: 10.3390/molecules180910548

22. Burnley, B. T.; Afonine, P. V.; Adams, P. D.; Gros, P. Modelling dynamics in protein crystal structures by ensemble refinement eLife 2012, 1, e00311DOI: 10.7554/eLife.00311

23. Hillisch, A.; Lorenz, M.; Diekmann, S. Recent advances in FRET: distance determination in protein-DNA complexes Curr. Opin. Struct. Biol. 2001, 11, 201– 207DOI: 10.1016/S0959-440X(00)00190-1

24. Altenbach, C.; Kusnetzow, A. K.; Ernst, O. P.; Hofmann, K. P.; Hubbell, W. L. High-resolution distance mapping in rhodopsin reveals the pattern of helix movement due to activation Proc. Natl. Acad. Sci. U. S. A. 2008, 105, 7439– 7444DOI: 10.1073/pnas.0802515105

25. Klug, C. S.; Feix, J. B. Methods and applications of site-directed spin labeling EPR spectroscopy Methods Cell Biol. 2008, 84, 617– 658DOI: 10.1016/S0091-679X(07)84020-9

26. Cavalli, A.; Montalvao, R. W.; Vendruscolo, M. Using chemical shifts to determine structural changes in proteins upon complex formation J. Phys. Chem. B 2011, 115, 9491– 9494DOI: 10.1021/jp202647q

27. Bonvin, A. M.; Boelens, R.; Kaptein, R. Time- and ensemble-averaged direct NOE restraints J. Biomol. NMR 1994, 4, 143– 149DOI: 10.1007/BF00178343

28. Kemmink, J.; van Mierlo, C. P.; Scheek, R. M.; Creighton, T. E. Local structure due to an aromatic-amide interaction observed by 1H-nuclear magnetic resonance spectroscopy in peptides related to the N terminus of bovine pancreatic trypsin inhibitor J. Mol. Biol. 1993, 230, 312– 322DOI: 10.1006/jmbi.1993.1144

29. Vogeli, B.; Kazemi, S.; Guntert, P.; Riek, R. Spatial elucidation of motion in proteins by ensemble-based structure calculation using exact NOEs Nat. Struct. Mol. Biol. 2012, 19, 1053– 1057DOI: 10.1038/nsmb.2355

30. Schwieters, C. D.; Kuszewski, J. J.; Clore, G. M. Using Xplor-NIH for NMR molecular structure determination Prog. Nucl. Magn. Reson. Spectrosc. 2006, 48, 47– 62DOI: 10.1016/j.pnmrs.2005.10.001

31. Clore, G. M.; Schwieters, C. D. Concordance of residual dipolar couplings, backbone order parameters and crystallographic B-factors for a small alpha/beta protein: a unified picture of high probability, fast atomic motions in proteins J. Mol. Biol. 2006, 355, 879– 886DOI: 10.1016/j.jmb.2005.11.042

32. Best, R. B.; Vendruscolo, M. Determination of protein structures consistent with NMR order parameters J. Am. Chem. Soc. 2004, 126, 8090– 8091DOI: 10.1021/ja0396955

33. Batta, G.; Barna, T.; Gaspari, Z.; Sandor, S.; Kover, K. E.; Binder, U.; Sarg, B.; Kaiserer, L.; Chhillar, A. K.; Eigentler, A.; Leiter, E.; Hegedus, N.; Pocsi, I.; Lindner, H.; Marx, F. Functional aspects of the solution structure and dynamics of PAF--a highly-stable antifungal protein from Penicillium chrysogenum FEBS J. 2009, 276, 2875– 2890DOI: 10.1111/j.1742-4658.2009.07011.x

34. Lee, J. H.; Li, F.; Grishaev, A.; Bax, A. Quantitative residue-specific protein backbone torsion angle dynamics from concerted measurement of 3J couplings J. Am. Chem. Soc. 2015, 137, 1432– 1435DOI: 10.1021/ja512593s

35. Lindorff-Larsen, K.; Best, R. B.; Depristo, M. A.; Dobson, C. M.; Vendruscolo, M. Simultaneous determination of protein structure and dynamics Nature 2005, 433, 128– 132DOI: 10.1038/nature03199

36. Clore, G. M.; Schwieters, C. D. How much backbone motion in ubiquitin is required to account for dipolar coupling data measured in multiple alignment media as assessed by independent cross-validation? J. Am. Chem. Soc. 2004, 126, 2923–2938DOI: 10.1021/ja0386804

37. Hess, B.; Scheek, R. M. Orientation restraints in molecular dynamics simulations using time and ensemble averaging J. Magn. Reson. 2003, 164, 19– 27DOI: 10.1016/S1090-7807(03)00178-2

38. Louhivuori, M.; Otten, R.; Lindorff-Larsen, K.; Annila, A. Conformational fluctuations affect protein alignment in dilute liquid crystal media J. Am. Chem. Soc. 2006, 128, 4371– 4376DOI: 10.1021/ja0576334

39. Montalvao, R. W.; De Simone, A.; Vendruscolo, M. Determination of structural fluctuations of proteins from structure-based calculations of residual dipolar couplings J. Biomol. NMR 2012, 53, 281– 292DOI: 10.1007/s10858-012-9644-3

40. De Simone, A.; Montalvao, R. W.; Vendruscolo, M. Determination of conformational equilibria in proteins using residual dipolar couplings J. Chem. Theory Comput. 2011, 7, 4189– 4195DOI: 10.1021/ct200361b

41. Clore, G. M.; Iwahara, J. Theory, practice, and applications of paramagnetic relaxation enhancement for the characterization of transient low-population states of biological macromolecules and their complexes Chem. Rev. 2009, 109, 4108– 4139DOI: 10.1021/cr900033p

42. Iwahara, J.; Clore, G. M. Detecting transient intermediates in macromolecular binding by paramagnetic NMR Nature 2006, 440, 1227– 1230DOI: 10.1038/nature04673

43. Iwahara, J.; Schwieters, C. D.; Clore, G. M. Ensemble approach for NMR structure refinement against 1H paramagnetic relaxation enhancement data arising from a flexible paramagnetic group attached to a macromolecule J. Am. Chem. Soc. 2004, 126, 5879– 5896DOI: 10.1021/ja031580d

44. Best, R. B.; Vendruscolo, M. Structural interpretation of hydrogen exchange protection factors in proteins: characterization of the native state fluctuations of CI2 Structure 2006, 14, 97– 106DOI: 10.1016/j.str.2005.09.012

45. Clore, G. M. Visualizing lowly-populated regions of the free energy landscape of macromolecular complexes by paramagnetic relaxation enhancement Mol. BioSyst. 2008, 4, 1058– 1069DOI: 10.1039/b810232e

46. Loquet, A.; Sgourakis, N. G.; Gupta, R.; Giller, K.; Riedel, D.; Goosmann, C.; Griesinger, C.; Kolbe, M.; Baker, D.; Becker, S.; Lange, A. Atomic model of the type III secretion system needle Nature 2012, 486, 276– 279DOI: 10.1038/nature11079

47. Schwieters, C. D.; Suh, J. Y.; Grishaev, A.; Ghirlando, R.; Takayama, Y.; Clore, G. M. Solution structure of the 128 kDa enzyme I dimer from Escherichia coli and its 146 kDa complex with HPr using residual dipolar couplings and small- and wide-angle X-ray scattering J. Am. Chem. Soc. 2010, 132, 13026– 13045DOI: 10.1021/ja105485b

48. Takayama, Y.; Schwieters, C. D.; Grishaev, A.; Ghirlando, R.; Clore, G. M. Combined use of residual dipolar couplings and solution X-ray scattering to rapidly probe rigid-body conformational transitions in a non-phosphorylatable active-site mutant of the 128 kDa enzyme I dimer J. Am. Chem. Soc. 2011, 133, 424– 427DOI: 10.1021/ja109866w

49. Shinsky, S. A.; Cosgrove, M. S. Unique role of the WD-40 repeat protein 5 (WDR5) subunit within the mixed lineage leukemia 3 (MLL3) histone methyltransferase complex J. Biol. Chem. 2015, 290, 25819– 25833DOI: 10.1074/jbc.M115.684142

50. Heinrich, F.; Chakravarthy, S.; Nanda, H.; Papa, A.; Pandolfi, P. P.; Ross, A. H.; Harishchandra, R. K.; Gericke, A.; Losche, M. The PTEN tumor suppressor forms homodimers in solution Structure 2015, 23, 1952– 1957DOI: 10.1016/j.str.2015.07.012

51. Xia, Z.; Bell, D. R.; Shi, Y.; Ren, P. RNA 3D structure prediction by using a coarse-grained model and experimental data J. Phys. Chem. B 2013, 117, 3135– 3144DOI: 10.1021/jp400751w

52. Hura, G. L.; Menon, A. L.; Hammel, M.; Rambo, R. P.; Poole, F. L., 2nd; Tsutakawa, S. E.; Jenney, F. E., Jr.; Classen, S.; Frankel, K. A.; Hopkins, R. C.; Yang, S. J.; Scott, J. W.; Dillard, B. D.; Adams, M. W.; Tainer, J. A. Robust, high-throughput solution structural analyses by small angle X-ray scattering (SAXS) Nat. Methods 2009, 6, 606– 612DOI: 10.1038/nmeth.1353

53. Tsutakawa, S. E.; Hura, G. L.; Frankel, K. A.; Cooper, P. K.; Tainer, J. A. Structural analysis of flexible proteins in solution by small angle X-ray scattering combined with crystallography J. Struct. Biol. 2007, 158, 214– 223DOI: 10.1016/j.jsb.2006.09.008

54. Horejs, C.; Pum, D.; Sleytr, U. B.; Tscheliessnig, R. SAXS: Structure Verification of an S-Layer Protein using a Fractal Mean Force Potential Biophys. J. 2010, 98, 760aDOI: 10.1016/j.bpj.2009.12.4171

55. Gutberlet, T.; Heinemann, U.; Steiner, M. Protein crystallography with neutrons--status and perspectives Acta Crystallogr., Sect. D: Biol. Crystallogr. 2001, 57, 349– 354DOI: 10.1107/S0907444900019934

56. Breyton, C.; Gabel, F.; Lethier, M.; Flayhan, A.; Durand, G.; Jault, J. M.; Juillan-Binard, C.; Imbert, L.; Moulin, M.; Ravaud, S.; Hartlein, M.; Ebel, C. Small angle neutron scattering for the study of solubilised membrane proteins Eur. Phys. J. E: Soft Matter Biol. Phys. 2013, 36, 71DOI: 10.1140/epje/i2013-13071-6

57. Rambo, R. P.; Tainer, J. A. Improving small-angle X-ray scattering data for structural analyses of the RNA world RNA 2010, 16, 638– 646DOI: 10.1261/rna.1946310

58. Grishaev, A.; Anthis, N. J.; Clore, G. M. Contrast-matched small-angle X-ray scattering from a heavy-atom-labeled protein in structure determination: application to a lead-substituted calmodulin-peptide complex J. Am. Chem. Soc. 2012, 134, 14686– 14689DOI: 10.1021/ja306359z

59. Putnam, C. D.; Hammel, M.; Hura, G. L.; Tainer, J. A. X-ray solution scattering (SAXS) combined with crystallography and computation: defining accurate macromolecular structures, conformations and assemblies in solution Q. Rev. Biophys. 2007, 40, 191– 285DOI: 10.1017/S0033583507004635

60. Grishaev, A. Sample preparation, data collection, and preliminary data analysis in biomolecular solution X-ray scattering. Curr. Protoc. Protein Sci. 2012, Chapter 17, Unit 17.14.DOI: 10.1002/0471140864.ps1714s70

61. Garman, E. F.; Weik, M. Radiation damage to macromolecules: kill or cure? J. Synchrotron Radiat. 2015, 22, 195– 200DOI: 10.1107/S160057751500380X

62. Lipfert, J.; Millett, I. S.; Seifert, S.; Doniach, S. Sample holder for small-angle x-ray scattering static and flow cell measurement Rev. Sci. Instrum. 2006, 77, 046108DOI: 10.1063/1.2194484

63. Koch, M. H. J. Instruments and methods for small-angle scattering with synchrotron radiation Makromol. Chem., Macromol. Symp. 1988, 15, 19– 90DOI: 10.1002/masy.19880150106

64. Pontoni, D.; Narayanan, T.; Rennie, A. High-dynamic range SAXS data acquisition with an X-ray image intensifier J. Appl. Crystallogr. 2002, 35, 207– 211DOI: 10.1107/S0021889802000493

65. Otendal, M.; Tuohimaa, T.; Vogt, U.; Hertz, H. M. A 9 keV electron-impact liquid-gallium-jet x-ray source Rev. Sci. Instrum. 2008, 79, 016102DOI: 10.1063/1.2833838

66. Koch, M. H. J. SAXS Instrumentation for Synchrotron Radiation then and now J. Phys. Conf. Ser. 2010, 247, 1– 23DOI: 10.1088/1742-6596/247/1/012001

67. Hamley, I. W.; Pedersen, J. S. Analysis of Neutron and X-ray Reflectivity Data. I. Theory J. Appl. Crystallogr. 1994, 27, 29– 35DOI: 10.1107/S0021889893006260

68. Chu, B.; Harney, P. J.; Li, Y.; Linliu, K.; Yeh, F. A laser-aided prealigned pinhole collimator for synchrotron x rays Rev. Sci. Instrum. 1994, 65, 597– 602DOI: 10.1063/1.1145123

69. Block, I. D.; Scheffold, F. Modulated 3D cross-correlation light scattering: improving turbid sample characterization Rev. Sci. Instrum. 2010, 81, 123107DOI: 10.1063/1.3518961

70. Lipfert, J.; Doniach, S. Small-angle X-ray scattering from RNA, proteins, and protein complexes Annu. Rev. Biophys. Biomol. Struct. 2007, 36, 307– 327DOI: 10.1146/annurev.biophys.36.040306.132655

71. Calmettes, P.; Durand, D.; Desmadril, M.; Minard, P.; Receveur, V.; Smith, J. C. How random is a highly denatured protein? Biophys. Chem. 1994, 53, 105– 113DOI: 10.1016/0301-4622(94)00081-6

72. Doniach, S. Changes in biomolecular conformation seen by small angle X-ray scattering Chem. Rev. 2001, 101, 1763– 1778DOI: 10.1021/cr990071k

73. Groenewegen, P.; Feil, D. Molecular Form Factors in X-ray Crystallography Acta Crystallogr., Sect. A: Cryst. Phys., Diffr., Theor. Gen. Crystallogr. 1969, 25, 444– 450DOI: 10.1107/S0567739469000891

74. Hilfiker, R.; Eicke, H.; Sager, W.; Steeb, C.; Hofmeier, U.; Gehrke, R. Form and structure factors of water/AOT/Oil microemulsions from synchrotron SAXS Ber. Bunsenges Phys. Chem. 1990, 94, 677– 683DOI: 10.1002/bbpc.19900940611

75. Huang, H. K.; Sheu, H. S.; Chuang, W. T.; Jeng, U. S.; Su, A. C.; Wu, W. R.; Liao, K. F.; Chen, C. Y.; Chang, S. Y.; Lai, H. M. Correlated changes in structure and viscosity during gelatinization and gelation of tapioca starch granules IUCrJ 2014, 1, 418– 428DOI: 10.1107/S2052252514019137

76. Zhang, F.; Skoda, M. W.; Jacobs, R. M.; Martin, R. A.; Martin, C. M.; Schreiber, F. Protein interactions studied by SAXS: effect of ionic strength and protein concentration for BSA in aqueous solutions J. Phys. Chem. B 2007, 111, 251– 259DOI: 10.1021/jp0649955

77. Grant, T. D.; Luft, J. R.; Carter, L. G.; Matsui, T.; Weiss, T. M.; Martel, A.; Snell, E. H. The accurate assessment of small-angle X-ray scattering data Acta Crystallogr., Sect. D: Biol. Crystallogr. 2015, 71, 45– 56DOI: 10.1107/S1399004714010876

78. Patterson, A. L. A Fourier series method for the determination of the components of interatomic distances in crystals Phys. Rev. 1934, 46, 372– 376DOI: 10.1103/PhysRev.46.372

79. Debye, P.; Bueche, A. M. Scattering by an Inhomogeneous Solid J. Appl. Phys. 1949, 20, 518– 525DOI: 10.1063/1.1698419

80. Glatter, O. A new method for the evaluation of small-angle scattering data J. Appl. Crystallogr. 1977, 10, 415– 421DOI: 10.1107/S0021889877013879

81. Sousa, M. C.; Trame, C. B.; Tsuruta, H.; Wilbanks, S. M.; Reddy, V. S.; McKay, D. B. Crystal and solution structures of an HslUV protease-chaperone complex Cell 2000, 103, 633– 643DOI: 10.1016/S0092-8674(00)00166-5

82. Mendillo, M. L.; Putnam, C. D.; Kolodner, R. D. Escherichia coli MutS tetramerization domain structure reveals that stable dimers but not tetramers are essential for DNA mismatch repair in vivo J. Biol. Chem. 2007, 282, 16345–16354DOI: 10.1074/jbc.M700858200

83. Vestergaard, B.; Sanyal, S.; Roessle, M.; Mora, L.; Buckingham, R. H.; Kastrup, J. S.; Gajhede, M.; Svergun, D. I.; Ehrenberg, M. The SAXS solution structure of RF1 differs from its crystal structure and is similar to its ribosome bound cryo-EM structure Mol. Cell 2005, 20, 929– 938DOI: 10.1016/j.molcel.2005.11.022

84. Petoukhov, M. V.; Eady, N. A.; Brown, K. A.; Svergun, D. I. Addition of missing loops and domains to protein models by x-ray solution scattering Biophys. J. 2002, 83, 3113– 3125DOI: 10.1016/S0006-3495(02)75315-0

85. Schwieters, C. D.; Clore, G. M. A physical picture of atomic motions within the Dickerson DNA dodecamer in solution derived from joint ensemble refinement against NMR and large-angle X-ray scattering data Biochemistry 2007, 46, 1152–1166DOI: 10.1021/bi061943x

86. Svergun, D. I. Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing Biophys. J. 1999, 76, 2879–2886DOI: 10.1016/S0006-3495(99)77443-6

87. Heller, W. T. ELLSTAT: shape modeling forsolution small-angle scattering of proteins and protein complexes with automated statistical characterization J. Appl. Crystallogr. 2006, 39, 671– 675DOI: 10.1107/S0021889806029591

88. Walther, D.; Cohen, F. E.; Doniach, S. Reconstruction of low-resolution three-dimensionaldensity maps from one-dimensional small-angle X-ray solution scattering data for biomolecules J. Appl. Crystallogr. 2000, 33, 350– 363DOI: 10.1107/S0021889899015976

89. Tolman, J. R.; Flanagan, J. M.; Kennedy, M. A.; Prestegard, J. H. Nuclear magnetic dipole interactions in field-oriented proteins: information for structure determination in solution Proc. Natl. Acad. Sci. U. S. A. 1995, 92, 9279– 9283DOI: 10.1073/pnas.92.20.9279

90. Eletsky, A.; Pulavarti, S. V.; Beaumont, V.; Gollnick, P.; Szyperski, T. Solution NMR experiment for measurement of 15N-1H residual dipolar couplings in large proteins and supramolecular complexes J. Am. Chem. Soc. 2015, 137, 11242– 11245DOI: 10.1021/jacs.5b07010

91. Smith, S. A.; Palke, W. E.; Gerig, J. T. The Hamiltonians of NMR Part II Concepts Magn. Reson. 1992, 4, 181– 204DOI: 10.1002/cmr.1820040302

92. Bax, A.; Grishaev, A. Weak alignment NMR: a hawk-eyed view of biomolecular structure Curr. Opin. Struct. Biol. 2005, 15, 563– 570DOI: 10.1016/j.sbi.2005.08.006

93. Prestegard, J. H.; Bougault, C. M.; Kishore, A. I. Residual dipolar couplings in structure determination of biomolecules Chem. Rev. 2004, 104, 3519– 3540DOI: 10.1021/cr030419i

94. Lipsitz, R. S.; Tjandra, N. Residual dipolar couplings in NMR structure analysis Annu. Rev. Biophys. Biomol. Struct. 2004, 33, 387– 413DOI: 10.1146/annurev.biophys.33.110502.140306

95. Tjandra, N.; Bax, A. Direct measurement of distances and angles in biomolecules by NMR in a dilute liquid crystalline medium Science 1997, 278, 1111– 1114DOI: 10.1126/science.278.5340.1111

96. Clore, G. M.; Starich, M. R.; Gronenborn, A. M. Measurement of residual dipolar couplings of macromolecules aligned in the nematic phase of a colloidal suspension of rod-shaped viruses J. Am. Chem. Soc. 1998, 120, 10571– 10572DOI: 10.1021/ja982592f

97. Losonczi, J. A.; Prestegard, J. H. Improved dilute bicelle solutions for high-resolution NMR of biological macromolecules J. Biomol. NMR 1998, 12, 447– 451DOI: 10.1023/A:1008302110884

98. Hansen, M. R.; Hanson, P.; Pardi, A. Filamentous bacteriophage for aligning RNA, DNA, and proteins for measurement of nuclear magnetic resonance dipolar coupling interactions Methods Enzymol. 2000, 317, 220– 240DOI: 10.1016/S0076-6879(00)17017-X

99. Bertini, I.; Felli, I. C.; Luchinat, C. Lanthanide induced residual dipolar couplings for the conformational investigation of peripheral 15NH2 moieties J. Biomol. NMR 2000, 18, 347– 355DOI: 10.1023/A:1026785228634

100. Sass, H. J.; Musco, G.; Stahl, S. J.; Wingfield, P. T.; Grzesiek, S. Solution NMR of proteins within polyacrylamide gels: diffusional properties and residual alignment by mechanical stress or embedding of oriented purple membranes J. Biomol. NMR 2000, 18, 303– 309DOI: 10.1023/A:1026703605

101. Prestegard, J. H.; Kishore, A. I. Partial alignment of biomolecules: an aid to NMR characterization Curr. Opin. Chem. Biol. 2001, 5, 584– 590DOI: 10.1016/S1367-5931(00)00247-7

102. Bax, A.; Kontaxis, G.; Tjandra, N. Dipolar couplings in macromolecular structure determination Methods Enzymol. 2001, 339, 127– 174DOI: 10.1016/S0076-6879(01)39313-8

103. Maltsev, A. S.; Grishaev, A.; Roche, J.; Zasloff, M.; Bax, A. Improved cross validation of a static ubiquitin structure derived from high precision residual dipolar couplings measured in a drug-based liquid crystalline phase J. Am. Chem. Soc. 2014, 136, 3752– 3755DOI: 10.1021/ja4132642

104. Hansen, M. R.; Mueller, L.; Pardi, A. Tunable alignment of macromolecules by filamentous phage yields dipolar coupling interactions Nat. Struct. Biol. 1998, 5, 1065– 1074DOI: 10.1038/4176

105. Chou, J. J.; Gaemers, S.; Howder, B.; Louis, J. M.; Bax, A. A simple apparatus for generating stretched polyacrylamide gels, yielding uniform alignment of proteins and detergent micelles J. Biomol. NMR 2001, 21, 377– 382DOI: 10.1023/A:1013336502594

106. Tycko, R.; Blanco, F. J.; Y, I. Alignment of biopolymers in strained gels: a new way to create detectable dipole-dipole couplings in high-resolution biomolecular NMR J. Am. Chem. Soc. 2000, 122, 9340– 9341DOI: 10.1021/ja002133q

107. Ulmer, T. S.; Ramirez, B. E.; Delaglio, F.; Bax, A. Evaluation of backbone proton positions and dynamics in a small protein by liquid crystal NMR spectroscopy J. Am. Chem. Soc. 2003, 125, 9179– 9191DOI: 10.1021/ja0350684

108. Brunner, E. Residual Dipolar Couplings in Protein NMR Concepts Magn. Reson. 2001, 13, 238– 259DOI: 10.1002/cmr.1012

109. Ottiger, M.; Delaglio, F.; Bax, A. Measurement of J and dipolar couplings from simplified two-dimensional NMR spectra J. Magn. Reson. 1998, 131, 373– 378DOI: 10.1006/jmre.1998.1361

110. Tjandra, N.; Marquardt, J.; Clore, G. M. Direct refinement against proton-proton dipolar couplings in NMR structure determination of macromolecules J. Magn. Reson. 2000, 142, 393– 396DOI: 10.1006/jmre.1999.1985

111. Kontaxis, G.; Bax, A. Multiplet component separation for measurement of methyl 13C-1H dipolar couplings in weakly aligned proteins J. Biomol. NMR 2001, 20, 77– 82DOI: 10.1023/A:1011280529850

112. Meier, S.; Haussinger, D.; Jensen, P.; Rogowski, M.; Grzesiek, S. High-accuracy residual 1HN-13C and 1HN-1HN dipolar couplings in perdeuterated proteins J. Am. Chem. Soc. 2003, 125, 44– 45DOI: 10.1021/ja028740q

113. Permi, P. Measurement of residual dipolar couplings from 1Hα to 13Cα and 15N using a simple HNCA-based experiment J. Biomol. NMR 2003, 27, 341– 349DOI: 10.1023/A:1025866606252

114. Vogeli, B.; Yao, L.; Bax, A. Protein backbone motions viewed by intraresidue and sequential HN-Hα residual dipolar couplings J. Biomol. NMR 2008, 41, 17– 28DOI: 10.1007/s10858-008-9237-3

115. Tjandra, N.; Grzesiek, S.; Bax, A. Magnetic Field Dependence of Nitrogen-Proton J Splittings in 15N-Enriched Human Ubiquitin Resulting from Relaxation Interference and Residual Dipolar Coupling J. Am. Chem. Soc. 1996, 118, 6264– 6272DOI: 10.1021/ja960106n

116. Tjandra, N.; Bax, A. Large variations in 13Cα chemical shift anisotropy in proteins correlate with secondary structure J. Am. Chem. Soc. 1997, 119, 9576– 9577DOI: 10.1021/ja9721374

117. Bax, A.; Vuister, G. W.; Grzesiek, S.; Delaglio, F.; Wang, A. C.; Tschudin, R.; Zhu, G. Measurement of homo- and heteronuclear J couplings from quantitative J correlation Methods Enzymol. 1994, 239, 79– 105DOI: 10.1016/S0076-6879(94)39004-5

118. Liu, Y.; Prestegard, J. H. Measurement of one and two bond N-C couplings in large proteins by TROSY-based J-modulation experiments J. Magn. Reson. 2009, 200, 109– 118DOI: 10.1016/j.jmr.2009.06.010

119. Fitzkee, N. C.; Bax, A. Facile measurement of 1H-15N residual dipolar couplings in larger perdeuterated proteins J. Biomol. NMR 2010, 48, 65– 70DOI: 10.1007/s10858-010-9441-9

120. Rosen, M. K.; Gardner, K. H.; Willis, R. C.; Parris, W. E.; Pawson, T.; Kay, L. E. Selective methyl group protonation of perdeuterated proteins J. Mol. Biol. 1996, 263, 627– 636DOI: 10.1006/jmbi.1996.0603

121. Gardner, K. H.; Kay, L. E. The use of 2H, 13C, 15N multidimensional NMR to study the structure and dynamics of proteins Annu. Rev. Biophys. Biomol. Struct. 1998, 27, 357– 406DOI: 10.1146/annurev.biophys.27.1.357

122. Godoy-Ruiz, R.; Guo, C.; Tugarinov, V. Alanine methyl groups as NMR probes of molecular structure and dynamics in high-molecular-weight proteins J. Am. Chem. Soc. 2010, 132, 18340– 18350DOI: 10.1021/ja1083656

123. Clark, L.; Zahm, J. A.; Ali, R.; Kukula, M.; Bian, L.; Patrie, S. M.; Gardner, K. H.; Rosen, M. K.; Rosenbaum, D. M. Methyl labeling and TROSY NMR spectroscopy of proteins expressed in the eukaryote Pichia pastoris J. Biomol. NMR 2015, 62, 239–245DOI: 10.1007/s10858-015-9939-2

124. Ollerenshaw, J. E.; Tugarinov, V.; Kay, L. E. Methyl TROSY: explaination and experimental verification Magn. Reson. Chem. 2003, 41, 843– 852DOI: 10.1002/mrc.1256

125. Tugarinov, V.; Hwang, P. M.; Ollerenshaw, J. E.; Kay, L. E. Cross-correlated relaxation enhanced 1H-13C NMR spectroscopy of methyl groups in very high molecular weight proteins and protein complexes J. Am. Chem. Soc. 2003, 125, 10420– 10428DOI: 10.1021/ja030153x

126. Guo, C.; Godoy-Ruiz, R.; Tugarinov, V. High resolution measurement of methyl 13Cm-13C and 1Hm-13Cm residual dipolar couplings in large proteins J. Am. Chem. Soc. 2010, 132, 13984– 13987DOI: 10.1021/ja1041435

127. Clore, G. M.; Gronenborn, A. M.; Bax, A. A robust method for determining the magnitude of the fully asymmetric alignment tensor of oriented macromolecules in the absence of structural information J. Magn. Reson. 1998, 133, 216– 221DOI: 10.1006/jmre.1998.1419

128. Warren, J. J.; Moore, P. B. Application of dipolar coupling data to the refinement of the solution structure of the sarcin-ricin loop RNA J. Biomol. NMR 2001, 20, 311– 323DOI: 10.1023/A:1011214214552

129. Clore, G. M.; Garrett, D. S. R-factor, free R, and complete cross-validation for dipolar coupling refinement of NMR structures J. Am. Chem. Soc. 1999, 121, 9008–9012DOI: 10.1021/ja991789k

130. Zweckstetter, M.; Bax, A. Prediction of sterically induced alignment in a dilute liquid crystalline phase: aid to protein structure determination by NMR J. Am. Chem. Soc. 2000, 122, 3791– 3792DOI: 10.1021/ja0000908

131. Huang, J. R.; Grzesiek, S. Ensemble calculations of unstructured proteins constrained by RDC and PRE data: a case study of urea-denatured ubiquitin J. Am. Chem. Soc. 2010, 132, 694– 705DOI: 10.1021/ja907974m

132. Zweckstetter, M. NMR: prediction of molecular alignment from structure using the PALES software Nat. Protoc. 2008, 3, 679– 690DOI: 10.1038/nprot.2008.36

133. Zweckstetter, M.; Hummer, G.; Bax, A. Prediction of charge-induced molecular alignment of biomolecules dissolved in dilute liquid-crystalline phases Biophys. J. 2004, 86, 3444– 3460DOI: 10.1529/biophysj.103.035790

134. Camilloni, C.; Vendruscolo, M. A tensor-free method for the structural and dynamical refinement of proteins using residual dipolar couplings J. Phys. Chem. B 2015, 119, 653– 661DOI: 10.1021/jp5021824

135. Camilloni, C.; Vendruscolo, M. Reply to ”Comment on ’A tensor-free method for the structural and dynamic refinement of proteins using residual dipolar couplings’” J. Phys. Chem. B 2015, 119, 8225– 8226DOI: 10.1021/acs.jpcb.5b04166

136. Clore, G. M.; Gronenborn, A. M. Determining the structures of large proteins and protein complexes by NMR Trends Biotechnol. 1998, 16, 22– 34DOI: 10.1016/S0167-7799(97)01135-9

137. Tjandra, N.; Omichinski, J. G.; Gronenborn, A. M.; Clore, G. M.; Bax, A. Use of dipolar 1H-15N and 1H-13C couplings in the structure determination of magnetically oriented macromolecules in solution Nat. Struct. Biol. 1997, 4, 732–738DOI: 10.1038/nsb0997-732

138. Clore, G. M.; Gronenborn, A. M. New methods of structure refinement for macromolecular structure determination by NMR Proc. Natl. Acad. Sci. U. S. A. 1998, 95, 5891– 5898DOI: 10.1073/pnas.95.11.5891

139. Clore, G. M. Accurate and rapid docking of protein-protein complexes on the basis of intermolecular nuclear overhauser enhancement data and dipolar couplings by rigid body minimization Proc. Natl. Acad. Sci. U. S. A. 2000, 97, 9021– 9025DOI: 10.1073/pnas.97.16.9021

140. Ramirez, B. E.; Bax, A. Modulation of the alignment tensor of macromolecules dissolved in a dilute liquid crystalline medium J. Am. Chem. Soc. 1998, 120, 9106–9107DOI: 10.1021/ja982310b

141. Al-Hashimi, H. M.; Valafar, H.; Terrell, M.; Zartler, E. R.; Eidsness, M. K.; Prestegard, J. H. Variation of molecular alignment as a means of resolving orientational ambiguities in protein structures from dipolar couplings J. Magn. Reson. 2000, 143, 402– 406DOI: 10.1006/jmre.2000.2049

142. Venditti, V.; Schwieters, C. D.; Grishaev, A.; Clore, G. M. Dynamic equilibrium between closed and partially closed states of the bacterial Enzyme I unveiled by solution NMR and X-ray scattering Proc. Natl. Acad. Sci. U. S. A. 2015, 112, 11565–11570DOI: 10.1073/pnas.1515366112

143. Wang, Y. X.; Zuo, X.; Wang, J.; Yu, P.; Butcher, S. E. Rapid global structure determination of large RNA and RNA complexes using NMR and small-angle X-ray scattering Methods 2010, 52, 180– 191DOI: 10.1016/j.ymeth.2010.06.009

144. Burke, J. E.; Butcher, S. E. Nucleic acid structure characterization by small angle X-ray scattering (SAXS). Curr. Protoc. Nucleic Acid Chem. 2012, Chapter 7, Unit 7.18.DOI: 10.1002/0471142700.nc0718s51

145. Burke, J. E.; Sashital, D. G.; Zuo, X.; Wang, Y. X.; Butcher, S. E. Structure of the yeast U2/U6 snRNA complex RNA 2012, 18, 673– 683DOI: 10.1261/rna.031138.111

146. Rossi, P.; Shi, L.; Liu, G.; Barbieri, C. M.; Lee, H. W.; Grant, T. D.; Luft, J. R.; Xiao, R.; Acton, T. B.; Snell, E. H.; Montelione, G. T.; Baker, D.; Lange, O. F.; Sgourakis, N. G. A hybrid NMR/SAXS-based approach for discriminating oligomeric protein interfaces using Rosetta Proteins: Struct., Funct., Genet. 2015, 83, 309– 317DOI: 10.1002/prot.24719

147. Deshmukh, L.; Schwieters, C. D.; Grishaev, A.; Ghirlando, R.; Baber, J. L.; Clore, G. M. Structure and dynamics of full-length HIV-1 capsid protein in solution J. Am. Chem. Soc. 2013, 135, 16133– 16147DOI: 10.1021/ja406246z

148. Clore, G. M.; Szabo, A.; Bax, A.; Kay, L. E.; Driscoll, P. C.; Gronenborn, A. M. Deviations from the simple two parameter model free approach to the interpretation of 15N nuclear magnetic relaxation of proteins J. Am. Chem. Soc. 1990, 112, 4989– 4991DOI: 10.1021/ja00168a070

149. Baber, J. L.; Szabo, A.; Tjandra, N. Analysis of slow interdomain motion of macromolecules using NMR relaxation data J. Am. Chem. Soc. 2001, 123, 3953– 3959DOI: 10.1021/ja0041876

150. Braddock, D. T.; Cai, M.; Baber, J. L.; Huang, Y.; Clore, G. M. Rapid identification of medium- to large-scale interdomain motion in modular proteins using dipolar couplings J. Am. Chem. Soc. 2001, 123, 8634– 8635DOI: 10.1021/ja016234f

151. Braddock, D. T.; Louis, J. M.; Baber, J. L.; Levens, D.; Clore, G. M. Structure and dynamics of KH domains from FBP bound to single-stranded DNA Nature 2002, 415, 1051– 1056DOI: 10.1038/4151051a

152. Schwieters, C. D.; Clore, G. M. Using small angle solution scattering data in Xplor-NIH structure calculations Prog. Nucl. Magn. Reson. Spectrosc. 2014, 80, 1– 11DOI: 10.1016/j.pnmrs.2014.03.001

153. Chauvin, F.; Brand, L.; Roseman, S. Enzyme I: the first protein and potential regulator of the bacterial phosphoenolpyruvate: glycose phosphotransferase system Res. Microbiol. 1996, 147, 471– 479DOI: 10.1016/0923-2508(96)84001-0

154. Meadow, N. D.; Fox, D. K.; Roseman, S. The bacterial phosphoenolpyruvate: glycose phosphotransferase system Annu. Rev. Biochem. 1990, 59, 497– 542DOI: 10.1146/annurev.bi.59.070190.002433

155. Venditti, V.; Ghirlando, R.; Clore, G. M. Structural basis for Enzyme I inhibition by alpha-ketoglutarate ACS Chem. Biol. 2013, 8, 1232– 1240DOI: 10.1021/cb400027q

156. LiCalsi, C.; Crocenzi, T. S.; Freire, E.; Roseman, S. Sugar transport by the bacterial phosphotransferase system. Structural and thermodynamic domains of enzyme I of Salmonella typhimurium J. Biol. Chem. 1991, 266, 19519– 19527

157. Garrett, D. S.; Seok, Y. J.; Liao, D. I.; Peterkofsky, A.; Gronenborn, A. M.; Clore, G. M. Solution structure of the 30 kDa N-terminal domain of enzyme I of the Escherichia coli phosphoenolpyruvate:sugar phosphotransferase system by multidimensional NMR Biochemistry 1997, 36, 2517– 2530DOI: 10.1021/bi962924y

158. Garrett, D. S.; Seok, Y. J.; Peterkofsky, A.; Clore, G. M.; Gronenborn, A. M. Identification by NMR of the binding surface for the histidine-containing phosphocarrier protein HPr on the N-terminal domain of enzyme I of the Escherichia coli phosphotransferase system Biochemistry 1997, 36, 4393– 4398DOI: 10.1021/bi970221q

159. Seok, Y. J.; Lee, B. R.; Zhu, P. P.; Peterkofsky, A. Importance of the carboxyl-terminal domain of enzyme I of the Escherichia coli phosphoenolpyruvate: sugar phosphotransferase system for phosphoryl donor specificity Proc. Natl. Acad. Sci. U. S. A. 1996, 93, 347– 351DOI: 10.1073/pnas.93.1.347

160. Oberholzer, A. E.; Schneider, P.; Siebold, C.; Baumann, U.; Erni, B. Crystal structure of enzyme I of the phosphoenolpyruvate sugar phosphotransferase system in the dephosphorylated state J. Biol. Chem. 2009, 284, 33169– 33176DOI: 10.1074/jbc.M109.057612

161. Marquez, J.; Reinelt, S.; Koch, B.; Engelmann, R.; Hengstenberg, W.; Scheffzek, K. Structure of the full-length enzyme I of the phosphoenolpyruvate-dependent sugar phosphotransferase system J. Biol. Chem. 2006, 281, 32508– 32515DOI: 10.1074/jbc.M513721200

162. Teplyakov, A.; Lim, K.; Zhu, P. P.; Kapadia, G.; Chen, C. C.; Schwartz, J.; Howard, A.; Reddy, P. T.; Peterkofsky, A.; Herzberg, O. Structure of phosphorylated enzyme I, the phosphoenolpyruvate:sugar phosphotransferase system sugar translocation signal protein Proc. Natl. Acad. Sci. U. S. A. 2006, 103, 16218– 16223DOI: 10.1073/pnas.0607587103

163. Venditti, V.; Tugarinov, V.; Schwieters, C. D.; Grishaev, A.; Clore, G. M. Large interdomain rearrangement triggered by suppression of micro- to millisecond dynamics in bacterial Enzyme I Nat. Commun. 2015, 6, 5960DOI: 10.1038/ncomms6960

164. Bewley, C. A.; Clore, G. M. J. Am. Chem. Soc. 2000, 122, 6009– 6016DOI: 10.1021/ja000858o

165. Clore, G. M.; Bewley, C. A. Using conjoined rigid body/torsion angle simulated annealing to determine the relative orientation of covalently linked protein domains from dipolar couplings J. Magn. Reson. 2002, 154, 329– 335DOI: 10.1006/jmre.2001.2489

166. Pornillos, O.; Ganser-Pornillos, B. K.; Yeager, M. Atomic-level modelling of the HIV capsid Nature 2011, 469, 424– 427DOI: 10.1038/nature09640

167. Dahirel, V.; Shekhar, K.; Pereyra, F.; Miura, T.; Artyomov, M.; Talsania, S.; Allen, T. M.; Altfeld, M.; Carrington, M.; Irvine, D. J.; Walker, B. D.; Chakraborty, A. K. Coordinate linkage of HIV evolution reveals regions of immunological vulnerability Proc. Natl. Acad. Sci. U. S. A. 2011, 108, 11530– 11535DOI: 10.1073/pnas.1105315108

168. Zhao, G.; Perilla, J. R.; Yufenyuy, E. L.; Meng, X.; Chen, B.; Ning, J.; Ahn, J.; Gronenborn, A. M.; Schulten, K.; Aiken, C.; Zhang, P. Mature HIV-1 capsid structure by cryo-electron microscopy and all-atom molecular dynamics Nature 2013, 497, 643– 646DOI: 10.1038/nature12162

169. Pornillos, O.; Ganser-Pornillos, B. K.; Kelly, B. N.; Hua, Y.; Whitby, F. G.; Stout, C. D.; Sundquist, W. I.; Hill, C. P.; Yeager, M. X-ray structures of the hexameric building block of the HIV capsid Cell 2009, 137, 1282– 1292DOI: 10.1016/j.cell.2009.04.

170. van den Bedem, H.; Fraser, J. S. Integrative, dynamic structural biology at atomic resolution--it's about time Nat. Methods 2015, 12, 307– 318DOI: 10.1038/nmeth.3324

171. Ward, A. B.; Sali, A.; Wilson, A. I. Integrative structural biology Science 2013, 339, 913– 915DOI: 10.1126/science.1228565

172. Valentini, E.; Kikhney, A. G.; Previtali, G.; Jeffries, C. M.; Svergun, D. I. SASBDB, a repository for biological small-angle scattering data Nucleic Acids Res. 2015, 43, D357– D363DOI: 10.1093/nar/gku1047

173. Bewley, C. A.; Gustafson, K. R.; Boyd, M. R.; Covell, D. G.; Bax, A.; Clore, G. M.; Gronenborn, A. M. Solution structure of cyanovirin-N, a potent HIV-inactivating protein Nat. Struct. Biol. 1998, 5, 571– 578DOI: 10.1038/828

174. Liao, D. I.; Silverton, E.; Seok, Y. J.; Lee, B. R.; Peterkofsky, A.; Davies, D. R. The first step in sugar transport: crystal structure of the amino terminal domain of enzyme I of the E. coli PEP: sugar phosphotransferase system and a model of the phosphotransfer complex with HPr Structure 1996, 4, 861– 872DOI: 10.1016/S0969-2126(96)00092-5

175. Schwieters, C. D.; Clore, G. M. Reweighted atomic densities to represent ensembles of NMR structures J. Biomol. NMR 2002, 23, 221– 225DOI: 10.1023/A:1019875223132

176. Du, S.; Betts, L.; Yang, R.; Shi, H.; Concel, J.; Ahn, J.; Aiken, C.; Zhang, P.; Yeh, J. I. Structure of the HIV-1 full-length capsid protein in a conformationally trapped unassembled state induced by small-molecule binding J. Mol. Biol. 2011, 406, 371–386DOI: 10.1016/j.jmb.2010.11.027

**Figures**

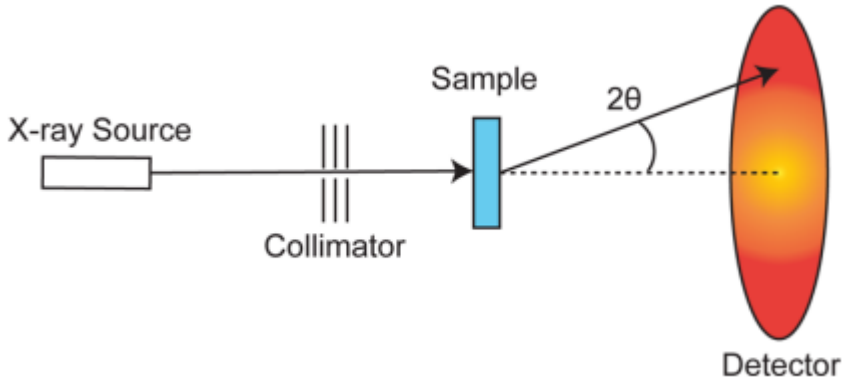


**Figure 3.1.** Schematic representation of a SAXS instrument.

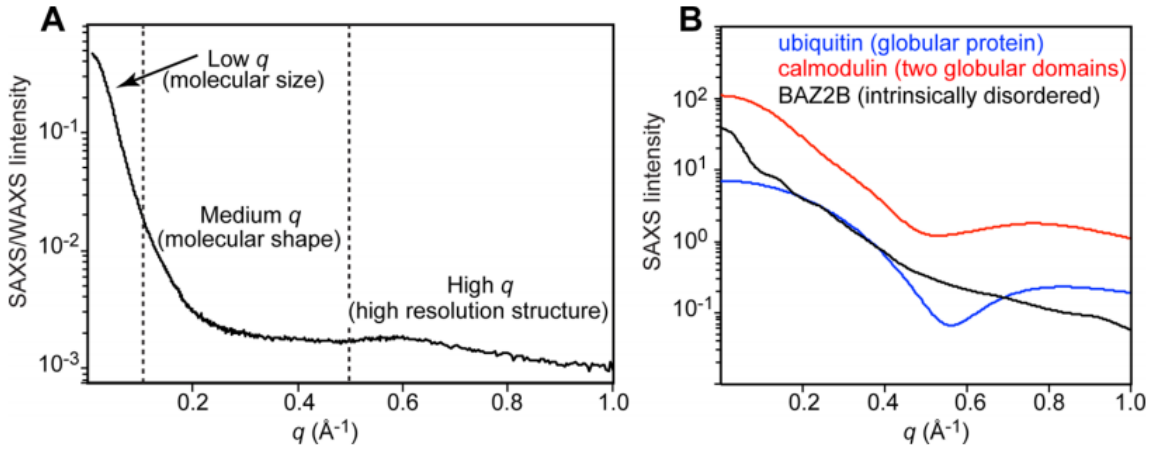


**Figure 3.2.** (A) SAXS/WAXS curve acquired for Enzyme I of the bacterial phosphoenolpyruvate:sugar phosphotransferase system. Vertical dashed lines separate the low, medium and high q regions. (B) SAXS curves acquired for ubiquitin (blue), calmodulin (red), and the bromodomain of protein 2B (BAZ2B – black). Data displayed in (B) were downloaded from the Small Angle Scattering Biological Data Bank (SASBDB).[172]

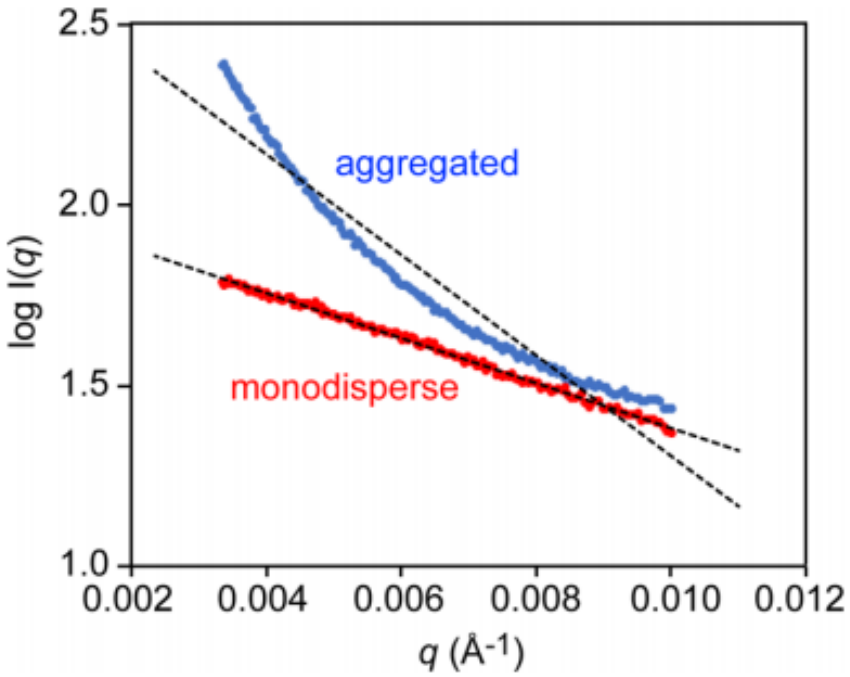


**Figure 3.3.** Examples of Guinier plots for monodisperse (red) and aggregated (blue) samples.

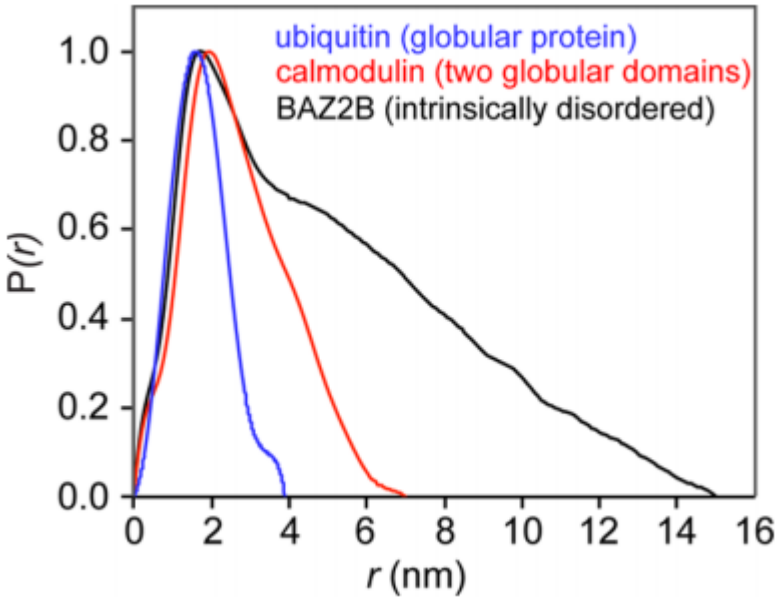


**Figure 3.4.** Pair-distribution functions, P(r), obtained from the SAXS curves acquired for ubiquitin (blue), calmodulin (red), and the bromodomain of protein 2B (BAZ2B – black). The experimental SAXS data were obtained from the SASBDB.[172]

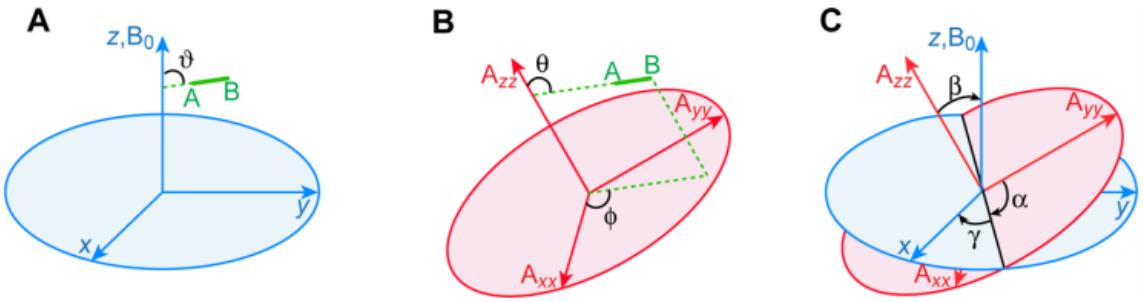


**Figure 3.5.** Relationship between the bond vector A-B, the alignment tensor and the external magnetic field (B0). (A) The angle ϑ describes the orientation of the bond vector relative to $B_0$; (B) the angles θ and φ define the orientation of the bond vector relative to the alignment tensor; and (C) the relationship between the alignment tensor and the external magnetic field is given by the Euler angles α, β and γ. The A-B vector is displayed green; the alignment tensor is colored red. The magnetic field $B_0$ is taken parallel to an external reference frame (blue).

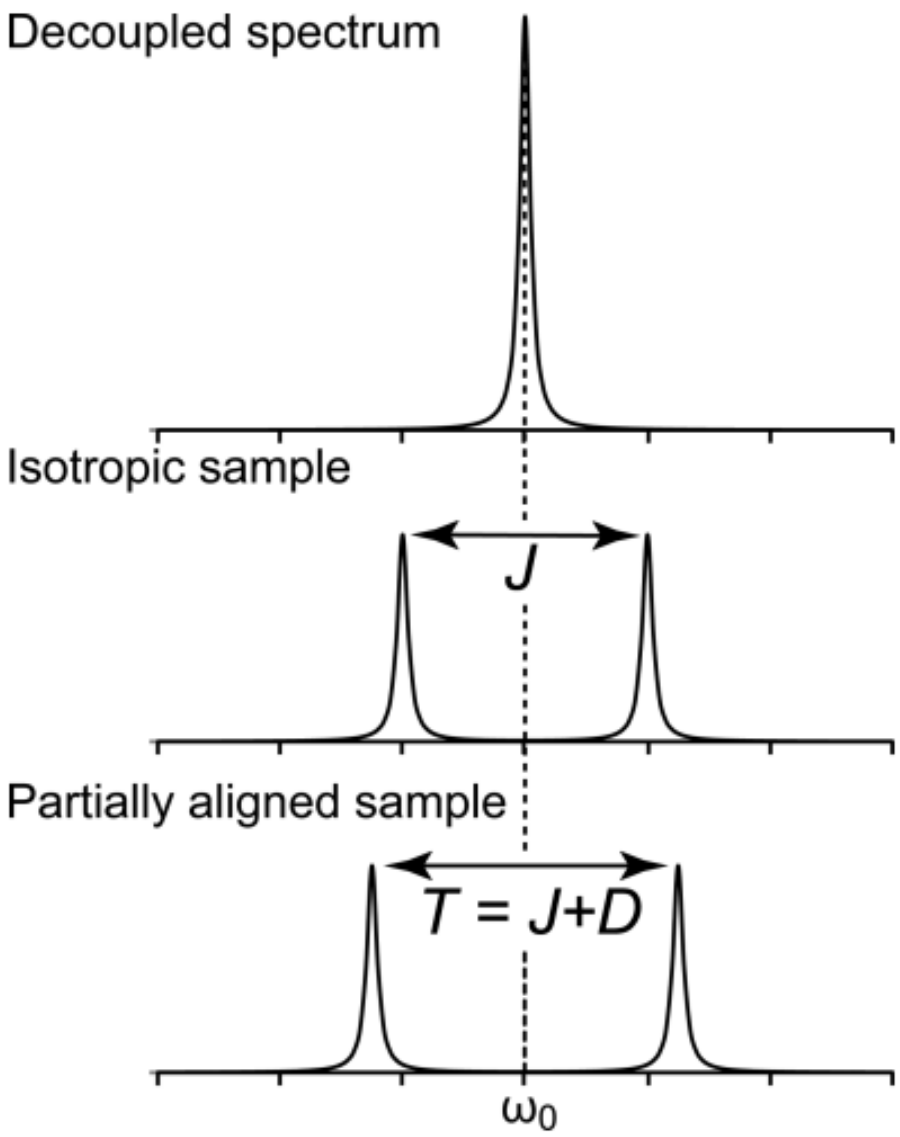


**Figure 3.6.** Comparison between a NMR peak measured in the (A) absence of coupling (i.e., decoupled), (B) presence of scalar (*J*) coupling (isotropic sample), and (C) presence of scalar (*J*) and dipolar (D) coupling (partially aligned sample)

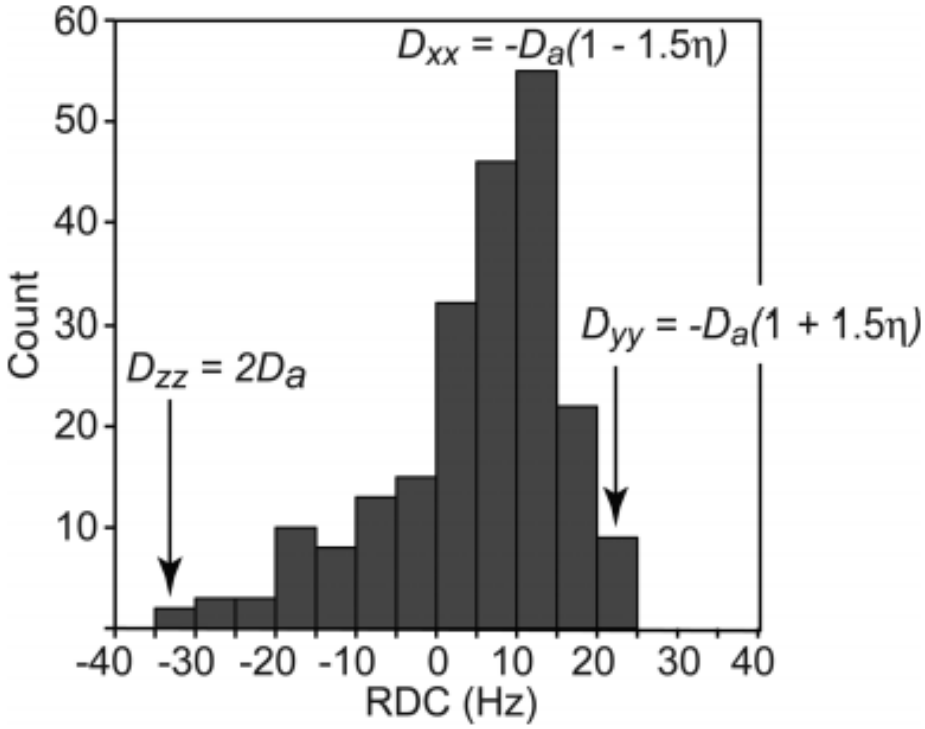


**Figure 3.7.** Deriving the magnitude of the axial component and rhombicity of the alignment tensor from a histogram of normalized RDCs. The data are taken for the protein cyanovirin[173], and the backbone Cα-Hα and Cα-C′ RDCs are normalized relative to the backbone N–H RDCs. The maximum, minimum and mode of the distribution correspond to $D_{zz} = 2D_a$, $Dyy = -D_a(1 + 1.5\eta)$, and $D_{xx} = -D_a(1 - 1.5\eta)$, where D*a* (scaled for N–H bond vectors) is the magnitude of the axial component of the alignment tensor and η the rhombicity. The sum of the three orthogonal components of the alignment tensor is equal to zero ($D_{zz} + D_{yy} + D_{xx} = 0$).

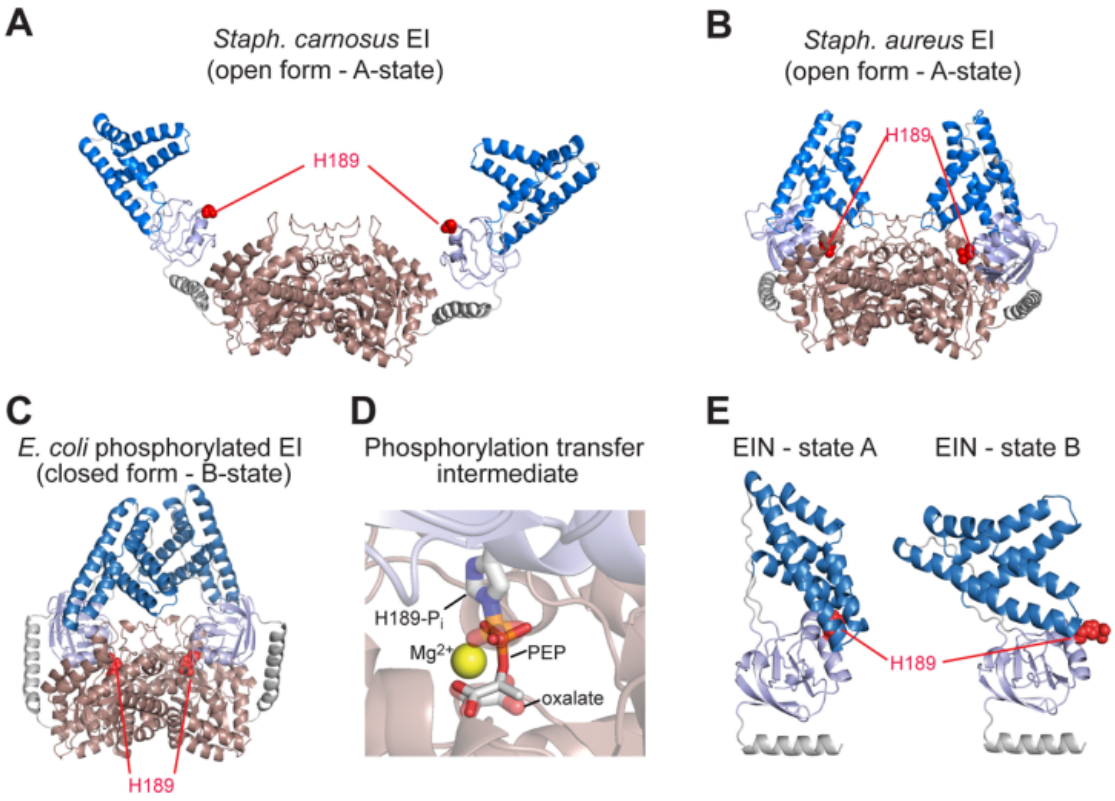


**Figure 3.8.** X-ray structures of bacterial Enzyme I (EI). (A) EI from Staph. carnosus (2HRO)[161] (B) EI from Staph. aureus (2WQD)[160] (C) phosphorylated EI from E. coli (2HWG)[162] The C-terminal dimerization domain (EIC) is colored pink; the EINα and $EIN^{\alpha/\beta}$ subdomains of the N-terminal domain (EIN) are colored blue and light blue, respectively; the active site His189, located in the EINα/β subdomain, is shown as red spheres. (D) Structural model of phosphoenolpyruvate (PEP) bound to the active site of the E. coli EI structure. PEP is shown as solid sticks, and the $Mg^{2+}$ ion is displayed as a yellow sphere. The phosphorylated His189 and the oxalate molecule in the X-ray structure of the phosphoryl transfer intermediate in the closed state (2HWG) are displayed as transparent sticks. (E) A [157, 174] and B [162] conformations of the EIN domain seen in the open and closed states of EI, respectively.

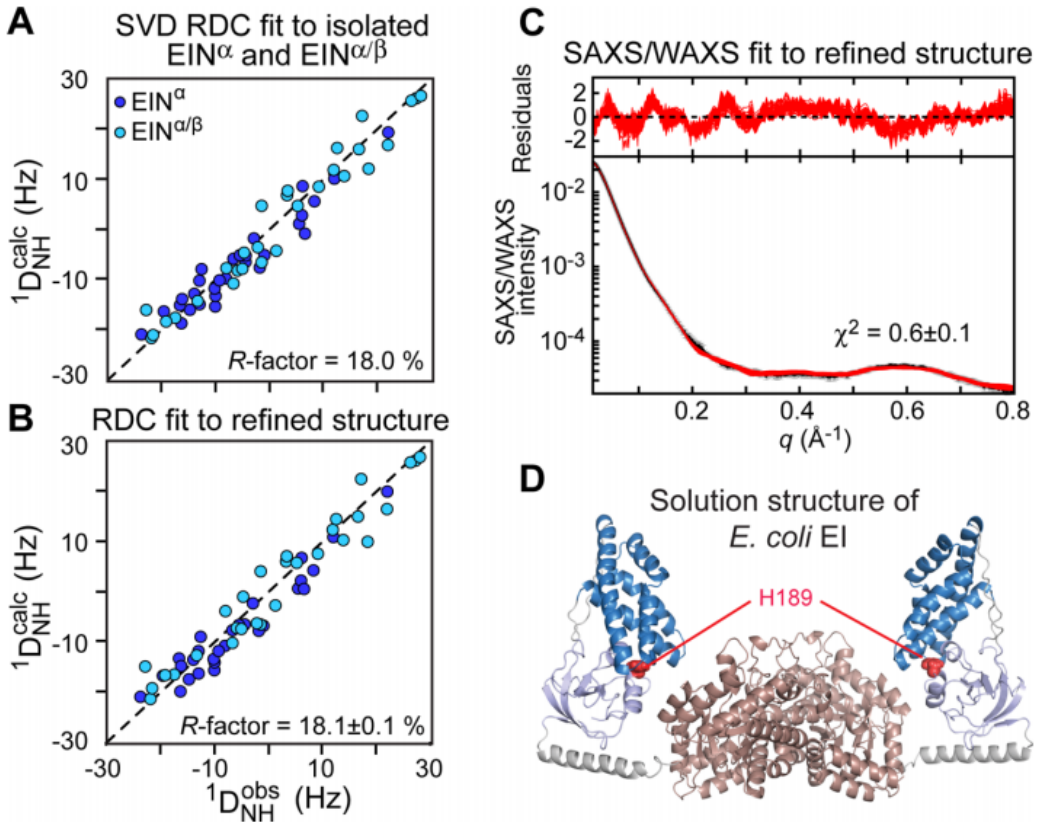


**Figure 3.9.** Combined SAXS/RDC refinement of the E. coli EI structure. (A) Comparison of the observed and calculated RDCs obtained by SVD fits to the individual $EIN^{\alpha}$ (blue circles) and $EIN^{\alpha/\beta}$ (light blue circles) subdomains. (B) Comparison of the observed and calculated RDCs for the refined structure of the EI dimer. The resulting RDC R-factor is the same (within experimental error) as the RDC R-factor obtained from the SVD fits to the individual subdomains. (C) Agreement between the experimental and back calculated SAXS/WAXS curve. (D) Solution structure of unliganded E. coli EI obtained by combined SAXS/RDC refinement (2KX9)[47]. Color coding as in Figure 8. Adapted from Schwieters et al. (2010).[47]

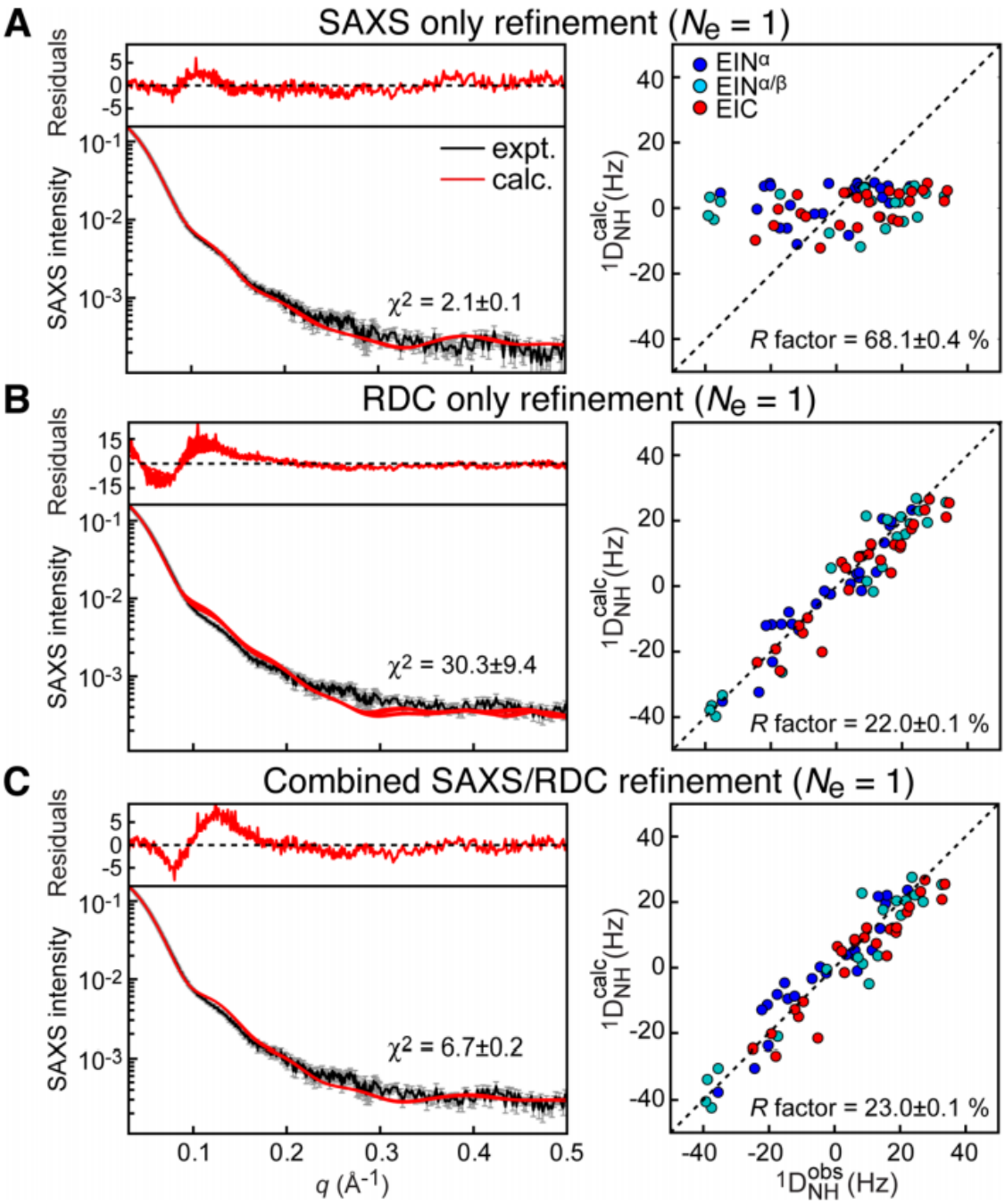


**Figure 3.10.** Comparison between experimental and back-calculated SAXS (left) and RDC (right) data for a single-structure refinement of the EI$^{A}$-PEP complex. (A) Refinement using only SAXS data. (B) Refinement using only RDC data. (C) Combined SAXS/RDC refinement. Adapted from Venditti et al. (2015).[142]

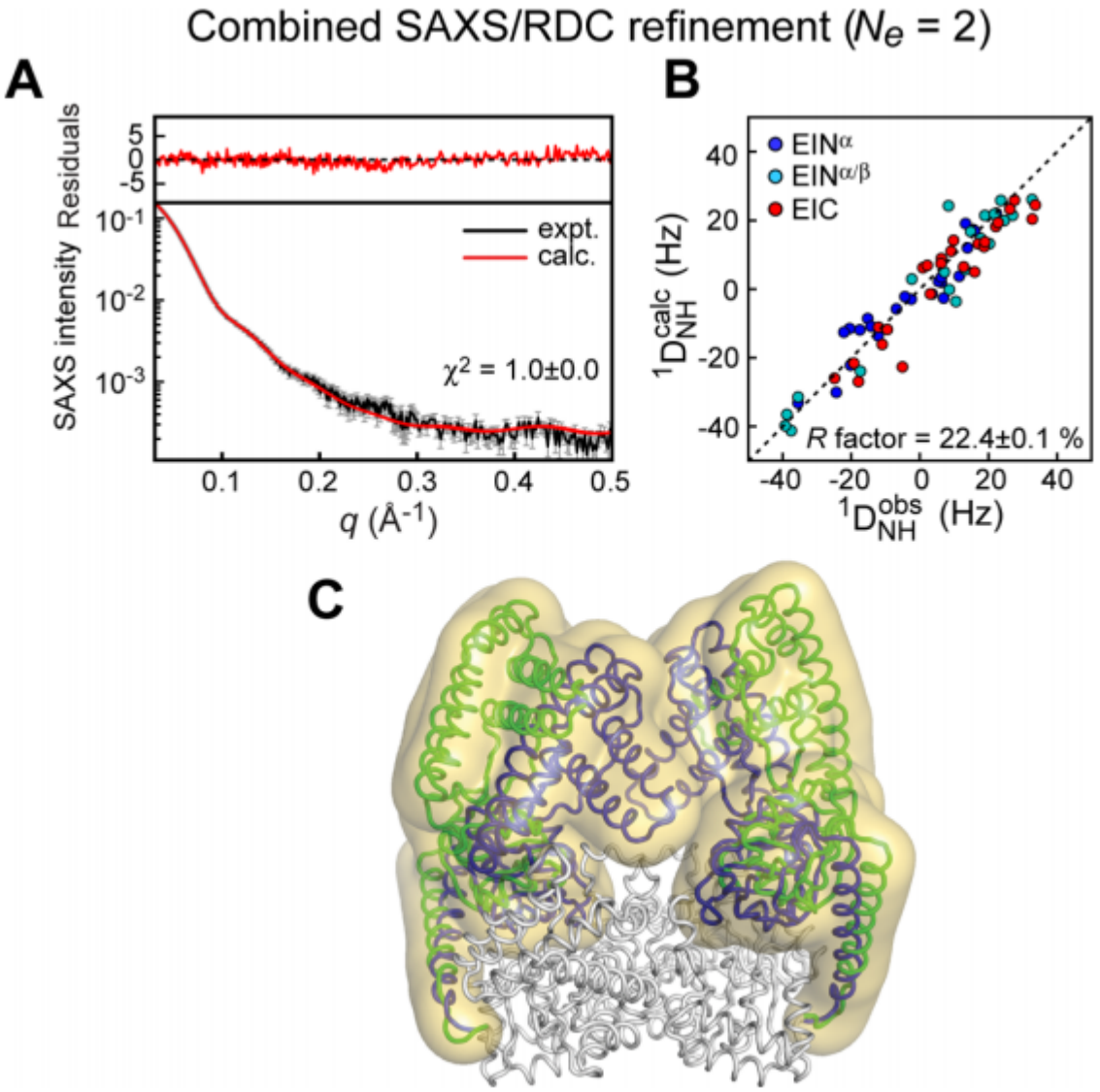


**Figure 3.11.** Combined SAXS/RDC refinement of the EI$^{A}$-PEP complex using a two-member ensemble representation. (A) Agreement between experimental and back calculated SAXS curve. (B) Comparison of observed and calculated RDCs. (C) Structural ensemble obtained for EIA-PEP complex. The overall distribution of EIN relative to EIC is shown as a reweighted atomic probability map[175] plotted at 2% of maximum (transparent yellow surface). Representative structures for the closed and partially closed EI conformations (PDB code 2N5T) are shown as blue and green ribbons, respectively. Adapted from Venditti et al. (2015).[142]

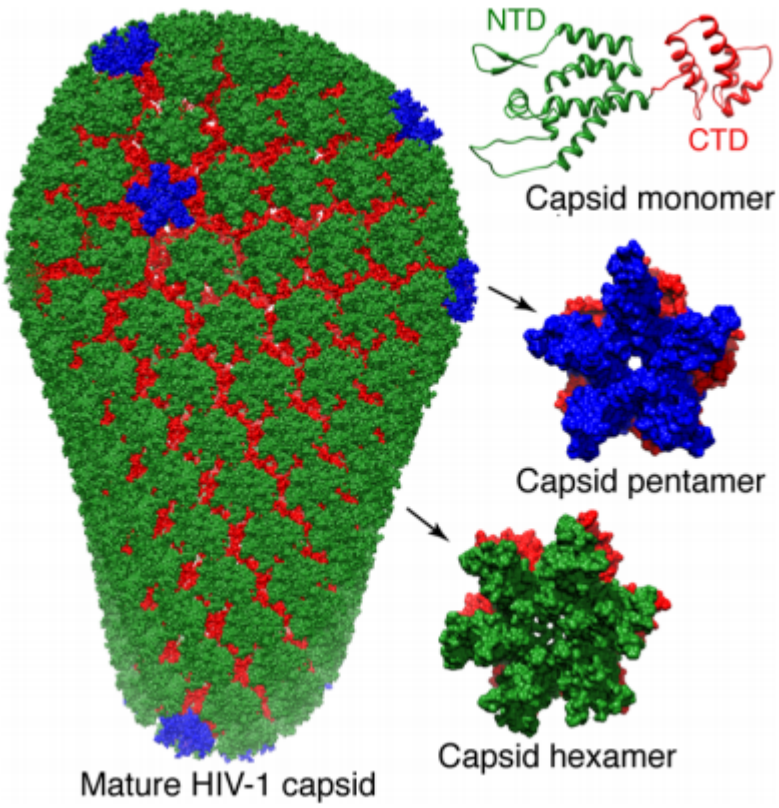


**Figure 3.12**. HIV-1 capsid assembly. The capsid protein comprises N (green) and C (red) terminal domains (top right)[176] The N-terminal domains associate to form either pentamers[166] (middle right with N-terminal domains in blue) or hexamers[169] (bottom right with N-terminal domains in green) which assemble via the C-terminal domain dimers to form a cone comprising ~250 hexamers and exactly 12 pentamers (left).[166] Adapted from Deshmukh et al.[147]

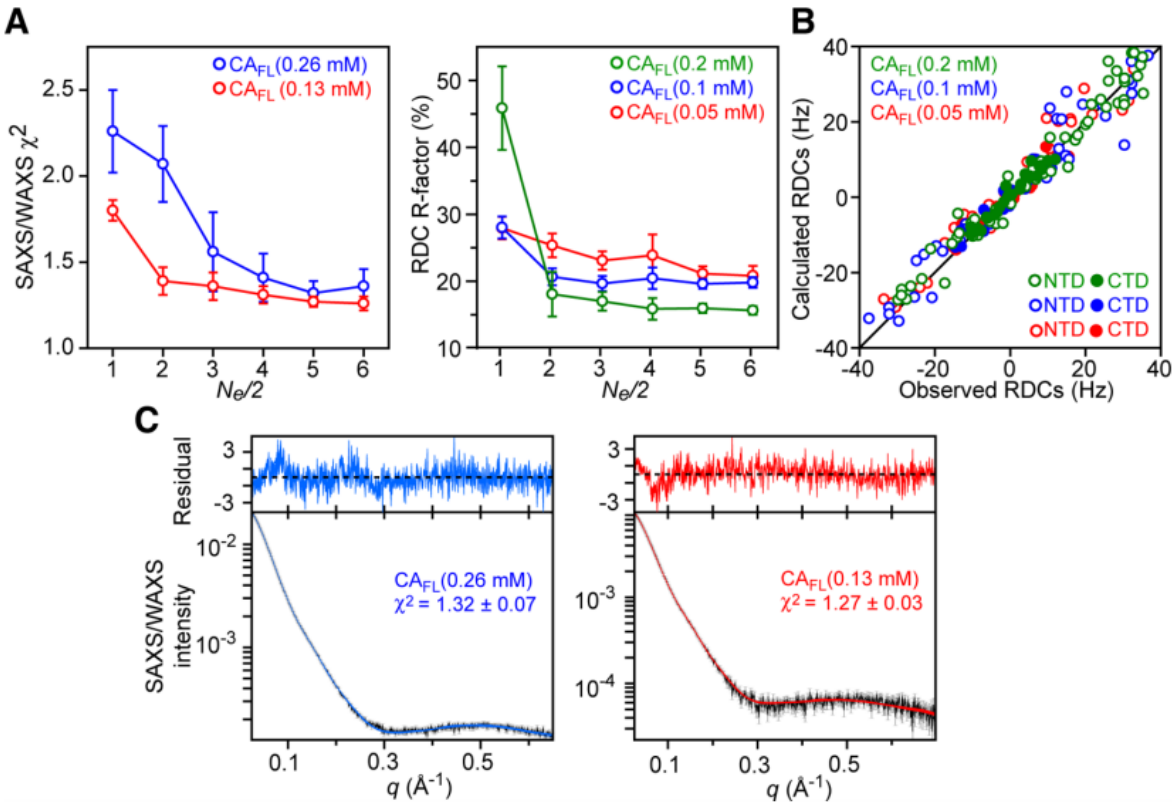


**Figure 3.13.** RDC and SAXS/WAXS driven ensemble simulated annealing refinement of the HIV-1 capsid protein. Both monomer and dimer are included in the calculations and are represented by an equal number of ensemble members ($N_e$/2). RDC and SAXS/WAXS data at several different concentrations are treated simultaneously. (A) SAXS/WAXS χ2 and RDC R-factor as a function of ensemble size. (B) Correlation between observed and back-calculated RDCs based on molecular shape at three concentrations of capsid protein. (C) Agreement between observed (black) and calculated SAXS/WAXS curves at two capsid protein concentrations. The residuals, given by $(I_i^{calc} - I_i^{obs})/I_i^{err}$, are plotted above the curves. Error bars: ± 1 standard deviation. Adapted from Deshmukh et al.[147]

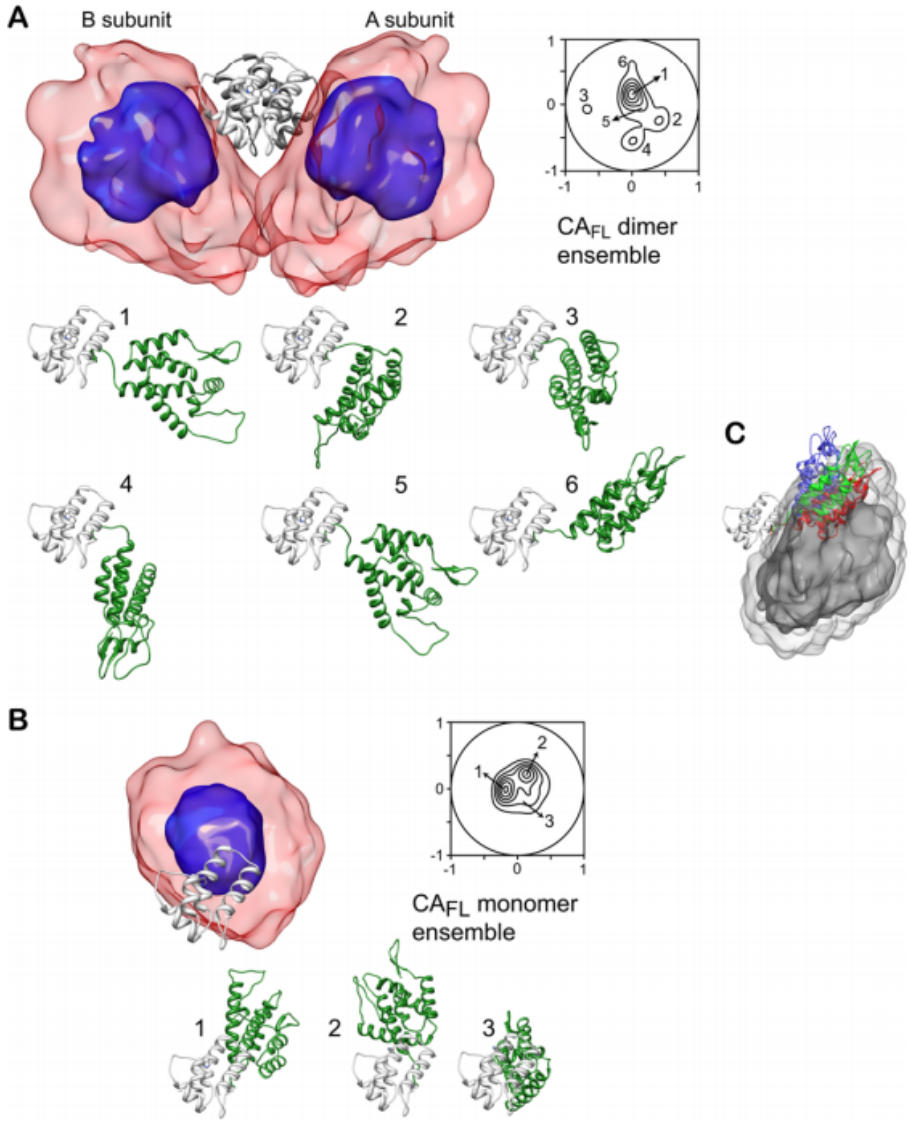


**Figure 3.14.** Structural ensembles calculated for full-length wild type HIV-1 capsid protein. The dimer and monomer ensembles are shown in (A) and (B), respectively. The overall distribution of the N-terminal domain relative to the C-terminal domain (gray ribbon) is displayed as a reweighted atomic probability density map plotted at 50% (blue) and 10% (red transparent) of maximum. Projection contour maps showing the distribution of the position of the centroid of the N-terminal domain relative to the C-terminal domain are also shown. The dimer and monomer ensembles are characterized by six and three main clusters, respectively. (For clarity only a single subunit is shown for the six dimer clusters; the orientation of the C-terminal domain is the same throughout). (C) Position of the N-terminal domain in cluster 6 (red) of the capsid protein dimer ensemble compared to that in the pentamer (green) and hexamer (blue) with the C-terminal domain shown as a gray ribbon, and the atomic probability density map of the N-terminal domain in the capsid dimer plotted at 10% (dark gray) and 2% (light gray) of maximum. Adapted from Deshmukh et al.[147]

# CHAPTER 4

## MECHANISTIC INSIGHT INTO NANOPARTICLE SURFACE ADSORPTION BY SOLUTION NMR SPECROSCOPY IN AN AQUEOUS GEL

Published in *Angewandte Chemie International Edition* in 2017.

Timothy Egner[1], Pranjali Naik[1,2], Nicholas Nelson[1,2], Igor Slowing[1,2], Vincenzo Venditti[1]

[1]Department of Chemistry, Iowa State University
[2] US Department of Energy, Ames Laboratory

### Abstract

Engineering nanoparticle (NP) functions at the molecular level requires a detailed understanding of the dynamic processes occurring at the NP surface. Herein we show that a combination of dark-state exchange saturation transfer (DEST) and relaxation dispersion (RD) NMR experiments on gel-stabilized NP samples enables the accurate determination of the kinetics and thermodynamics of adsorption. We used the former approach to describe the interaction of cholic acid (CA) and phenol (PhOH) with ceria NPs with a diameter of approximately 200nm. Whereas CA formed weak interactions with the NPs, PhOH was tightly bound to the NP surface. Interestingly, we found that the adsorption of PhOH proceeds via an intermediate, weakly bound state in which the small molecule has residual degrees of rotational diffusion. We believe the use of aqueous gels for stabilizing NP samples will increase the applicability of solution NMR methods to the characterization of nanomaterials.

### Introduction, Results, and Discussion

Nanoparticles (NPs) are central to several areas of nanoscience, including optoelectronics,[1] sensing,[2] medicine,[3] and catalysis.[4] NPs are usually made from inorganic and/or

organic materials and owe some of their unique properties to their high surface-area-to-volume ratio, which maximizes the probability of constructive encounters with their target ligands. Thus, the final properties and ultimate application of a NP depend on its surface composition and morphology. Capping agents can be applied to the surface of a NP to increase its solubility and direct its targeting.[5] Surface functionalization with enzymes or other catalysts can be used to produce NPs for industrial or therapeutic applications.[6] Therefore, elucidation of the interactions that occur at the NP surface is crucial for the understanding and engineering of its function at the molecular level.

Although solid-state NMR methods are routinely employed to gain structural insight into small molecule/NP systems,[7] kinetic and thermodynamic aspects of the interaction (i.e. rate constants and populations for the adsorption/desorption equilibrium) are difficult to obtain by such techniques. Thus, there is a critical need to develop new strategies to gain mechanistic insight into the processes occurring at the NP surface. Several solution NMR techniques have been developed to probe and characterize interactions between small, NMR-visible molecules and large, NMR-invisible systems. Among these, dark-state exchange saturation transfer (DEST)[8] and relaxation dispersion (RD)[9] experiments are particularly attractive because they can provide kinetic, thermodynamic, structural, and dynamic information on the interaction.[10] The combined analysis of DEST and RD data acquired at multiple saturation field strengths (for DEST) and spectrometer frequencies (for RD) has been used to provide fundamental insight into the binding of small proteins to large molecular machines,[11] aggregates,[8a,12] biological membranes,[13] and nanoparticles.[14] When using these techniques, two conditions must apply for the interaction to be detected and characterized accurately: The NMR-visible molecule has to exchange between

the free and bound states on the millisecond–second timescale, and the NMR-visible and NMR-invisible molecules have to remain homogeneously suspended in the NMR sample throughout the NMR data acquisition period.[10] Although the first condition is often satisfied by NP–small-molecule interactions, the generation of NP suspensions that remain stable and homogeneous for several hours is often a challenge that limits the applicability of solution NMR methods for the characterization of NPs.

Herein, we show that a 1 wt % agarose gel can be employed to prevent NP sedimentation and enable NMR characterization of the kinetics of the adsorption/desorption of NMR-visible ligands to the surface of NMR-invisible NPs. The ability of agarose to provide a chemically inert matrix for studying ligand adsorption was tested by investigating the interaction between cerium oxide (ceria) NPs and two small molecules with different sizes and chemical properties: phenol (PhOH) and cholic acid (CA; see **Figure 4.S1** in the Supporting Information). The study was conducted with NPs of cubic morphology and approximately 200 nm in size (see **Figure 4.S2**). Ceria NPs were recently shown to actively participate in the palladium-catalyzed hydrogenation of PhOH by coadsorbing the small molecule and the metal on the NP surface, activating the PhOH molecule for the hydrogenation reaction, and possibly participating in the electron-transfer process during catalysis.[6b]

The addition of ceria (1 wt %) to a 10 mM solution of PhOH or CA in 99.9 % $D_2O$ resulted in extensive line broadening of the NMR resonances of the small molecules (Figure **4.1b**), thus indicating that PhOH and CA interact with ceria and exchange between a free and a NP-bound state. However, sedimentation of ceria caused progressive reduction of the amount of NP in the

NMR active volume (**Figure 4.1a**). This local change in the small molecule/NP ratio effectively reduced the fraction of the small molecule bound to the NP that was detected by the NMR experiment and resulted in a progressive shift of the small-molecule NMR spectrum toward the free state (**Figure 4.1b**). The rate of sedimentation depends strongly on the sample conditions. Indeed, the presence of moderate amounts (≥20 mM) of a buffer (Tris-HCl or phosphate at pH 7.4) and/or NaCl resulted in complete sedimentation of ceria in less than 30 min at room temperature. Samples prepared in the absence of a buffer and salts are more stable: Complete ceria sedimentation was observed after approximately 1 h or 1 week when NPs (1 wt %) were added to 10 mM CA (final pH 10.0) or PhOH (final pH 7.0), respectively. However, even in the case of the PhOH sample, for which NP sedimentation could be detected by visual inspection only after 2 days, the effects of sedimentation on the NMR spectra were observed after a much shorter time. Indeed, an approximately 15 % increase in signal intensity was observed about 30 min after the addition of NPs to a 10 mM PhOH solution (**Figure 4.1b**). Such a dramatic change in spectral properties over time hampers the acquisition of extensive and consistent sets of DEST and RD data to study small-molecule binding to ceria.

To limit NP sedimentation, we prepared NMR samples containing 10 mM PhOH or CA and 1 wt % NPs in a 1 wt % agarose gel matrix (see the Supporting Information). With a pore distribution in the 100–500 nm range,[15] the gel matrix kept the NPs in suspension (**Figure 4.1a**) and enabled the acquisition of reproducible NMR spectra for more than 30 days (**Figure 4.1b**). Furthermore, very weak signals originating from the gel matrix were detected in 1D proton spectra (**see Figure 4.S3**), which reduces the chance of signal overlap with the analyte of interest. Importantly, in the absence of NPs, the gel matrix did not perturb the $^{1}$H NMR spectra of PhOH

and CA (**Figure 4.1c,d**), and the $\Delta R_2^{gel}$ values, measured as the increase in transverse relaxation rate ($R_2$) caused by incorporating the small molecules in the agarose gel ($\Delta R_2^{gel} = R_2^{gel} - R_2^{water}$), were near 0 ($\Delta R_2^{gel} \approx 0$) for all NMR signals analyzed (**Figure 4.1e**). These findings indicate the absence of interactions between the investigated small molecules and the gel matrix. On the other hand, $\Delta R_2^{NP}$ values (i.e. the increase in $R_2$ caused by adding NPs to the small molecules in the gel matrix; $R_2^{gel+NP} - R_2^{gel}$) larger than 0 were measured for small-molecule samples prepared with 1 wt % agarose and 1 wt % NPs (**Figure 4.1e**), which reflects the interaction of PhOH and CA with the large and slowly tumbling NPs. The assessment of whether the gel matrix affects the binding properties of the NP surface is an important, but not trivial task. Indeed, in the absence of agarose, the fast sedimentation of ceria does not permit accurate description of the binding of CA and PhOH to the NP surface. However, for the PhOH/NP system (for which NP sedimentation in the absence of the gel matrix is slow), we observed very good agreement between $^1$H-DEST data collected in the absence and in the presence of agarose within about 60 min from sample preparation (**see Figure 4.S4**). This finding suggests that the NP surface displays similar PhOH-binding properties when in water or suspended in the gel matrix.

Binding of PhOH and CA to the surface of ceria NPs was investigated by $^1$H-DEST and $^1$H-RD experiments. **Figure 4.2a** shows examples of DEST profiles measured for the NMR signals of $H_{12}$ and $H_{meta}$ of CA and PhOH, respectively, at different spectrometer frequencies (600 and 800 MHz) and different saturation field strengths (150 and 300 Hz; see **Figures 4.S5** and **4.S6** for the full set of DEST data). The addition of NPs to CA and PhOH samples resulted in broadening of the DEST profiles (**Figure 4.2a**), thus indicating the presence of high-molecular-weight small-

molecule–NP adducts in exchange with the free ligands. Interestingly, only a small increase in the width of the DEST profile was observed for CA upon the addition of ceria, thus suggesting that this molecule interacts weakly with the NPs. In contrast, the very broad "wings" in the DEST profiles measured for PhOH (**see Figure 4.S6**) suggest that the association of PhOH with the NPs has a much longer lifetime.[8a, 12]

The existence of small-molecule–NP interactions was confirmed by $^{1}$H-RD experiments (**Figure 4.2b**; see also **Figures 4.S7** and 4.**S8**) acquired with a modified Carr–Purcell–Meinboom–Gill (CPMG) scheme that suppresses modulations from scalar couplings.[16] If a small molecule is in exchange on the microsecond to millisecond timescale between the free state (characterized by small $R_2$) and the NP-bound state (characterized by large $R_2$), and the free and bound states have different proton chemical shifts, the shape of the $^{1}$H-RD profiles (**Figure 4.2b**) measured for the small-molecule NMR signals will be affected by the exchange. At a high CPMG field, the observed $R_2$ rate will be enhanced by lifetime line broadening ($R_2^{obs} = R_2^{free} + R_{llb}$, in which $R_2^{free}$ is the $R_2$ value in the absence of NPs and $R_{llb}$ is lifetime line broadening).[10] At low CPMG fields, the observed $R_2$ rate will be enhanced by both lifetime line broadening and chemical-exchange contribution to $R_2$ ($R_2^{obs} = R_2^{free} + R_{llb} + R_{ex}$, in which $R_{ex}$ is the exchange contribution to the relaxation rate).[10] Increasing the CPMG field results in the progressive suppression of $R_{ex}$ and a decrease in $R_2$ (the value of which levels off when $R_{ex}$ was fully suppressed by the CPMG field), which introduces a curvature (dispersion) in the RD profile. In the absence of NPs, the $R_2$ rates measured for CA and PhOH were independent of the CPMG field (**Figure 4.2b**), consistent with the absence of small-molecule–NP interactions. Conversely,

samples prepared in the presence of 1 wt % NPs displayed clear relaxation dispersion (i.e. $R_{ex} > 0$) and $R_{llb} > 0$ (i.e. the $R_2$ value at a high CPMG field was higher for the NP-containing sample).

The data from the $^1$H-DEST and $^1$H-RD experiments were analyzed simultaneously by using a two- or three-site exchange model (**Figure 4.3**) in which all experimental observables are described by solutions to the McConnell equations (see the Supporting Information).[10, 17] For each small molecule, global nonlinear least-squares fitting was carried out by optimizing the values of two ($p_F$, $k_{RF}$) and five ($p_F$, $p_R$, $k_{RF}$, $k_{BR}$, $k_{BF}$) global parameters for the two- and three-site exchange, respectively (Figure **4.3**). Proton $R_2$ rates and chemical-shift changes upon binding (Δ) were treated as peak-specific parameters (see Tables S1 and S2 in the Supporting Information for the best-fit parameters for CA and PhOH, respectively). The agreement between the experimental and back-calculated DEST and RD profiles is shown in **Figures 4.2** (see also **Figures 4.S5–4.S8**). CA binding to ceria is fully described by a two-site exchange model involving the free state (population, $p_F$≈96 %; $R_2$(average)≈3 $s^{-1}$) and a NP-associated state with a population of approximately 4 % and an average $R_2$ value of approximately 60 $s^{-1}$. By using the Stoke law and approximating the NP to a sphere of 100 nm radius, we calculated an effective correlation time ($\tau_{eff} = 1/[1/\tau_c + k_{FR}^{app} + k_{RF}]$,[13b] in which $\tau_c$ is the correlation time of the NP) for a rigid CA–NP adduct in water at 25 °C of approximately 150 μs, which would translate into an $R_2$ value in the tens of thousands per second for a pair of isolated $^1$H nuclei located 2.5 Å apart that undergo dipolar relaxation. The fact that CA in the NP-associated state has an $R_2$ value that is three orders of magnitude smaller than the $R_2$ value simulated for a rigid CA–NP adduct indicates that residual degrees of rotational diffusion increase the mobility of the small molecule associated with the NP surface, thus decreasing its $\tau_{eff}$. In this weakly associated state, CA does not form a rigid

adduct with the NP, but its rotational diffusion is only marginally restricted as compared to the free state (**Figure 4.3**). We refer to such "dynamic" small-molecule–NP association as the "restricted state", as opposed to a rigidly bound species. The ζ potential measured for the NP at pH 10.0 (corresponding to the pH value of the NMR sample containing CA and NP) is −1.6 mV (see the Supporting Information), thus suggesting that formation of the restricted state is driven by hydrophobic and/or hydrogen-bonding interactions between CA and the NP surface.

$^{1}$H-DEST data acquired for PhOH in the presence of 1 wt % NPs can not be fit by using the two-site exchange model employed for CA (see **Figures 4.S6**), but a three-site exchange has to be invoked to fully account for the experimental DEST and RD profiles (**Figure 4.2**; see also Figures S6 and S8). The three states have average $R_2$ values of approximately 0.3, 90, and 10 000 $s^{-1}$, and represent the free, restricted, and rigidly bound states, respectively (**Figure 4.3**; see also **Table 4.S2**). Interestingly, analysis of the DEST and RD data reveals that there is no direct communication between the free and rigidly bound states (i.e. $k_{FB}^{app} = k_{BF} = 0$; see **Table 4.S2**), and that the restricted state is a weakly associated intermediate on the pathway describing PhOH binding to ceria. Given the neutral electric charge of the NP surface at pH 7.0 (ζ potential: ca. 1.7 mV; see the Supporting Information), formation of the restricted state is most likely driven by hydrophobic and/or hydrogen-bonding interactions between PhOH and NPs. On the other hand, the tightly bound state might be the result of the interaction of the phenolic oxygen atom with a cerium atom from the NP surface.[6b] Establishment of this rigidly associated PhOH/NP adduct mediates PhOH activation and is an essential step in PhOH hydrogenation catalyzed by Pd supported on ceria NPs.[6b] However, assessment of the role played by the equilibrium between

the restricted and rigidly bound state in the heterogeneous catalytic hydrogenation of PhOH is beyond the scope of the present study and will be the subject of future investigations.

In summary, we have shown that the preparation of nanoparticle samples in aqueous gels is a convenient strategy to avoid nanoparticle sedimentation for extended periods of time (up to several months), and allows the acquisition of extensive sets of solution NMR experiments for investigation of the thermodynamics and kinetics of binding to a NP surface. Interestingly, by using this approach we detected a motionally restricted intermediate state in PhOH binding to ceria NPs, thus demonstrating that the method is capable of providing detailed mechanistic insight into surface adsorption processes. Agarose seems an ideal choice as the gel matrix, since it is easy to prepare and does not introduce additional NMR signals. Most importantly, our data indicate that the agarose gel is an inert matrix that does not have detectable interactions with organic compounds (i.e. CA and PhOH) or proteins (i.e. ubiquitin; see **Figure 4.S10**), and does not perturb small-molecule binding to the surface of the ceria NPs tested in this study. The preparation of homogeneous and stable samples is crucial for highly quantitative NMR measurements, especially for insensitive experiments that require extensive signal averaging (i.e. NMR experiments involving nuclei with a low gyromagnetic ratio). We expect the use of aqueous gels to expand considerably the portfolio of NMR methods applicable to the characterization of NPs in solution, and to become a very valuable tool in understanding and guiding NP engineering.

**References**

[1] A. K. Podborska,M. F. Oszajca, S.A.Gaweda, K. T. Szacitowski, IET Circuits Devices Syst. 2011, 5, 103 –114.

[2] S. M. Ng, M. Koneswaran, R. Narayanaswamy, RSC Adv. 2016,6,21624–21661.

[3] S. K. Murthy ,Int. J. Nanomed.2007,2,129 –141.

[4] Y. Xia,H. Yang,C.T.Campbell,Acc.Chem.Res.2013,46,1671 –1672.

[5] a) B. I. Kharisov, H.V.R.Dias, O.V.Kharissova, A. Vazquez,Y.Pena ,I. Gómez, RSC Adv. 2014, 4,45354–45381; b)S.Liang, Q.Zhou, M. Wang, Y.Zhu, Q. Wu, X.Yang, Int. J. Nanomed. 2015, 10, 2325 –2333; c)O.Veiseh, J. W. Gunn, M. Zhang, Adv. Drug Delivery Rev. 2010,62,284 – 304.

[6] a) R. T. Hill,J . B. Shear,Anal.Chem.2006,78,7022 –7026; b) N. C. Nelson, J. S. Manzano, A.D.Sadow, S.H.Overbury, I.I.Slowing, ACS Catal. 2015,5,2051 –2061; c) K.E.Sapsford, W. R. Algar, L.Berti, K. B. Gemmill,B. J. Casey, E.Oh, M. H. Stewart, I. L. Medintz, Chem. Rev. 2013,113,1904 –2074.

[7] a) T. M. Alam, J. E. Jenkinsin Advanced Aspects of Spectroscopy (Ed.:M.A. Farrukh), INTECH, 2012,pp. 279 –306; b)T.Blasco,Chem.Soc.Rev.2010,39,4685 –4702.

[8] a) N. L. Fawzi, J. Ying, D.A.Torchia,G. M. Clore, J. Am.Chem.Soc.2010,132,9948 –9951; b)N.L.Fawzi, J. Ying,D.A.Torchia, G. M. Clore,Nat.Protoc.2012,7,1523 –1533.

[9] F. A. Mulder, N.R.Skrynnikov, B.Hon, F. W. Dahlquist, L. E. Kay, J. Am. Chem. Soc. 2001,123,967 –975.

[10] N. J. Anthis, G.M.Clore, Q. Rev.Biophys.2015,48,35–116.

[11] a) D. S. Libich, N. L. Fawzi, J. Ying, G.M.Clore, Proc. Natl. Acad. Sci. USA 2013, 110,11361–11366; b)D.S. Libich, V. Tugarinov, G.M.Clore, Proc. Natl. Acad. Sci. USA 2015, 112,8817–8823.

[12] N. L. Fawzi, J. Ying, R.Ghirlando, D.A.Torchia, G. M. Clore,Nature2011,480,268 –272

[13] a) G. Fusco, T.Pape, A.D.Stephens, P.Mahou,A. R. Costa, C. F. Kaminski, G. S. Kaminski Schierle, M.Vendruscolo, G.Veglia,C. M. Dobson,A. De Simone, Nat. Commun. 2016,7,12563; b)A.Ceccon, V.Tugarinov, A.Bax, G. M. Clore,J. Am. Chem. Soc. 2016, 138,5789 –5792; c) A.Ceccon, M. Lelli, M.D Onofrio, H. Molinari, M. Assfalg, J. Am. Chem. Soc. 2014, 136,13158–13161.

[14] a) S. Zanzoni, A. Ceccon,M. Assfalg, R.K.Singh, D. Fushman, M. D'Onofrio, Nanoscale 2015,7,7197 –7205; b) S.Zanzoni, M.Pedroni, M.D'Onofrio, A.Speghini, M.Assfalg,J. Am. Chem. Soc. 2016, 138,72–75.

[15] B. H. Zimm, S. D. Levene, Q. Rev.Biophys.1992,25,171 –204.

[16] J. A. Aguilar, M.Nilsson, G. Bodenhausen, G. A. Morris, Chem. Commun. 2012, 48,811 –813.

[17] M. Helgstrand, T. Hard, P. Allard, J. Biomol. NMR 2000,18,49–63.

**Figures**

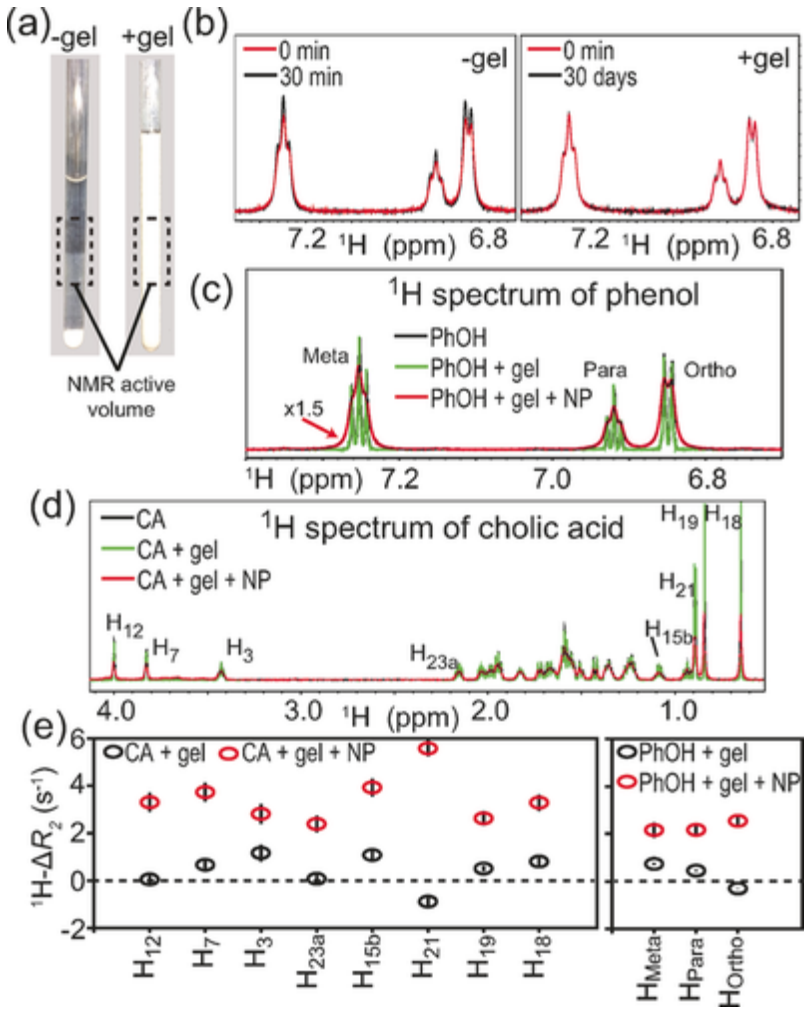


**Figure 4.1.** a) NMR samples containing 10 mM PhOH and 1 wt % NPs in the absence (left) and in the presence (right) of the gel matrix. Pictures were taken after 1 week and 6 months from sample preparation, respectively. b) $^{1}$H NMR spectra of PhOH in the presence of 1 wt % NPs were acquired 0 (red) and 30 min (black) after sample preparation. Samples were prepared in the absence (left) and in the presence (right) of the agarose gel matrix. c, d) $^{1}$H NMR spectra of PhOH (c) and CA (d) in the absence of a gel (black), in the presence of the gel matrix (green), and in the presence of the gel matrix and 1 wt % NPs (red). The spectrum acquired for PhOH in the presence of agarose and NPs is multiplied by 1.5 in the figure. NMR signals used for the analysis of $^{1}$H-DEST and $^{1}$H-RD profiles are labeled according to the numbering scheme reported Figure S1 of the Supporting Information. e) Increase in the $^{1}$H $R_2$ rate ($\Delta R_2$) upon addition of the agarose gel matrix ($\Delta R_2^{gel}$ , black circles) and of the agarose gel matrix and 1 wt % NPs ($\Delta R_2^{NP}$ , red circles) to 10 mM CA (left) and PhOH (right).

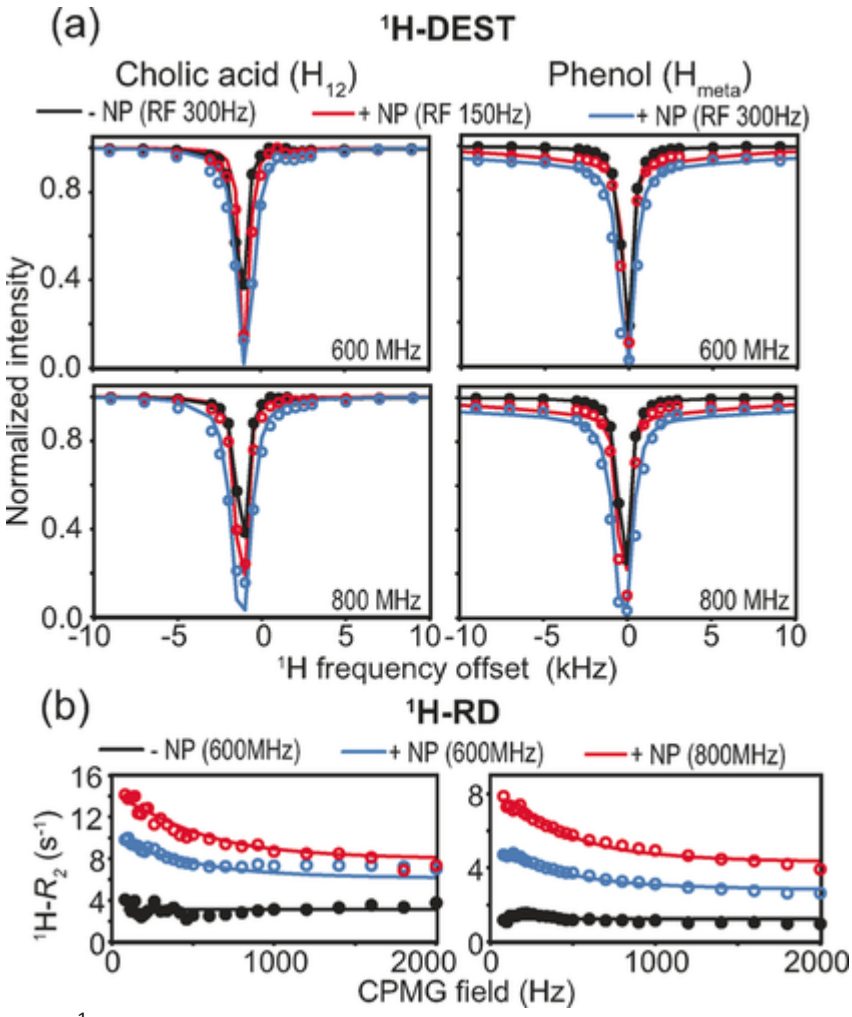


**Figure 4.2.** a) $^{1}$H-DEST profiles for the $H_{12}$ and $H_{meta}$ positions of CA (left) and PhOH (right), respectively. Data were acquired with 600 (top) and 800 MHz (bottom) spectrometers. Experimental data are shown as circles. DEST profiles simulated from the best-fit parameters (see main text) are shown as solid lines. Color code: 0 wt % NPs and 300 Hz saturation field, black; 1 wt % NPs and 150 Hz saturation field, red; 1 wt % NPs and 300 Hz saturation field, blue. See Figures S5 and S6 for the complete set of $^{1}$H-DEST profiles. b) $^{1}$H-RD profiles for the $H_{12}$ and $H_{meta}$ positions of CA (left) and PhOH (right), respectively. Experimental data are shown as circles. RD profiles simulated from the best-fit parameters (see main text) are shown as solid lines. Color code: 0 wt % NPs and 600 MHz spectrometer, black; 1 wt % NPs and 600 MHz spectrometer, blue; 1 wt % NPs and 800 MHz spectrometer, red. See Figures S7 and S8 for the complete set of $^{1}$H-RD profiles.

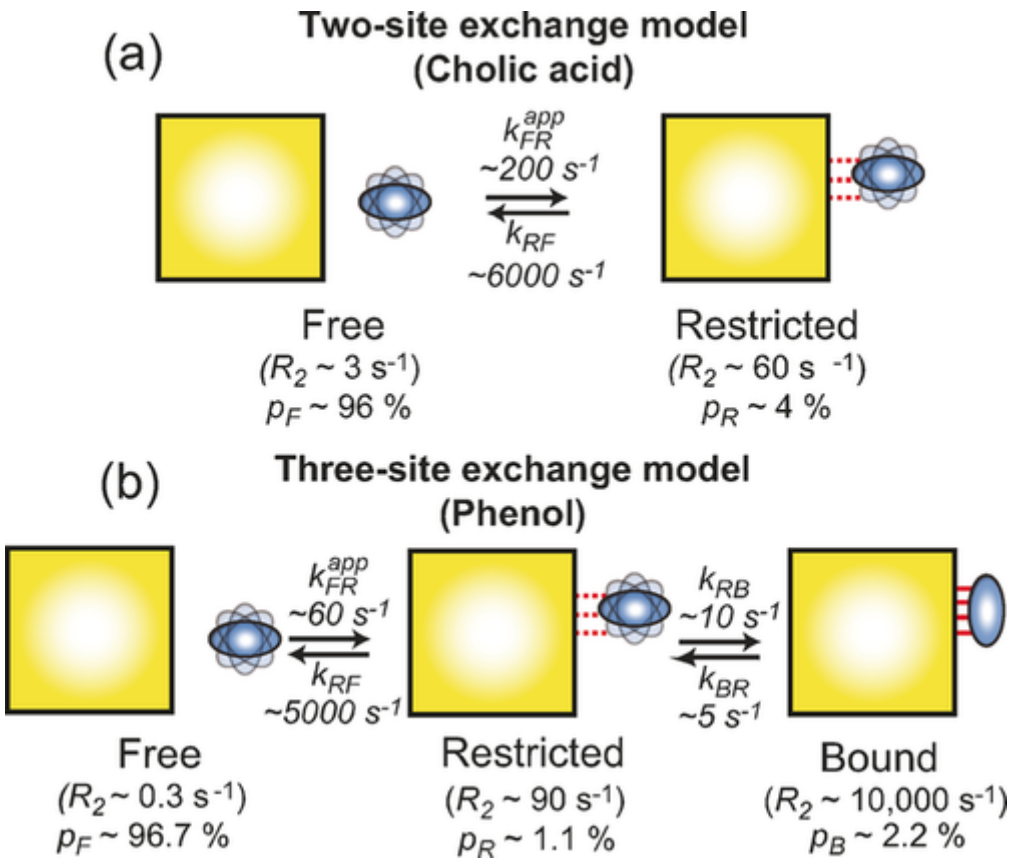


**Figure 4.3.** a) Interpretation of the $^1$H-DEST and $^1$H-RD data measured for CA in the presence of 1 wt % NPs by using a two-site exchange model. The NP-associated state forms a weak interaction with ceria (see main text) and is referred to as the "restricted state". b) Interpretation of the $^1$H-DEST and $^1$H-RD data measured for PhOH in the presence of 1 wt % NPs by using a three-site exchange model. The NP-associated state exists as a combination of restricted and bound states (see main text). The NP is represented as a yellow square. The small molecule is shown as a blue oblate ellipsoid. Transparent blue oblate ellipsoids are used to indicate rotational diffusion of the small molecule. Solid and dashed red lines represent strong and weak interactions, respectively. Rate constants for each equilibrium ($k_{xy}$ , in which *x* and *y* are the states in thermodynamic equilibrium; $k_{xy}^{app}$ indicates the apparent rate constant) and populations of each state ($p_F$, $p_R$, and $p_B$ for the free, restricted, and bound states, respectively) were fit as global parameters. $R_2$ rates for each state and chemical-shift changes associated with each equilibrium were fit as peak-specific parameters (see the Supporting Information). Best-fit parameters are shown. For the $R_2$ rate of each state, the average over the analyzed NMR peaks is reported. The fit reveals that the free and bound states of PhOH are not kinetically connected ($k_{FB}^{app} = k_{BF} = 0$ ). See Tables S1 and S2 for a full report of the fitted parameters for CA and PhOH, respectively; see Figure S10 for results of the individual, atom-based fits of the $^1$H-DEST and $^1$H-RD data.

## Supplemental Information

***Synthesis of ceria nanoparticles***. Cerium (III) nitrate hexahydrate (99.99%) and sodium hydroxide (reagent grade) were purchased from Sigma-Aldrich and used without further purification, deionized water (18 MΩ-cm) was obtained in house from a Barnstead e-pure system. Ceria nanocubes were synthesized following a reported method[1]. In brief, a solution of $Ce(NO_3)_3 \cdot 6H_2O$ (5.0 mL, 0.4 M) was mixed with aqueous NaOH (35.0 mL, 6.9 M) under constant stirring for 30 min at room temperature to produce a milky slurry. The mixture was then transferred into a Teflon lined autoclave and placed for 24 h into an oven preheated at 180 °C. The precipitate was collected by centrifugation, washed several times with deionized water and ethanol and dried at 60 °C overnight. The dried sample was then calcined at 450 °C for 4 h at a rate 2 °C/min.

***Preparation of NMR samples.*** Uniformly 15N-labeled ubiquitin was prepared as described previously.[2] Phenol (PhOH), cholic acid (CA), and agarose were purchased from Sigma Aldrich and used without further purification. PhOH and CA NMR samples were prepared in 99.9% D2O. Ubiquitin samples were prepared in 30 mM phosphate buffer (pH 6.5), 1 mM EDTA, and 10 v/v % D2O. PhOH and and CA were 10 mM in each prepared sample. The concentration of ubiquitin was 1 mM. Samples that contained ceria nanocubes in the absence of gel were prepared adding 1 wt % ceria nanocubes to a 10 mM solution of ligand (PhOH or CA) in D2O. The mixture was sonicated for 60 seconds (1 s/2 s on/off duty cycle) using a 500 W sonicator set to 25 % amplitude. The suspension was transferred to an NMR tube and measurements were acquired immediately. Samples that contained ceria nanocubes in the presence of gel were prepared mixing 1 wt % ceria

nanocubes and 1 wt % agarose in D2O (for PhOH or CA) or in the ubiquitin NMR buffer. The mixture was sonicated for 60 seconds (1 s/2 s on/off duty cycle) using a 500 W sonicator set to 25 % amplitude, and then microwaved for 30-60 seconds to bring the mixture to a boil. The container was transferred to a 50 °C water bath. Ligands were added to their final concentrations. The warm solution was transferred to an NMR tube and allowed to cool at room temperature for 30 minutes. The reproducibility of the sample preparation was evaluated by collecting NMR experiments on two different preparations of the same NMR sample (**Figure 4.S11**).

***NMR spectroscopy.*** All NMR data were recorded at 25 °C on Bruker 800 and 600 MHz spectrometers equipped with triple resonance z-gradient cryoprobes. 1H-DEST experiments were measured at two saturation fields (150 and 300 Hz) using the pulse scheme in **Figure 4.S12**. A set of 29 offsets from the 1H carrier frequency were used (ranging from -25 kHz to +25 kHz) at both saturation field strengths and supplemented by three control experiments with offsets of -25, 0 and +25 kHz and a continuous-wave (CW) field of 0 Hz that were used for normalization and error evaluation. The saturation field was applied for 1 s. 15N-DEST experiments were measured at two saturation fields (250 and 500 Hz) using the pulse scheme described by Clore and co-workers.[3] A set of 15 offsets from the 15N carrier frequency were used (ranging from -35 kHz to +35 kHz) at both saturation field strengths and supplemented by two control experiments with offsets of -15 and +15 kHz and a continuous-wave (CW) field of 0 Hz that were used for normalization and error evaluation. The saturation field was applied for 0.7 s. 1H-RD experiments were measured using a modified CPMG sequence that suppresses modulation from scalar couplings.[4] Experiments were performed at 800 and 600 MHz with a fixed relaxation delay (100 ms) but a changing number of refocusing pulses to achieve different effective CPMG fields.[5]

Repeats at three CPMG field were acquired for error evaluation. 15N-RD experiments were measured using 15N-CPMG scheme with amide proton decoupling.[6] Experiments were performed at 800 and 600 MHz with a fixed relaxation delay (50 ms) but a changing number of refocusing pulses to achieve different effective CPMG fields.[5] Proton longitudinal relaxation rates ($^1$H-$R_1$) for PhOH and CA in the absence and in the presence of nanoparticles were measured using the inversion recovery pulse scheme. NMR spectra were processed using Mnova NMR (http://mestrelab.com/software/mnova/nmr/) and NMRPipe.[7] Spectra were analyzed using Mnova NMR and SPARKY (http://www.cgl.ucsf.edu/home/sparky).

***Dynamic Light Scattering.*** The hydrodynamic diameter distribution of the nanocubes was measured on a Malvern Zetasizer Nano ZS90 instrument with 90° scattering optics. NP suspensions were prepared by sonicating small amount of sample in deionized water (~0.1 mg/mL). Samples of CA were prepared by diluting a concentrated stock of CA pH 10.0 into deionized water (final pH = 10.0). Samples (0.1 mL) were placed in a low volume polystyrene cuvette and general purpose analysis model. Measurements were performed in triplicate.

***Transmission Electron Microscopy.*** Cerium nanoparticle imaging was performed in a FEI Tecnai G2 F20 field emission transmission electron microscope (TEM) operating at 200 kV. Samples were prepared by placing 3-4 drops of ethanol suspension (~0.1 mg/mL) onto lacey-carbon-coated copper grids and dried in air.

***ζ-potential.*** The ζ-potential of 0.1 wt % ceria nanocube suspensions were measured on a Malvern Zetasizer Nano ZS90 instrument. Measurements were performed at 25°C at either pH 7.0 or pH 10.0 to simulate NMR sample conditions for PhOH and CA, respectively. pH was adjusted by

addition of 1 µl aliquots of 0.1 or 1.0 M NaOH directly to the NP sample. Measurements were performed in triplicate to evaluate the experimental error. ζ –potentials of 1.7 ± 6.9 and -1.6 ± 13.5 mV were measured for the ceria nanocubes at pH 7.0 and 10.0, respectively.

***Combined Fitting of 1H-DEST and 1H-RD data.*** The 1H-DEST and 1H-RD data collected at two spectrometer frequencies were globally fit using the two- or three-site exchange process described in the main text (see **Figure 4.S13** for definition of the parameters). Fitting was performed using an in-house Matlab program that minimizes the following χ2 function:

$$\chi^2 = 0.1\chi^2_{RD} + \chi^2_{DEST} \quad \text{(S1)}$$

Where

$$\chi^2_{RD} = \sum_i \sum_j \sum_k \left(R_2^{obs,i,j,k} - R_2^{calc,i,j,k}\right)^2 \quad \text{(S2)}$$

$$\chi^2_{C-DEST} = \sum_i \sum_j \sum_n \sum_p \left(I^{obs,i,j,l,m} - I^{calc,i,j,l,m}\right)^2 \quad \text{(S3)}$$

Here, the indices i, j, k, n and p refer to the peak index, the spectrometer frequency (600 and 800 MHz), the CPMG field, the saturation field strength and the DEST frequency offset, respectively; the superscripts obs and calc stand for experimentally observed and calculated, respectively; I is the normalized intensity of the DEST experiment; the 0.1 factor accounts for the fact that the experimental RD values are ~10 larger than the DEST ones (compare **Figures 4.S5-S8**).

R2's for any specific value of i, j and k were calculated as:

$$R_2^{calc,i,j,k} = \frac{\ln(M(t_1)/M(t_2))}{t_2 - t_1} \quad \text{(S4)}$$

Here, t2 corresponds to the constant CPMG delay, and t1 is set to 0 (see NMR data acquisition).

The transverse magnetizations M(t1) and M(t2) are given by:

$$M(t) = (AA^*A^*A)^n M(0) \quad \text{(S5)}$$

where $M=[M_F;\ M_R]^T$ and $M=[M_F;\ M_R;\ M_B]^T$ for the two- and three-site exchange processes, respectively (see **Figure 4.S13**); $M_i$ denotes the transverse magnetization of state i; T denotes transposition; $A = \exp(-R\tau_{CP}/2)$; A* is the complex conjugate of A; n is the number of CPMG cycles used; $\tau_{CP}$ is the time interval between two consecutive 180° pulses in the CPMG pulse train; and R=RCS+Rrel+Rex, where

$$R^{CS} = i\begin{bmatrix} 0 & 0 \\ 0 & -\Delta\omega^{FR} \end{bmatrix} \quad \text{(S6)}$$

$$R^{CS} = i\begin{bmatrix} R_2^F & 0 \\ 0 & R_2^R \end{bmatrix} \quad \text{(S7)}$$

$$R^{CS} = i\begin{bmatrix} k_{FR}^{app} & -k_{RF} \\ -k_{FR}^{app} & k_{RF} \end{bmatrix} \quad \text{(S8)}$$

or

$$R^{CS} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -\Delta\omega^{FR} & 0 \\ 0 & 0 & -\Delta\omega^{FB} \end{bmatrix} \quad \text{(S9)}$$

$$R^{rel} = \begin{bmatrix} R_2^F & 0 & 0 \\ 0 & R_2^R & 0 \\ 0 & 0 & R_2^B \end{bmatrix} \quad \text{(S10)}$$

$$R^{ex} = \begin{bmatrix} k_{FR}^{app} + k_{FB}^{app} & -k_{RF} & -k_{BF} \\ -k_{FR}^{app} & k_{RF} + k_{RB} & -k_{BR} \\ -k_{FB}^{app} & -k_{RB} & k_{BR} + k_{BF} \end{bmatrix} \quad \text{(S11)}$$

for the two- or three-site exchange processes in **Figure 4.S13**, respectively. Here, $\Delta\omega_{i,j}$ is the difference between chemical shifts (in rad $s^{-1}$) of states i and j ($\omega_i - \omega_j$), and is the transverse relaxation of state i in the absence of exchange.

The DEST experimental observable, $I_{obs}$, is the ratio of the signal intensity as a function of saturation offset and saturation field to that without saturation. Here, the NMR signal intensities at each combination of saturation offset Ω, saturation field $\omega_x$, peak index, and spectrometer frequency were obtained by solving the following equation by matrix exponentiation:

$$\frac{d}{dt}\begin{bmatrix} E/2 \\ I_x^A \\ I_y^A \\ I_z^A \\ I_x^B \\ I_y^B \\ I_z^B \\ I_x^C \\ I_y^C \\ I_z^C \end{bmatrix} = -\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & R_2^{A*} & \Omega^A & -\omega_y & -k_{BA} & 0 & 0 & -k_{CA} & 0 & 0 \\ 0 & -\Omega^A & R_2^{A*} & \omega_x & 0 & -k_{BA} & 0 & 0 & -k_{CA} & 0 \\ -2\Theta^A & \omega_y & -\omega_x & R_1^{A*} & 0 & 0 & -k_{BA} & 0 & 0 & -k_{CA} \\ 0 & -k_{AB}^{app} & 0 & 0 & R_2^{B*} & \Omega^B & -\omega_y & -k_{CB} & 0 & 0 \\ 0 & 0 & -k_{AB}^{app} & 0 & -\Omega^B & R_2^{B*} & \omega_x & 0 & -k_{CB} & 0 \\ -2\Theta^B & 0 & 0 & -k_{AB}^{app} & \omega_y & -\omega_x & R_1^{B*} & 0 & 0 & -k_{CB} \\ 0 & -k_{AC}^{app} & 0 & 0 & -k_{BC} & 0 & 0 & R_2^{C*} & \Omega^C & -\omega_y \\ 0 & 0 & -k_{AC}^{app} & 0 & 0 & -k_{BC} & 0 & -\Omega^C & R_2^{C*} & \omega_x \\ -2\Theta^C & 0 & 0 & -k_{AC}^{app} & 0 & 0 & -k_{BC} & \omega_y & -\omega_x & R_1^{C*} \end{bmatrix} \quad \text{(S13)}$$

for the two- and three-site exchange processes in **Figure 4.3** (see main text), respectively. Here, $\Theta^p = \mathrm{R}_1^p I_z^p$, $R_2^{p*} = R_2^p + k_{pq} + k_{pr}$, $R_1^{p*} = R_1^p + k_{pq} + k_{pr}$; I represents the magnetization of a 1H nucleus in the rotating frame; is the difference between the resonant frequency of state n and the frequency of the applied saturation field; ω is strength of the continuous-wave (CW) saturation field about the given axis (x or y); E is unity; and , where is the equilibrium longitudinal magnetization of state n. The value of Icacl was computed by normalizing the signal intensity calculated by equation S12 or S13 for the two- or three-site exchange models, respectively, to the solution without saturation.

The optimization was run using two ($p_F$, $k_{RF}$) and five ($p_F$, $p_R$, $k_{RF}$, $k_{BR}$, $k_{BF}$) global parameters for the two- and three-site exchange, respectively. Remaining values of populations and rate constants were calculated using the following equations:

$$p_R = 1 - p_F \tag{S14}$$

$$k_{FR}^{app} = k_{RF}\, {p_R}/{p_F} \tag{S15}$$

or

$$p_R = 1 - p_F - p_R \tag{S16}$$

$$k_{FR}^{app} = k_{RF}\, {p_R}/{p_F} \tag{S17}$$

$$k_{RB} = k_{BR}\, {p_B}/{p_R} \tag{S18}$$

$$k_{FB}^{app} = k_{BF}\, {p_B}/{p_F} \tag{S19}$$

for the two- or three-site exchange, respectively.

Error of the fitted parameters was computed by Monte Carlo simulation of synthetic datasets from the estimated experimental error (see NMR spectroscopy section). Convergence of the fits was evaluated by using a grid search procedure (**Figure 4.S14**).

**References**

[1] H.-X. Mai, L.-D. Sun, Y.-W. Zhang, R. Si, W. Feng, H.-P. Zhang, H.-C. Liu, C.-H. Yan, *J Phys Chem B* **2005**, *109*, 24380-24385.

[2] R. Varadan, M. Assfalg, A. Haririnia, S. Raasi, C. Pickart, D. Fushman, *J Biol Chem* **2004**, *279*, 7055-7063.

[3] N. L. Fawzi, J. Ying, D. A. Torchia, G. M. Clore, *J Am Chem Soc* **2010**, *132*, 9948-9951.

[4] J. A. Aguilar, M. Nilsson, G. Bodenhausen, G. A. Morris, *Chem Commun (Camb)* **2012**, *48*, 811-813.

[5] F. A. Mulder, N. R. Skrynnikov, B. Hon, F. W. Dahlquist, L. E. Kay, *J Am Chem Soc* **2001**, *123*, 967-975.

[6] D. F. Hansen, P. Vallurupalli, L. E. Kay, *J Phys Chem B* **2008**, *112*, 5898-5904.

[7] F. Delaglio, S. Grzesiek, G. W. Vuister, G. Zhu, J. Pfeifer, A. Bax, *J Biomol NMR* **1995**, *6*, 277-293.

[8] A. F. Hofmann, A. Roda, *J Lipid Res* **1984**, *25*, 1477-1489.

[9] N. L. Fawzi, J. Ying, D. A. Torchia, G. M. Clore, *Nat Protoc* **2012**, *7*, 1523-1533.

[10] N. Bloembergen, E. M. Purcell, R. V. Pound, *Phys Rev* **1948**, *73*, 679-712.

[11] N. J. Anthis, G. M. Clore, *Q Rev Biophys* **2015**, *48*, 35-116.

**Supplemental Figures**

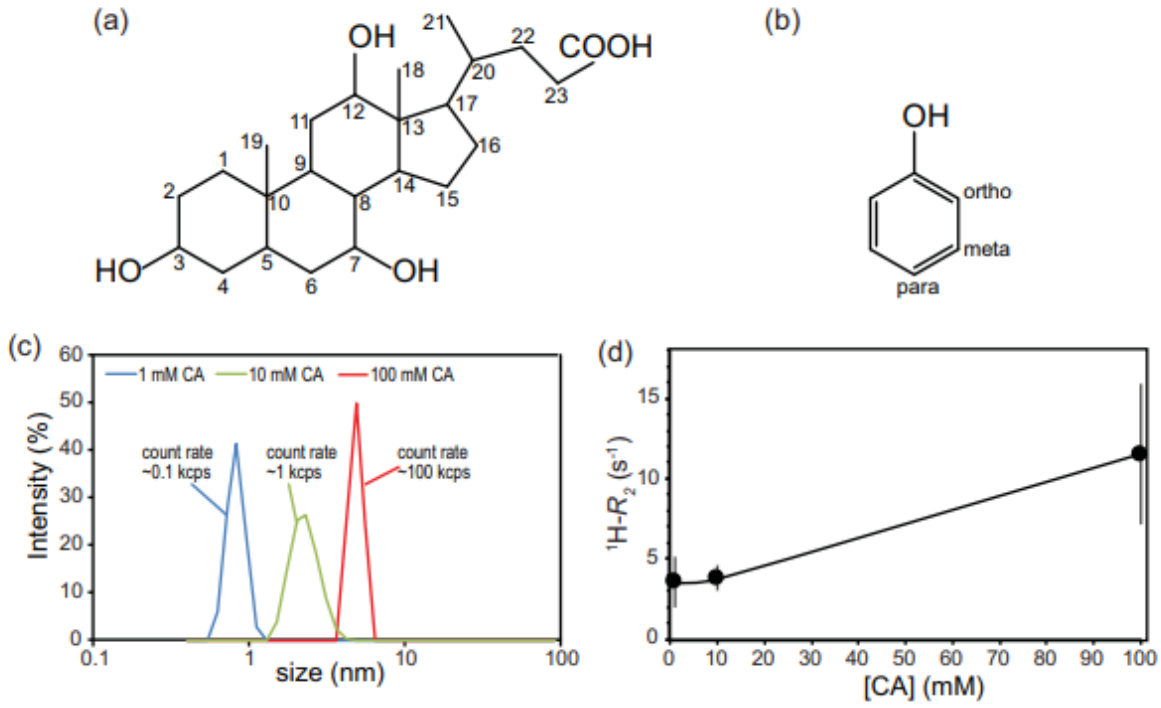


**Figure 4.1S.** Molecular structures of (a) Cholic acid (CA) and (b) Phenol (PhOH). The critical micellar concentration of CA (13 mM)[8] is close to the concentration used for NMR data acquisition (10 mM). DLS (c) and transverse relaxation NMR (d) experiments acquired at multiple concentrations of CA indicate the absence of large CA species at 10 mM (the count rate for the DLS measurement at 10 mM CA is close to the detection limit, and the 1 H-R2's measured for CA at 1 and 10 mM are the same within experimental error). Therefore, CA is monodisperse at the experimental conditions used to probe its interaction with ceria.

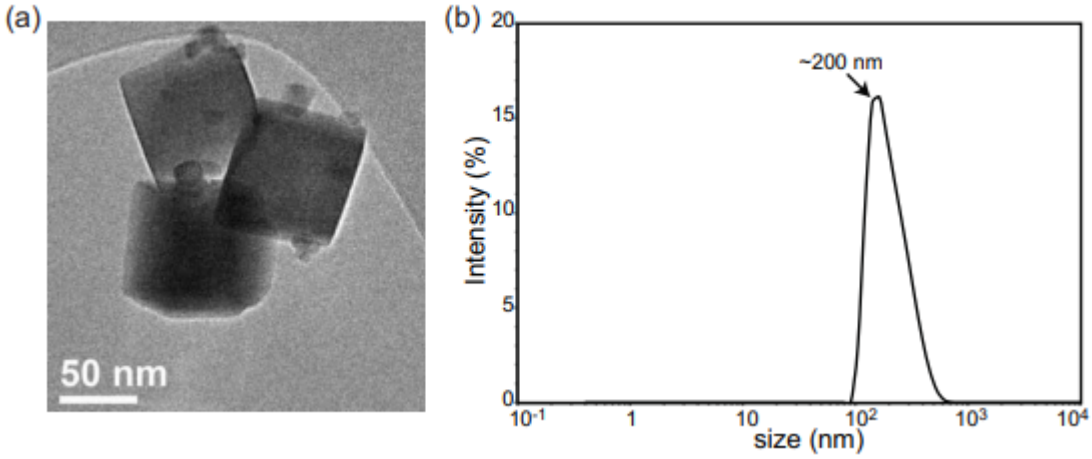


**Figure 4.2S.** (a) TEM image and (b) DLS spectrum of the ceria nanoparticles used in the present study.

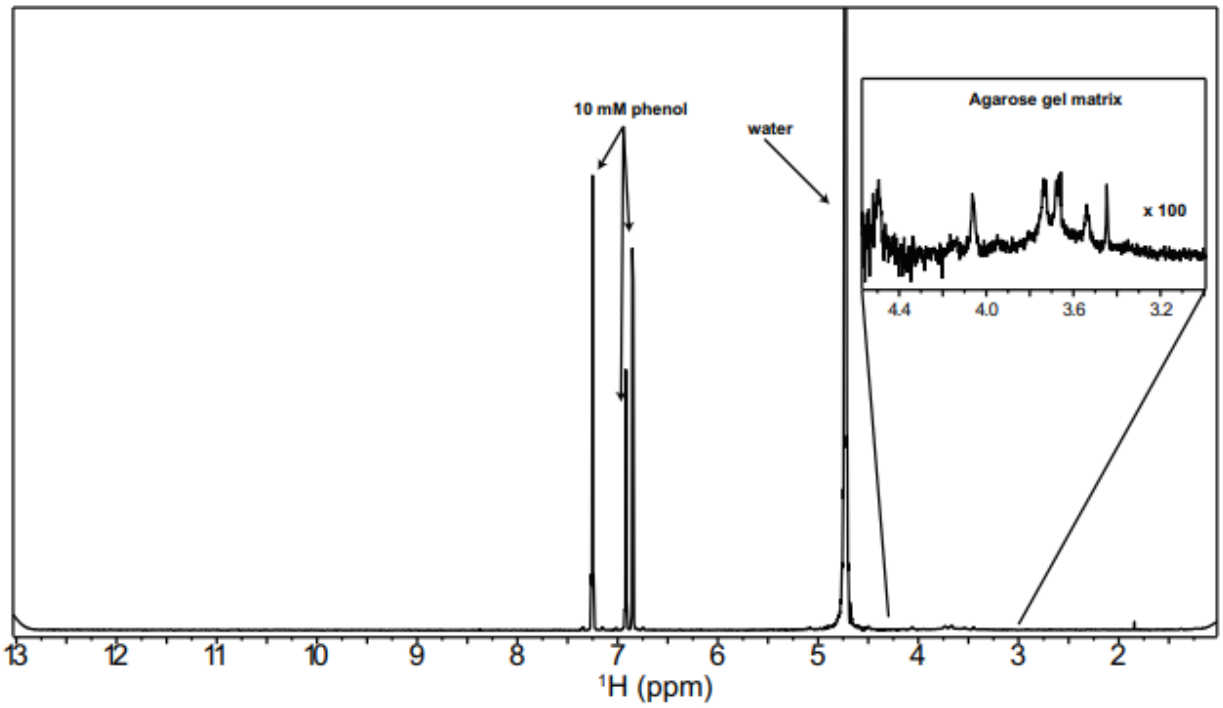


**Figure 4.3S**. $^{1}$H-NMR spectrum of 10 mM Phenol in a 1 wt % agarose gel matrix. Residual signals from the gel are enhanced 100 times in the inset.

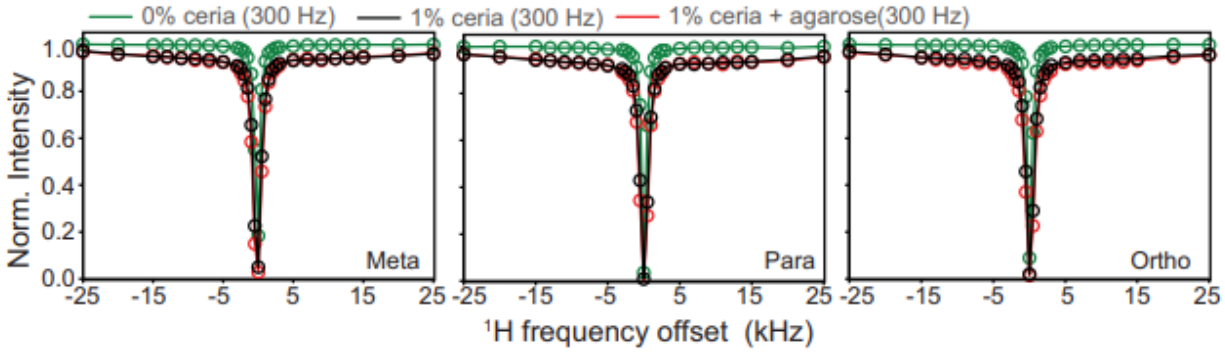


**Figure 4.4S.** $^{1}$H-DEST profiles measured for the $H_{meta}$ (left), $H_{para}$ (center) and $H_{ortho}$ (right) positions of phenol in the presence of 0 wt % ceria and 0 wt % agarose (green), 1 wt % ceria and 0 wt % agarose (black), and 1 wt % ceria and 1 wt % agarose (red). Data were measured with a saturation field strength of 300 Hz on a 600 MHz spectrometer. The data were acquired within 60 minutes from sample preparation (to limit ceria sedimentation in the absence of agarose). The perfect match between the DEST profiles measured for the phenol/nanoparticle system in the absence and in the presence of agarose (i.e. the black and red curves, respectively) suggests that the gel matrix does not perturb the small molecule-nanoparticle interaction. The DEST and lifetime line broadening effects depend upon the difference in transverse relaxation ($R_2$) between the exchanging species[9] (the free and NP-bound states in our case). $R_2$ increases with increasing rotational correlation time ($\tau_C$).[10] However, approaching the solid state limit ($\tau C > 10\text{-}5$ s), the $R_2$ dependence on $\tau_C$ reaches a plateau, and there is no appreciable change in R2 by further increasing $\tau_C$. [10] Based on the Stokes-Einstein equation and approximating the NP to a sphere with a 100 nm radius, we can calculate $\tau_C \sim 1$ ms for the unhindered nanoparticle. Therefore, an additional increase in $\tau_C$ originating from embedding the NP into the gel matrix will not affect the $R_2$ of the NP-bound state substantially, and, consequently, will not affect the DEST and lifetime line broadening data.

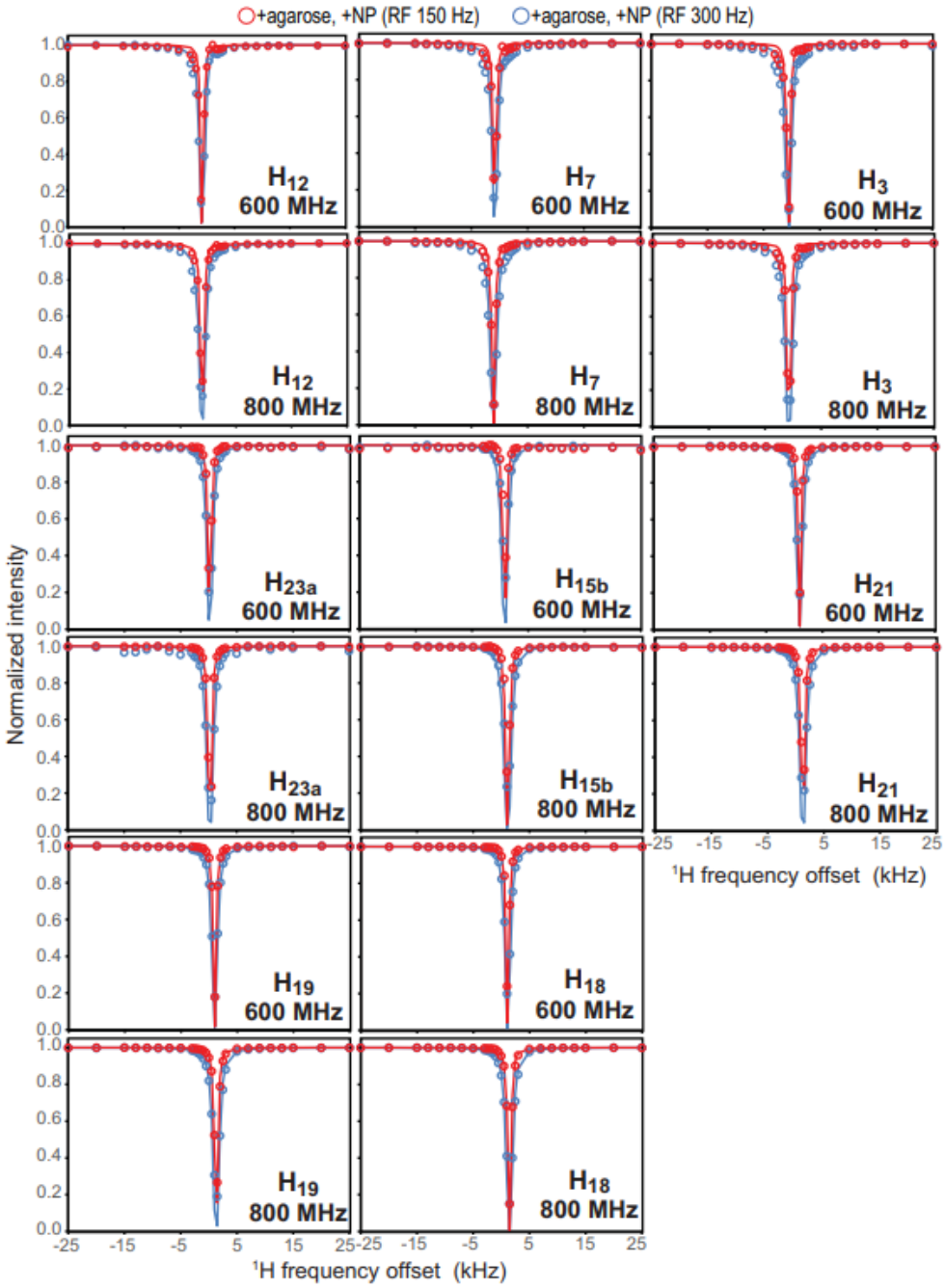


**Figure 4.5S.** $^{1}$H-DEST profiles measured for eight resolved NMR signals of CA (Figure 1d, main text) in the presence of 1 wt % agarose and 1 wt % ceria. Experimental data are shown as circles. Results of global fit of the $^{1}$H-DEST and $^{1}$H-RD data (reported in Figure S7) are shown as solid lines. Color code: 150 Hz saturation field, red; 300 Hz saturation field, blue. Data were collected on 800 and 600 MHz spectrometers.

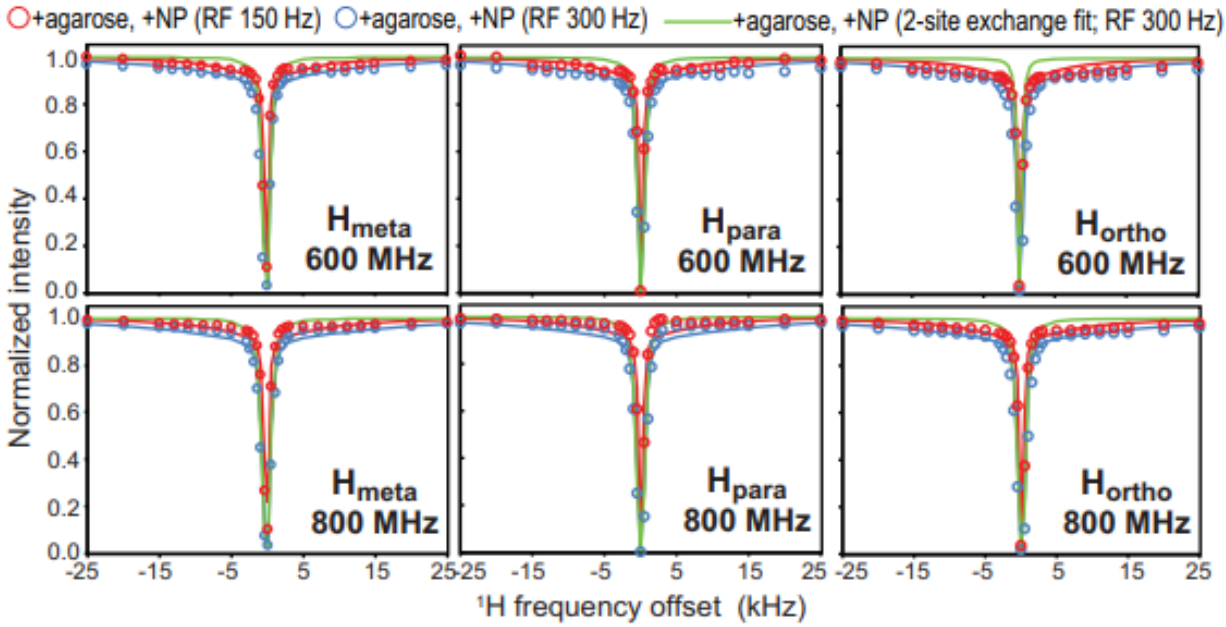


**Figure 4.6S.** [1]H-DEST profiles measured for the NMR signals of PhOH (Figure 1c, main text) in the presence of 1 wt % agarose and 1 wt % ceria. Experimental data are shown as circles. Results of global fit of the [1] H-DEST and [1] H-RD data (reported in Figure S8) are shown as solid lines. Color code: 150 Hz saturation field, red; 300 Hz saturation field, blue. Data were collected on 800 and 600 MHz spectrometers. Results from the fit using a two-site exchange model are shown for a saturation field strength of 300 Hz as green solid line. The poor agreement between the experimental DEST profile and the DEST profile simulated using the two-site exchange model justify the use of a three-site exchange model to interpret the NMR data.

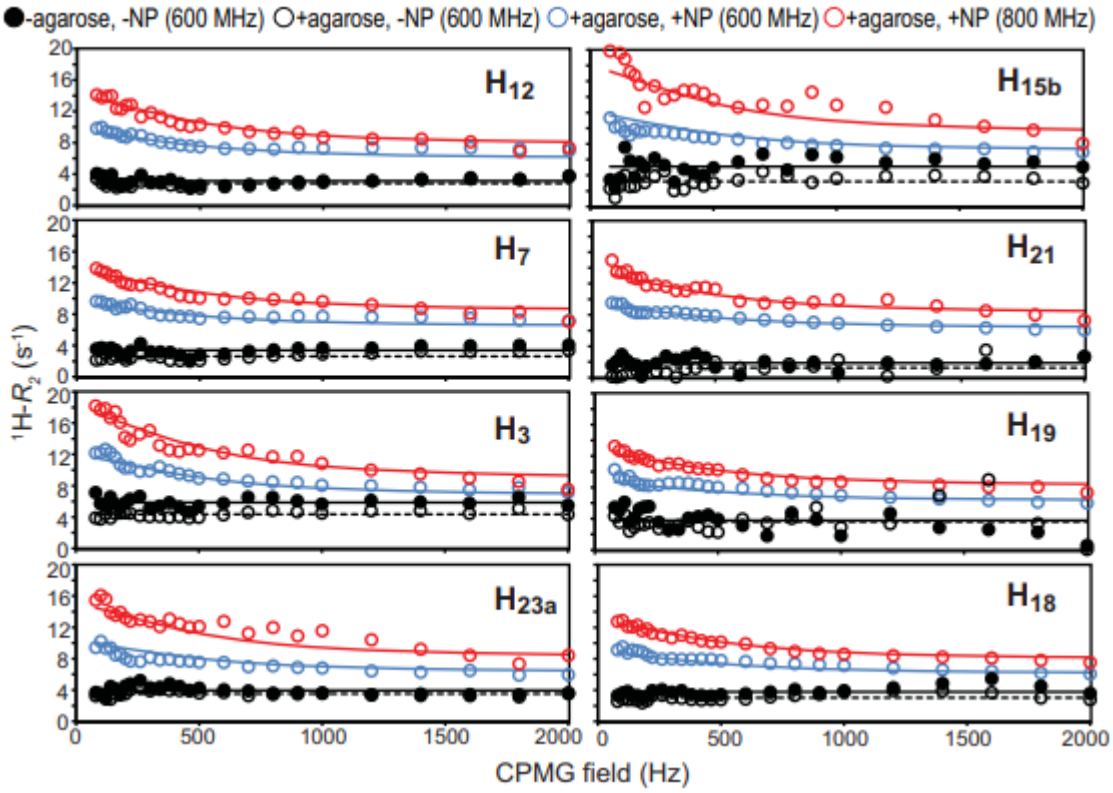


**Figure 4.7S.** $^1$H-RD profiles measured at two spectrometer frequencies (600 and 800 MHz) for eight resolved NMR signals of CA (Figure 1d, main text). Color code: 0 wt % agarose and 0 wt % ceria at 600 MHz, filled black circles; 1 wt % agarose and 0 wt % ceria at 600 MHz, open black circles; 1 wt % agarose and 1 wt % ceria at 600 MHz, open blue circles; 1 wt % agarose and 1 wt % ceria at 800 MHz, open red circles. The data acquired in the presence of ceria were fit together with the $^1$H-DEST profiles (Figure S5) to a two-site exchange model (Figure S12a). Results of the global fit are shown as blue and red solid lines for the 600 and 800 MHz data, respectively. Average (over the different CPMG fields) $R_2$'s measured in the absence of ceria with and without the gel matrix are shown as dashed and solid black line, respectively.

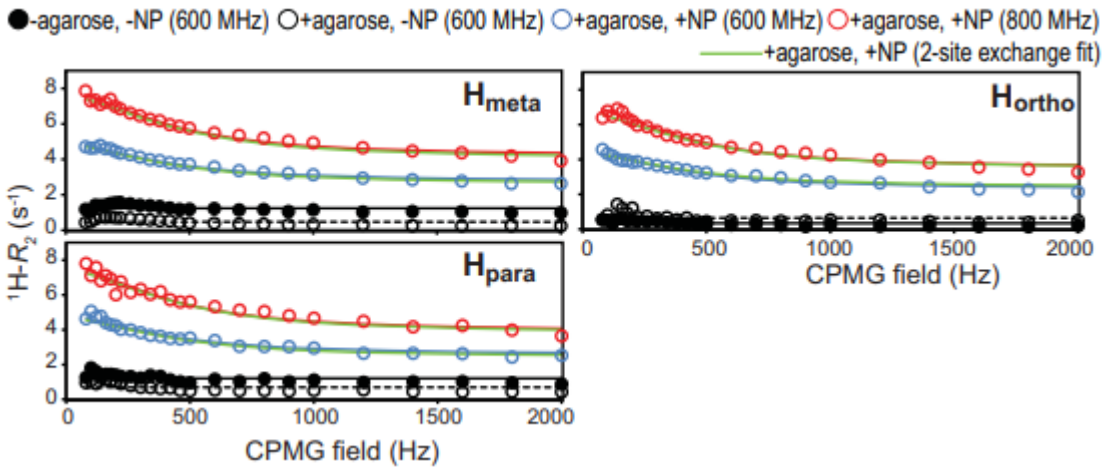


**Figure 4.8S.** [1] H-RD profiles measured at two spectrometer frequencies (600 and 800 MHz) for the NMR signals of phenol (Figure 1c, main text). Color code: 0 wt % agarose and 0 wt % ceria at 600 MHz, filled black circles; 1 wt % agarose and 0 wt % ceria at 600 MHz, open black circles; 1 wt % agarose and 1 wt % ceria at 600 MHz, open blue circles; 1 wt % agarose and 1 wt % ceria at 800 MHz, open red circles. The data acquired in the presence of ceria were fit together with the $^1$H-DEST profiles (Figure S6) to a three-site exchange model (Figure S12b). Results of the global fit are shown as blue and red solid lines for the 600 and 800 MHz data, respectively. Results from the global fit to the two-site exchange model are shown as green solid lines for both the 600 and 800 MHz data. Note that the $^1$H-RD experiment is insensitive to the rigidly bound state of Figure S12b. Average (over the different CPMG fields) $R_2$'s measured in the absence of ceria with and without the gel matrix are shown as dashed and solid black line, respectively.

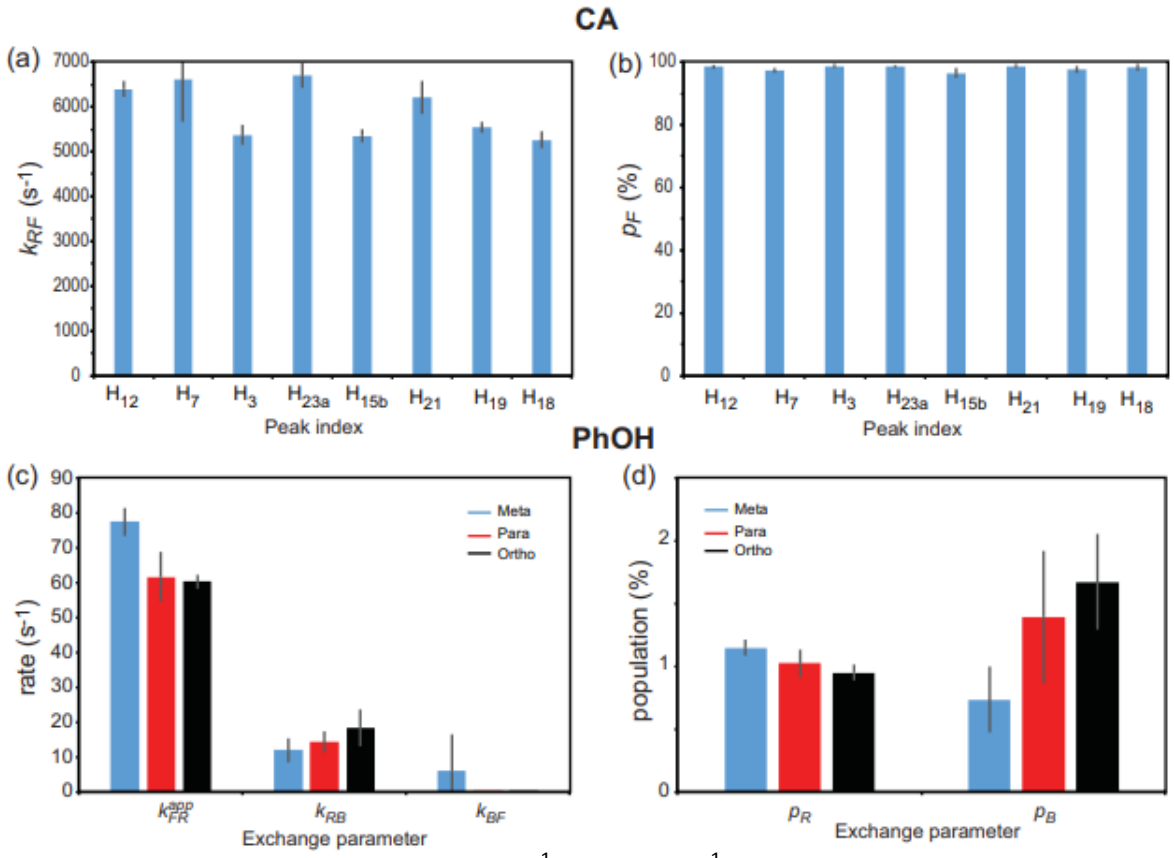


**Figure 4.9S.** Atom-based fitting of the $^{1}$H-DEST and $^{1}$H-RD data collected for CA (top) and PhOH (bottom). CA data were fit using a two-site exchange model. The fitted $k_{RF}$ and $p_F$ parameters are plotted versus the NMR peak index in (a) and (b), respectively. Similar $k_{RF}$ and $p_F$ values are fitted for the analyzed peaks of CA, therefore justifying the global fit procedure discussed in the main text. PhOH data were fit using a three-site exchange model. The optimized kinetic ($k^{app}_{FR}$, $k_{RB}$, $k_{BF}$,) and thermodynamic ($p_R$, $p_B$) parameters are shown in (c) and (d), respectively, for the $H_{meta}$ (blue), $H_{para}$ (red) and $H_{ortho}$ (black) positions. From the individual fits, similar exchange parameters are obtained for the NMR signals of PhOH, therefore justifying the global fit procedure discussed in the main text.

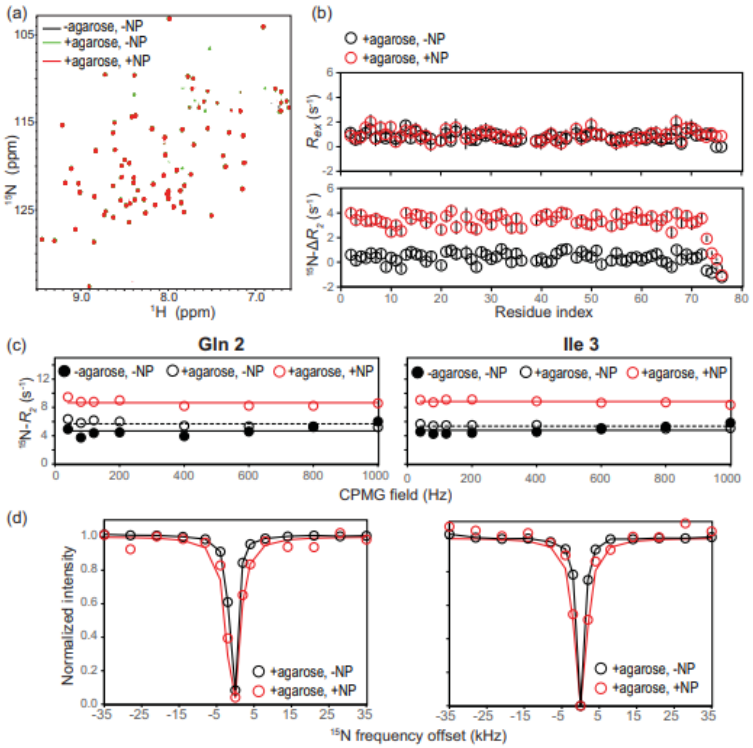


**Figure 4.10S**. (a) $^{1}$H-$^{15}$N Heteronuclear Single Quantum Coherence (HSQC) spectra of ubiquitin (ubq) in 0 wt % agarose and 0 wt % ceria (black), 1 wt % agarose and 0 wt % ceria (green), and 1 wt % agarose and 1 wt % ceria (red). (b) $^{15}$N-$\Delta R_2$ for ubq versus residue index (bottom). $^{15}$N-$\Delta R_2$ are calculated as the difference between the $R_2$ in the presence of agarose and the $R_2$ in the absence of agarose ($\Delta R_2^{gel} = R_2^{gel} - R_2^{water}$, black circles), or as the difference between the $R_2$ in the presence nanoparticles and the $R_2$ in the absence of nanoparticles ($\Delta R_2^{NP} = R_2^{gel + NP} - R_2^{gel}$, red circles). In the top panel the exchange contribution to the $^{15}$N-$R_2$ rate ($R_{ex}$) is plotted versus the residue index. $R_{ex}$ were measured as the $R_2$ at low CPMG filed (50 Hz) minus $R_2$ at high CPMG field (1000 Hz) for ubq samples prepared in the presence of the gel matrix ($R_{ex}^{gel}$, black circles) and in the presence of both gel matrix and nanoparticles ($R_{ex}^{NP}$, red cirlces). $R_{ex}^{gel}$ and $\Delta R_2^{gel} \sim 0$ throughout the protein sequence indicate the absence of protein-gel interaction. $\Delta R_2^{NP} > 0$ suggests the existence of large protein-nanoparticle adducts upon addition of ceria. (c) $^{15}$N relaxation dispersion profiles for two representative residues of ubq (Gln 2 and Ile 3) acquired on a 600 MHz spectrometer in the absence of agarose (filled black circles), in the presence of agarose (open black circles), and in the presence of both agarose gel and nanoparticles (open red circles). (d) $^{15}$N DEST profiles for two representative residues of ubq (Gln 2 and Ile 3) acquired in the presence of agarose (open black circles), and in the presence of both agarose gel and nanoparticles (open red circles). The displayed data were collected at 600 MHz using a saturation field of 500 Hz. The absence of clear relaxation dispersions ($R_{ex}^{NP} \sim 0$) and the very small increase in the width of the DEST profiles observed for ubq upon addition of nanoparticles indicate that the exchange process is too fast (exchange rate ≥ 10,000 s$^{-1}$) to be described quantitatively by $^{15}$NDEST and $^{15}$N-RD experiments.

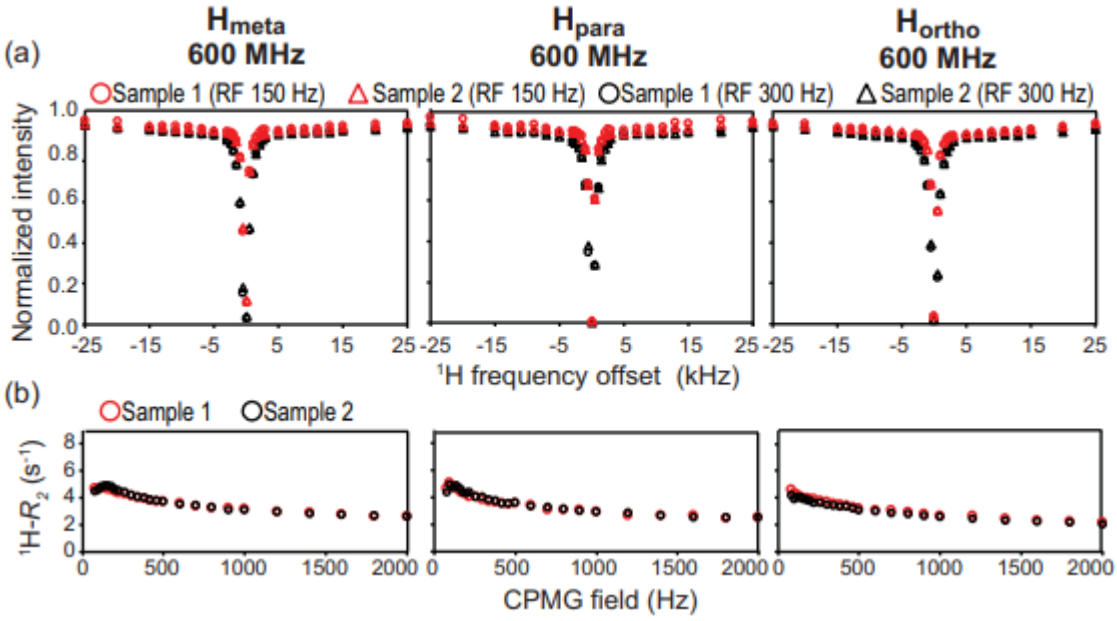


**Figure 4.11S.** (a) Experimental $^{1}$H-DEST profiles measured at saturation fields of 150 (red) and 300 (black) Hz for two different preparations of the NMR sample containing 10 mM PhOH, 1 wt % ceria and 1 wt % agarose. Data from sample 1 are shown as open circles. Data from sample 2 are shown as open triangles. (b) Experimental 1 H-RD profiles measured for two different preparations of the NMR sample containing 10 mM PhOH, 1 wt % ceria and 1 wt % agarose. Data from sample 1 are shown as red open circles. Data from sample 2 are shown as black open circles. $^{1}$H DEST and $^{1}$H-RD data were measured on a 600 MHz spectrometer. Data for the meta (left), para (center) and ortho (right) protons are shown. The error is ≤ 5% for all the measured data points.

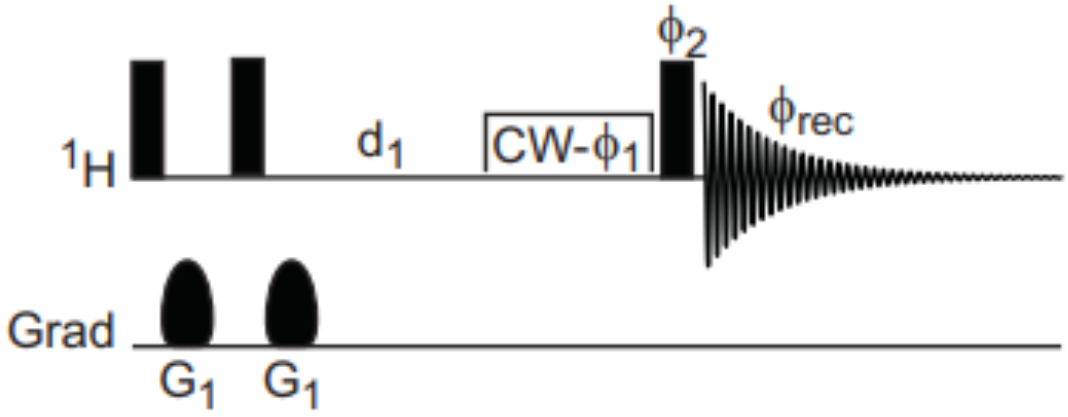


**Figure 4.12S.** Pulse scheme used for measuring $^{1}$H-DEST profiles for Ph and CA. Rectangular pulses are applied with flip angle of 90° along the x-axis unless indicated otherwise. $d_1$ (5 s) is the recycling delay. CW is the saturation radio frequency applied with phase $\phi_1$. The

phase cycling employed is: $\phi_1$ = (x, x, x, x, -x, -x, -x, -x); $\phi_2$ = (x, -x, -x, x, y, -y, -y, y); $\phi_{rec}$ = (x, -x, -x, x, y, -y, -y, y). The duration and strength of the gradients G1 are 1 ms and 50 G/cm, respectively. The first two pulses and the two $G_1$ gradients completely dephase the magnetization, which is then recovered during $d_1$. This scheme guarantees that the same equilibrium magnetization is present at the beginning of the saturation period for all frequency offsets.

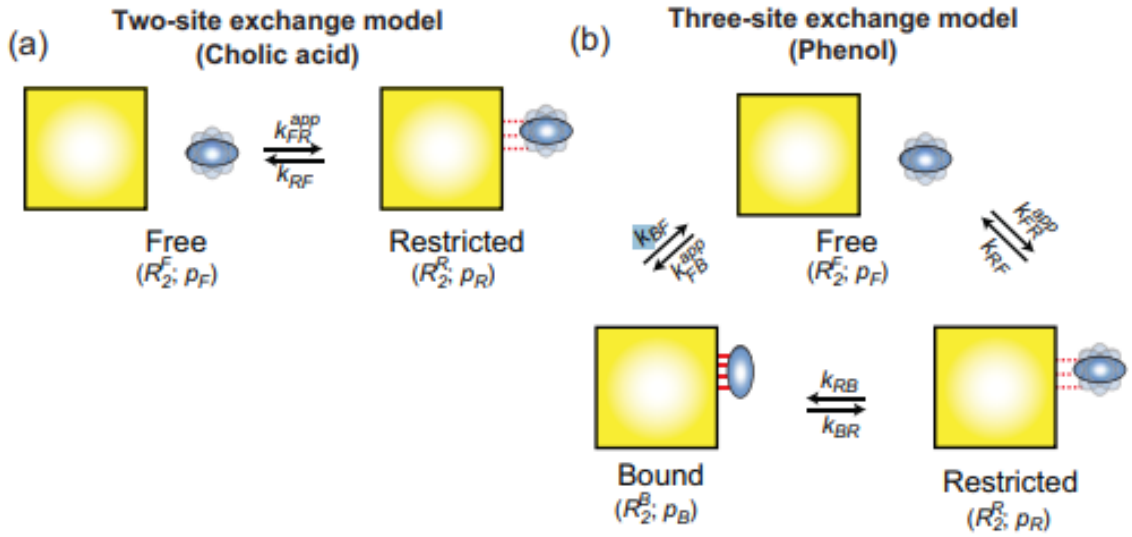


**Figure 4.13S.** Parameters for the (a) two- and (b) three-site exchange models used for fitting the $^1$H-DEST and $^1$H-RD data acquired for CA and PhOH, respectively. The NP is represented as a yellow square. The small molecule is shown as blue oblate ellipsoid. Transparent blue oblate ellipsoids are used to indicate rotational diffusion of the small molecule. Solid and dashed red lines represent strong and weak interactions, respectively (see main text). Rate constants for each equilibrium ($k_{xy}$ , where x and y are the states in thermodynamic equilibrium; $k_{xy}^{app}$ indicates the apparent rate constant) and populations of each state ($p_F$, $p_R$ and $p_B$ for the free, restricted and bound states, respectively) were fit as global parameters. $R_2$ rates for each state and chemical shift changes associated with each equilibrium were fit as peak specific parameters.

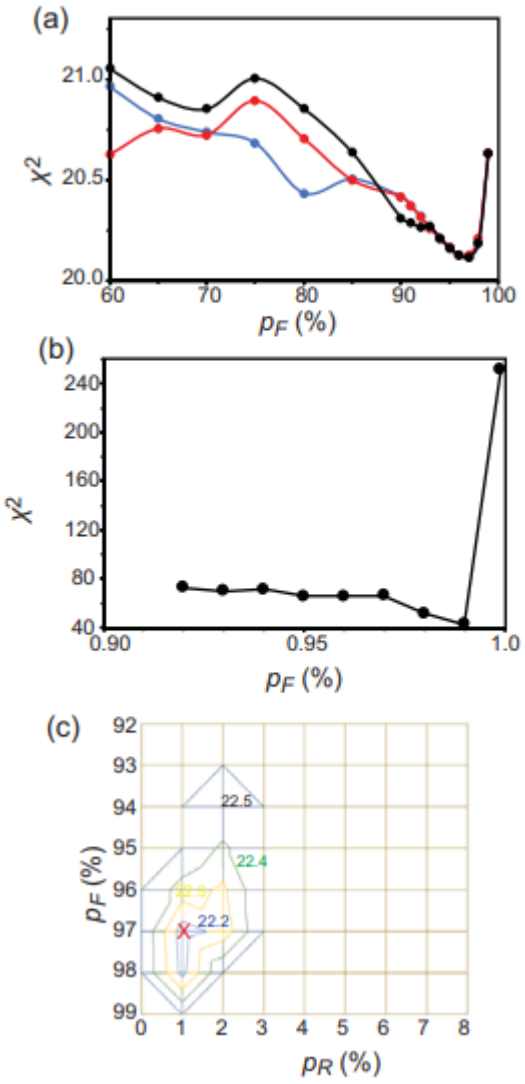


**Figure 4.14S.** (a) $^1$H-DEST and $^1$H-RD data acquired for CA in the presence of 1 wt % NP are fit using a two-site exchange model and keeping the population of the free state ($p_F$) to a constant value. Plotting *χ2* (Equation S1) versus $p_F$ a clear minimum at $p_F$ ~ 96 % can be observed. Calculations were repeated three times using different starting parameters (black, red and blue curves in the figure, respectively). (b) $^1$H-DEST and $^1$H-RD data acquired for PhOH in the presence of 1 wt % NP are fit to a two-site exchange model keeping the population of the free state ($p_F$) to a constant value. Here, a plot of *χ2* (Equation S1) versus $p_F$ is reported. (c) $^1$H-DEST and $^1$H-RD data acquired for PhOH in the presence of 1 wt % NP are fit using a three-site exchange model and keeping the populations of the free (pF) and restricted (pR) states to constant values. Plotting *χ2* (Equation S1) versus $p_F$ and $p_R$ a minimum at pF ~ 97 % and pR ~ 1 % can be observed (red cross). Contours are at χ2 = 22.2 (blue), 22.3 (yellow), 22.4 (green) and 22.5 (black). Note that fitting the PhOH data using a three-site exchange model returns lower *χ2* values if compared to the fits using a two-site exchange model.

## Supplementary Tables

**Table 4.S1.** Kinetic, population, chemical shift and transverse relaxation parameters for CA binding to ceria nanoparticles derived from global fitting of $^{1}$H-DEST and $^{1}$H-RD data using a two-site exchange model. Definitions for the global parameters are given in **Figure 4.S12a**. $R_2^{600}$ and $R_2^{800}$ are the fitted transverse relaxation rates for the motionally restricted state at 600 and 800 MHz, respectively. $\Delta_{FR}$ is the absolute value of the chemical shift difference in ppm between free and restricted state.

| | $H_{12}$ | $H_7$ | $H_3$ | $H_{23a}$ | $H_{15b}$ | $H_{21}$ | $H_{19}$ | $H_{18}$ |
|---|---|---|---|---|---|---|---|---|
| **Global Parameters** | | | | | | | | |
| $P_F$ (%) | | | | 96.0 ± 0.2 | | | | |
| $P_R$ (%) | | | | 4.0 ± 0.2 | | | | |
| $k_{FR}^{app}$ ($s^{-1}$) | | | | 250 ± 25 | | | | |
| $k_{RF}$ ($s^{-1}$) | | | | 6100 ± 350 | | | | |
| **Proton Specific Parameters** | | | | | | | | |
| $R_2^{600}$ ($s^{-1}$) | 52±3 | 58±3 | 62±3 | 56±4 | 66±4 | 70±10 | 56±3 | 53±3 |
| $R_2^{800}$ ($s^{-1}$) | 75±4 | 84±4 | 90±4 | 81±5 | 96±6 | 102±15 | 81±5 | 76±4 |
| $\Delta_{FR}$ (ppm) | 0.15±0.01 | 0.14±0.01 | 0.18±0.01 | 0.16±0.01 | 0.18±0.01 | 0.05±0.05 | 0.13±0.01 | 0.13±0.01 |

**Table 4.S2.** Kinetic, population, chemical shift and transverse relaxation parameters for PhOH binding to ceria nanoparticles derived from global fitting of $^{1}$H-DEST and $^{1}$H-RD data using a three-site exchange model. Definitions for the global parameters are given in Figure S12b. $R_2^{600R}$ and $R_2^{800R}$ the fitted transverse relaxation rates for the motionally restricted state at 600 and 800 MHz, respectively. $R_2^{600B}$ and $R_2^{800B}$ the fitted transverse relaxation rates for the bound state at 600 and 800 MHz, respectively. *ΔFR* and *ΔFB* are the absolute value of the chemical shift difference in ppm between free and restricted state, and free and bound state, respectively. Note that *ΔFB* values were fixed to 0 in the fit because $^{1}$H-RD experiments are insensitive to the bound state (the restricted and bound states exchange on the seconds timescale), and the DEST is a low-resolution experiment that does not provide chemical shift information.[11]

| | $H_{meta}$ | $H_{para}$ | $H_{ortho}$ |
|---|---|---|---|
| **Global Parameters** | | | |
| $P_F$ (%) | | 96.7 ± 0.2 | |
| $P_R$ (%) | | 1.1 ±0.1 | |
| $P_B$ (%) | | 2.2 ±0.2 | |
| $k_{FR}^{app}$ ($s^{-1}$) | | 60 ± 6 | |
| $k_{RF}$ ($s^{-1}$) | | 5346 ±176 | |
| $k_{RB}$ ($s^{-1}$) | | 9 ± 2 | |
| $k_{BR}$ ($s^{-1}$) | | 5 ± 1 | |
| $k_{BF}$ ($s^{-1}$) | | 0 ± 0 | |
| $k_{FB}^{app}$ ($s^{-1}$) | | 0 ± 0 | |
| **Proton Specific Parameters** | | | |
| $R_2^{600R}$ ($s^{-1}$) | 100±6 | 92±2 | 80±2 |
| $R_2^{600B}$ ($s^{-1}$) | 10,030±30 | 10,090±30 | 10,200±40 |
| $R_2^{800R}$ ($s^{-1}$) | 160±10 | 146±6 | 127±6 |
| $R_2^{800B}$ ($s^{-1}$) | 15,900±500 | 15,900±500 | 16,100±500 |
| $\Delta_{FR}$ (ppm) | 0.21±0.01 | 0.21±0.01 | 0.20±0.01 |
| $\Delta_{FB}$ (ppm) | 0.00±0.00 | 0.00±0.00 | 0.00±0.00 |

# CHAPTER 5

## 'SURFACE CONTRAST' NMR REVEALS NON-INNOCENT ROLE OF SUPPORT IN PD/CEO2 CATALYZED PHENOL HYDROGENATION

Submitted to Science Advances and is *under review* as of October 2019

Timothy K. Egner,[1,‡] Pranjali Naik,[1,2,‡] Yeongseo An,[1] Amrit Venkatesh,[1] Aaron J. Rossini,[1,2] Igor I. Slowing,[1,2,*] and Vincenzo Venditti[1,3,*]

*1 Department of Chemistry, Iowa State University, Ames, Iowa 50011, USA.*
*2 U.S. Department of Energy, Ames Laboratory, Ames, IA 50011, USA.*
*3 Roy J. Carver Department of Biochemistry, Biophysics and Molecular Biology, Iowa State University, Ames, Iowa 50011, USA.*
‡ Equal contribution

Ceria-supported metals are widely used as catalysts because of their exceptional redox properties. Here, we use surface contrast NMR methods to investigate the hydrogenation of phenol by Pd supported on ceria nanoparticles. We show that the rigid and planar binding of phenol to Pd is mediated by a weak and highly mobile association of the small molecule to ceria. Interestingly, while addition of phosphate to the mixture does not perturb the adsorption of phenol on Pd, it destabilizes its interaction with ceria and proportionally decreases the rate of catalytic conversion. Our data provide strong experimental evidence that weak interactions between adsorbate and ceria are catalytically competent and explain the exceptional performance of Pd/ceria compared to catalysts prepared using inactive supports.

### Introduction

Obtaining a comprehensive, atomic resolution knowledge of the exchange dynamics among desorbed, physisorbed, and chemisorbed molecules at the nanoparticle (NP)/solution interface is crucial toward the engineering of new and more efficient heterogeneous catalysts.

Yet, our understanding of the adsorption/desorption equilibria on NP surfaces and of their role in regulation of catalysis is still largely incomplete. This knowledge gap largely arises from the lack of experimental techniques able to provide a complete characterization of the kinetics and thermodynamics of sorption with atomic resolution. Here, we show that surface contrast (SC) solution NMR methods developed recently to investigate the binding of small, NMR visible proteins to large, NMR invisible molecular machines[1], aggregates[2], biological membranes[3, 4], and nanoparticles[5, 6] can be adapted to gain atomic resolution insight into the dynamics occurring at the surface of a heterogeneous catalyst. In brief, these methods utilize the contrast between the fast transverse relaxation ($R_2$) of the slowly tumbling complexed protein and the slower $R_2$ rate of the free protein to imprint information on the dynamics of the complexed state onto the spectrum of the NMR-observable free state (**Supplementary Figure 5.S1**)[7]. The approach is illustrated by the adsorption/desorption equilibrium of phenol (PhOH) on the surface of ceria-supported palladium ($Pd/CeO_2$) NPs. This heterogeneous catalyst has spawn significant attention in recent years due to its ability to catalyze the hydrogenation of phenolic compounds with very high selectivity at low temperature (35 °C) and hydrogen pressure (1 atm)[8]. However, it is still unclear if the exceptional performance of the $Pd/CeO_2$ catalyst compared to conventional Pd/C and $Pd/Al_2O_3$ NPs is the result of the higher dispersion of the metal NP, more efficient adsorption of the substrate on the catalyst surface, electronic effects of the support on the metal, or direct involvement of ceria in catalysis.

Here, we show that adsorption of PhOH on the metal NP proceeds through a weakly bound and highly flexible intermediate state in which the small molecule is likely hydrogen bonded to the support. Interestingly, we notice that addition of phosphate does not perturb the

interaction of PhOH with the metal catalyst, but results in a ~2-fold destabilization of its interaction with ceria and in a corresponding ~2-fold reduction in catalytic conversion. Our results provide strong experimental evidence that weak interactions between substrate and support are competent for the hydrogenation of PhOH catalyzed by the Pd/ceria catalyst, which explains its exceptional performance compared to catalysts prepared using other types of support.

**Results and discussion**

We recently reported that adsorption of PhOH on the surface of ceria ($CeO_2$) NPs is a two-step process in which an initial, weakly bound encounter complex evolves into a tight PhOH-NP adduct[9]. Evidence for this multi-step adsorption process was obtained by combined analysis of $^1$H dark-state exchange saturation transfer (DEST) and $^1$H relaxation dispersion (RD) experiments[9]. Although the high sensitivity of $^1$H NMR experiments allows for rapid acquisition of the required experimental data, they also present certain drawbacks that limit the accuracy and extent of information provided on the adsorption process. Chief among these, the $^1$H DEST experiment is highly susceptible to spin diffusion that can bias the quantitative analysis of the DEST profile by overemphasizing the relaxation contrast effect of a surface bound ligand[2]. Additionally, it is difficult to accurately simulate $^1$H relaxation due to complicated geometric consideration of other $^1$H nuclei within the molecule and the solvent. This aspect limits the ability of interpreting the results from $^1$H DEST and $^1$H RD experiments in terms of structure and dynamics of the adsorbate. Here, we show that simultaneous analysis of $^1$H and $^{13}$C RD, $^{13}$C DEST, and $^{13}$C longitudinal relaxation ($R_1$) data returns a description of the dynamics of the adsorbate at the NP surface with unprecedented details. Indeed, while the $^1$H and $^{13}$C RD experiments provide thermodynamic and kinetic information on adsorption/desorption equilibria occurring

on the µs-ms timescale, the $^{13}C$ DEST data report on slower (ms-s) processes[7]. In addition, the $^{13}C$ $R_1$ and $R_2$ rates of PhOH in the adsorbed state were obtained by quantitative analysis of the $^{13}C$ DEST, RD, and $R_1$ datasets and can be interpreted based on physical models that report atomic resolution details on the NP-adsorbate interaction (Supplementary Methods).

DEST, RD, and $R_1$ experiments were measured on 10 mM PhOH solutions with 1 wt% CeO2 or 1 wt% Pd/$CeO_2$ using the pulse sequences shown in **Supplementary Figure 5.S2**. Both NPs are of cubic morphology with a side length of ~25 nm (**Supplementary Figure S3**). Pd NPs deposited on ceria are of small size (≤ 2 nm) and are highly dispersed on the support (**Supplementary Figures 5.S3 and 5.S4**). A 1 wt% agarose gel matrix was used for all samples to prevent sedimentation of NPs (see Methods). As previously reported, the 1 wt% agarose gel (pore size 100-500 nm) provides a chemically inert matrix to study PhOH adsorption to ceria that allows for optimal homogeneity and stability of the NMR sample, and does not interfere with the SC-NMR experiments and analysis[9].

Addition of NPs to the PhOH solution results in broadening of the DEST profiles and a dramatic change in the RD curves. The RD curves move to a higher $R_2$ range and display curvature characteristic of exchange equilibria on the µs-ms timescale (**Supplementary Figure 5.S5**). Interestingly, while the data measured on the Pd/$CeO_2$ sample show homogeneous broadening of the DEST profiles and homogeneous changes in RD curves of the para, meta, and ortho positions, changes observed in the presence of CeO2 are more heterogeneous, with the para position experiencing a larger broadening of the DEST profile and an increased R2 rate compared to the meta and ortho carbons (**Figure 5.1** and **Supplementary Figure 5.S5**). These experimental

evidences indicate that adsorbed PhOH undergoes different dynamics when bound to $CeO_2$ or $Pd/CeO_2$ [2, 3].

Comprehensive understanding of the dynamic interaction between PhOH and NPs was obtained by quantitative modeling of the experimental DEST, RD, and $R_1$ data using an adsorption/desorption exchange model in which all experimental observables are described by solutions to the McConnell equation (Supplementary Methods). A three-site exchange model was required to fully account for the experimental data (**Figure 5.1**). The three sites involve the desorbed state (D), a weakly adsorbed state (W) (in fast/intermediate exchange with the desorbed state), and a tightly adsorbed state (T) (in intermediate/slow exchange with the weakly adsorbed state) (**Figure 5.2**). Interestingly, there is no direct transition between the desorbed and tightly adsorbed states ($k_{DT}^{app}$ ~ $k_{TD}$ ~ 0 s-1) (**Supplementary Tables 5.S1 and 5.S2**), and, therefore, the weakly associated state is an intermediate in the PhOH adsorption pathway. Model free (MF) analysis[10] of the $^{13}C$ $R_1$ and $R_2$ rates of weakly bound PhOH reveals that the adsorbate is highly mobile (order parameter, $S_W^2$ , < 0.01), which is suggestive of hydrogen bonding interaction between the hydroxyl group of PhOH and a surface exposed oxygen on the NP (**Figure 5.2**). Indeed, the existence of three rotatable bonds between the ceria surface and the rigid aromatic ring of PhOH (**Figure 5.2**) allows for extensive local flexibility of the adsorbate ($S^2 << 0.1$)[11]. Consistent with this hypothesis, dispersion of Pd on the support reduces the exposed ceria surface from 16.5 $m^2\ g^{-1}$ (in $CeO_2$) to 6.5 $m^2\ g^{-1}$ (in $Pd/CeO_2$) (**Supplementary Figure 5.S4**) and results in a corresponding ~2-fold reduction in the population ($p_w$) of the weakly bound intermediate (from 5.2 to 2.3 %; **Figure 5.2**).

MF analysis of the tightly adsorbed state reveals two different binding modes for PhOH to $CeO_2$ and Pd/$CeO_2$. Indeed, while the data measured in the presence of Pd/$CeO_2$ could be fit using a global order parameter for the tightly bound state ($S_T^2$ ~ 0.9), different $S_T^2$ for the para ($S_T^2$ ~ 1), meta ($S_T^2$ ~ 0), and ortho ($S_T^2$ ~ 0.1) positions are needed in order to account for the experimental data measured in the presence of $CeO_2$ (**Supplementary Figure 5.S7**). Given the small size and structural rigidity of PhOH, these $S_T^2$variations cannot be ascribed to local motion at the individual C-H bond vector level but rather to a global motional phenomenon involving the whole molecule on the surface of the NP[3, 11]. Interestingly, the use of a restricted rotation (RR) model[3, 11] in which PhOH is allowed fast (ps-ns timescale) rotation about an axis that coincides with the para C-H bond vector (see Supplementary Methods) can account for the experimental data (**Figures 5.1 and 5.3**). Indeed, such fast rotation will result in a dependency of the $^{13}C$ $R_2$ of the tightly bound state on the angle (θ) formed by the C-H bond vector and the axis of rotation, with maxima in $R_2$ when the bond vector is parallel (θ = 0°) or antiparallel (θ = 180°) with the axis of rotation and minima at the magic angles (θ =54.74° and 125.26°). The angular dependency of the transverse relaxation of tightly adsorbed PhOH is apparent when plotting the $^{13}C$ $\Delta R_2$ values (i.e. the increase in $R_2$ observed upon addition of NP, **Supplementary Figure 5.S1**) versus θ, and suggests that the tightly adsorbed species corresponds to PhOH bound to an oxygen vacancy present on the surface of $CeO_2$ (**Figure 5.3**). On the contrary, the global $S_T^2$ ~ 0.9 obtained for PhOH tightly adsorbed on Pd/$CeO_2$ indicates that the small molecule is rigidly associated with the NP, which provides experimental evidence for the flat binding of PhOH to Pd suggested by ab initio calculations[12] (**Figure 5.3**). Of note $p_T$ and $k_{WT}^{app}$ obtained for PhOH adsorption to $CeO_2$ and Pd/$CeO_2$ are similar (**Figure 5.2**), which implies a similar number of tight absorption sites on the

two NPs. This observation is in agreement with the hypothesis that vacant sites on ceria provide nucleation points for formation of small metal NPs[13, 14], and is supported by EPR measurement indicating that dispersion of Pd on ceria eliminates nearly all vacant sites on the support (**Supplementary Figure 5.S12**). It is also interesting to note that PhOH desorption from $CeO_2$ is faster than from Pd/CeO2 (compare $k_{TW}$ values in **Figure 5.2**), which suggests formation of a stable PhOH-Pd adduct and reflects the ability of water to displace PhOH from oxygen vacant sites.

To test the proposed binding models, additional NMR measurements were performed in the presence of 20 mM inorganic phosphate (Pi), which is known to bind efficiently to the surface of $CeO_2$[15]. As expected, addition of Pi reduces the population of the weakly adsorbed intermediate ($p_W$; **Figure 5.2**) and destabilizes the interaction of PhOH with oxygen vacancies on $CeO_2$ ($p_T$ < 0.01 %; **Figure 5.2c**). However, population ($p_T$) and exchange kinetics ($k_{WT}^{app}$ and $k_{TW}$) of PhOH tightly adsorbed on Pd/$CeO_2$ are unaffected by Pi (**Figures 5.2b,c** and **Supplementary Figure 5.S13**), confirming that this binding mode does not depend upon contacts between PhOH and CeO2. Importantly, addition of 20 mM Pi also results in a ~2-fold decrease in the rate of PhOH hydrogenation catalyzed by Pd/$CeO_2$ (**Figure 5.3f**), which correlates closely with the ~2-fold decrease in population of weakly adsorbed PhOH to ceria caused by Pi (**Figure 5.2**). These data provide evidence of a direct involvement of ceria in catalysis, specifically that the weak binding of PhOH to ceria is catalytically competent for the hydrogenation reaction.

## Conclusions

Metal NPs supported on ceria are widely used due to their superior catalytic efficiency compared to equivalent catalysts prepared on other supports[8, 16, 17]. However, the origins of the

exceptional performance of ceria-supported catalysts is unclear. Here, we have obtained a comprehensive description of the dynamic equilibria underlying PhOH adsorption on a $Pd/CeO_2$ catalyst. We have shown that the rigid and planar binding of PhOH to the metal NP inferred by ab initio calculations[12] is mediated by a weak PhOH/ceria adduct in which the small molecule is likely hydrogen bonded to the support. Interestingly, we noticed that disruption of this weakly bound intermediate state by addition of phosphate results in a corresponding disruption of catalytic conversion, indicating that the weak small molecule-ceria adduct is competent for PhOH hydrogenation. Such participation of ceria in catalysis revealed by our study explains the superior performances of the ceria supported catalyst compared to other supported Pd NPs.

More in general, we have demonstrated that SC-NMR experiments can provide unprecedented, atomic-resolution details on the adsorption/desorption equilibria at the surface of a NP catalyst. Indeed, SC-NMR data can be used to study the structure, dynamics, and thermodynamics of the adsorbed species, and are capable to differentiate among adsorbate species bound to different components of a hierarchical NP. In addition, our results have shown that SC-NMR data can guide the interpretation of reaction kinetics experiments, returning important mechanistic information on the catalytic process, such as the involvement of a specific component of a hierarchical NP in catalysis. Application of SC-NMR to a variety of heterogeneous catalyst systems could provide previously unattainable experimental insights into the role played by sorption equilibria in regulating the efficiency and selectivity of catalytic reactions.

## Methods

***Synthesis of CeO2 and Pd/CeO2 NPs.*** Ceria cubes were synthesized following a literature procedure[18]. Pd catalysts were prepared by impregnation with a 1 wt% Pd loading relative to the mass of the support. Further details are reported in Supplementary Methods.

***NMR samples.*** NMR samples were prepared by mixing together PhOH (10 mM), ceria nanoparticles (1 wt%), agarose (1 wt%), and sodium phosphate (0 or 20mM). The pD of the sample was adjusted to 7.0 ± 0.1 using DCl or NaOD. The mixture was sonicated for 60 seconds (1:2 on:off duty cycle) at 25% power on a 500 W horn-tipped sonicator to disperse the nanoparticles. The mixture was then heated to 85 °C for three minutes, transferred to a 5 mm NMR tube, and allowed to cool at room temperature. Reference samples that did not contain nanoparticles were prepared using the same procedure as above.

***Transmission Electron Microscopy.*** NP imaging was performed in a FEI Tecnai G2 F20 field emission TEM operating at 200 kV. Samples were prepared by placing 3-4 drops of ethanol suspension (~0.1 mg $mL^{-1}$) onto lacey-carbon-coated copper grids and dried in air. Size distribution histograms were obtained from TEM images by measuring the diameter of ~70 particles using Digital Micrograph® software. Average particle sizes were determined from this size distribution histogram.

***Electron Paramagnetic Resonance.*** EPR spectra were recorded at 25 °C on a Bruker ELEXYS E580 X-band spectrometer operating at 9.86 GHz. Each spectrum was centered at 3,450 G and acquired with 4 scans, a sweep width of 200 G, a sweep time of 20.94 s, and a digital resolution of < 0.1 G. Samples were prepared as dry powder (~100 mg) in 3 mm (ID) tubes.

***Inductively coupled plasma–optical emission spectroscopy.*** Pd loadings were analyzed in a Perkin Elmer Optima 2100 DV ICP-OES instrument. Samples (10 mg) were sonicated and vortexed for 24 h in 2 mL aqua regia. The solution was diluted to 12.0 mL with deionized water, filtered, and used to analyze the amount of palladium metal in the samples.

***Powder X-ray Diffraction (PXRD).*** Diffraction patterns were collected on a Bruker Siemens D500 X-ray Diffractometer (XRD) equipped with Cu Kα radiation source (40 kV, 44 mA) over the range of 10–100 2θ°. Samples were prepared by placing fine powders onto a background-less polycarbonate sample holder. Crystallite sizes were calculated using the Scherrer equation.

***Chemisorption.*** $H_2$-chemisorption analysis was carried out by pretreating the samples at 350 °C under $H_2$/Ar for 30 min, followed by flowing Ar for 15 min at 350 °C to remove surface-bound and dissolved hydrogen from Pd crystallites. The sample was then cooled under Ar to −20°C for hydrogen pulse chemisorption measurements. The palladium dispersion and surface area were calculated as described in Supplementary Methods.

***Kinetics of phenol hydrogenation.*** 10 mM PhOH solutions were prepared in water and 20 mM phosphate buffer, respectively. The pH of the above solutions was adjusted to 7.0 ± 0.1 using dilute HCl or NaOH. In a typical reaction, the catalyst (20 mg) and PhOH solution (4 mL, 0.010 M) were added to a 10 mL glass tube. The tube was sealed with a septum and purged with hydrogen for 10 min at rate of 20 $cm^3$ $min^{-1}$. After 10 min, the pressure relief needle was removed and the hydrogen supply was stopped. The reaction tube was then placed in an oil bath that was maintained at 35 °C. After the desired time, the septum was removed and the solution was extracted using ethyl acetate (1.0 mL 4 times). A 50 µL aliquot of the extracted solution was taken and added to 1.00 mL of ethanol along with a resorcinol internal standard (100 µL, 55 mM) and

analyzed in an Agilent GC-MS instrument (7890A, 5975C) with a HP-5MS column. The run started at 60 °C, and the temperature was then ramped to 150 °C at 5 °C $min^{-1}$. Then the temperature was ramped to 300 °C at 20 °C $min^{-1}$ and then it was held at 300 °C for 3 min.

**NMR Spectroscopy.** $^{13}C$ DEST, $^{13}C$ RD, $^{13}C$ $R_1$, and $^{1}H$ RD NMR data were collected at 25 °C on Bruker 600 and 800 MHz spectrometers equipped with triple resonance z-gradient cyroprobes and using the pulse sequences shown in **Supplementary Figure 5.S2**. Details on the NMR data acquisition and analysis are provided in Supplementary Methods.

## References


1. Libich, D.S., Fawzi, N.L., Ying, J. & Clore, G.M. Probing the transient dark state of substrate binding to GroEL by relaxation-based solution NMR. Proc Natl Acad Sci U S A 110, 11361-11366 (2013).

2. Fawzi, N.L., Ying, J., Ghirlando, R., Torchia, D.A. & Clore, G.M. Atomic-resolution dynamics on the surface of amyloid-beta protofibrils probed by solution NMR. Nature 480, 268-272 (2011).

3. Ceccon, A., Tugarinov, V., Bax, A. & Clore, G.M. Global Dynamics and Exchange Kinetics of a Protein on the Surface of Nanoparticles Revealed by Relaxation-Based Solution NMR Spectroscopy. J Am Chem Soc 138, 5789-5792 (2016).

4. Fusco, G. et al. Structural basis of synaptic vesicle assembly promoted by alpha-synuclein. Nat Commun 7, 12563 (2016).

5. Ceccon, A., Tugarinov, V. & Clore, G.M. TiO2 Nanoparticles Catalyze Oxidation of Huntingtin Exon 1-Derived Peptides Impeding Aggregation: A Quantitative NMR Study of Binding and Kinetics. J Am Chem Soc 141, 94-97 (2019).

6. Zanzoni, S., Pedroni, M., D'Onofrio, M., Speghini, A. & Assfalg, M. Paramagnetic Nanoparticles Leave Their Mark on Nuclear Spins of Transiently Adsorbed Proteins. J Am Chem Soc 138, 72-75 (2016).

7. Anthis, N.J. & Clore, G.M. Visualizing transient dark states by NMR spectroscopy. Q Rev Biophys 48, 35-116 (2015).

8. Nelson, N.C., Manzano, J.S., Sadow, A.D., Overbury, S.H. & Slowing, I.I. Selective Hydrogenation of Phenol Catalyzed by Palladium on High-Surface-Area Ceria at Room Temperature and Ambient Pressure. ACS Catal 5, 2051-2061 (2015).

9. Egner, T.K., Naik, P., Nelson, N.C., Slowing, II & Venditti, V. Mechanistic Insight into Nanoparticle Surface Adsorption by Solution NMR Spectroscopy in an Aqueous Gel. Angew Chem Int Ed Engl 56, 9802-9806 (2017).

10. Clore, G.M. et al. Deviations from the simple two parameter model free approach to the interpretation of 15N nuclear magnetic relaxation of proteins. J Am Chem Soc 112, 4989-4991 (1990).

11. Sachleben, J.R., Colvin, V., Emsley, L., Wooten, E.W. & Alivisatos, A.P. Solution-State NMR Studies of the Surface Structure and Dynamics of Semiconductor Nanocrystals. J Phys Chem B 102, 10117-10128 (1998).

12. Li, G., Han, J., Wang, H., Zhu, X.F. & Ge, Q. Role of Dissociation of Phenol in Its Selective Hydrogenation on Pt(111) and Pd(111). ACS Catal 5, 2009−2016 (2015).

13. Farmer, J.A. & Campbell, C.T. Ceria maintains smaller metal catalyst particles by strong metal-support bonding. Science 329, 933-936 (2010).

14. Ta, N. et al. Stabilized gold nanoparticles on ceria nanorods by strong interfacial anchoring. J Am Chem Soc 134, 20585-20588 (2012).

15. Wang, X., Liu, B. & Liu, J. DNA-Functionalized Nanoceria for Probing Oxidation of Phosphorus Compounds. Langmuir 34, 15871-15877 (2018).

16. Finocchio, E., Daturi, M., Binet, C., Lawalley, J.C. & Blanchard, G. Thermal evolution of the adsorbed methoxy species on CexZr1-xO2 solid solution samples: a FT-IR study. Catal Today 52, 53-63 (1999).

17. Leino, E. et al. Synthesis and characterization of ceria-supported catalysts for carbon dioxide transformation to diethyl carbonate. Catal Today 306, 128-137 (2018).

18. Mai, H.-X. et al. Shape-Selective Synthesis and Oxygen Storage Behavior of Ceria Nanopolyhedra, Nanorods, and Nanocubes. J Phys Chem B 109, 24380-24385 (2005).

## Figures

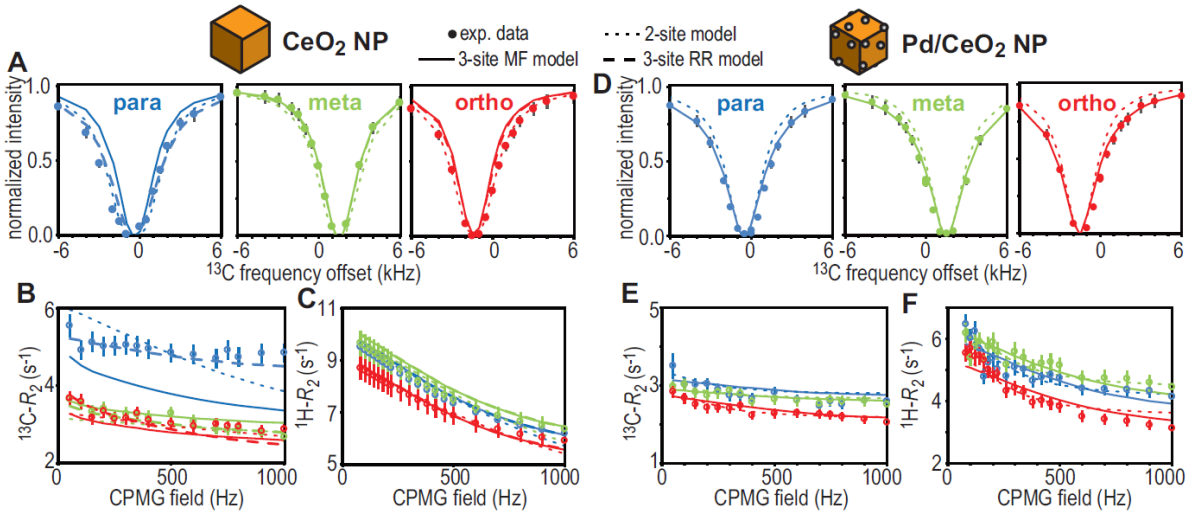


**Figure 5.1.** SC-NMR characterization of PhOH adsorption on $CeO_2$ and $Pd/CeO_2$ NPs. $^{13}C$ DEST profiles (saturation field = 1 kHz) measured for PhOH in the presence of (A) $CeO_2$ and (D) $Pd/CeO_2$. $^{13}C$ RD profiles measured for PhOH in the presence of (B) $CeO_2$ and (E) $Pd/CeO_2$. $^{1}H$ RD profiles measured for PhOH in the presence of (C) $CeO_2$ and (F) $Pd/CeO_2$. Data were acquired at 800 MHz. Experimental data for the para, meta, and ortho positions are blue, green, and red circles, respectively. The dotted line shows best fit to a two-site exchange model. The solid line shows best fit to a three-site exchange model in which local dynamics of the tightly adsorbed state are described using Model Free theory (with a global order parameter, S2). The dashed line shows best fit to a three-site exchange model in which local dynamics of the tightly adsorbed state are described using a Restricted Rotation model in which PhOH is allowed fast rotation about the C-O bond (Figure 3 and Supplementary Methods). An expanded version of this Figure is shown in Supplementary Figure S6. The full set of $^{13}C$ DEST, $^{13}C$ RD, and $^{1}H$ RD data is shown in Supplementary Figures S7-S11.

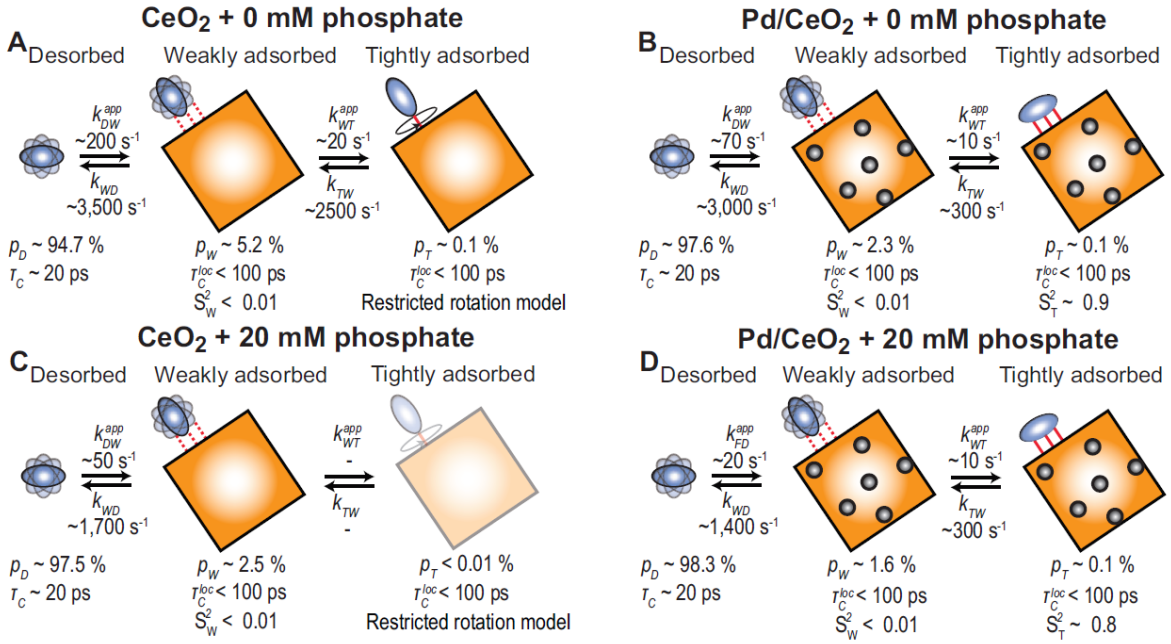


**Figure 5.2.** Quantitative interpretation of the SC-NMR data. SC-NMR data measured on 10 mM PhOH in the presence of (A) 1 wt% $CeO_2$, (B) 1wt% $Pd/CeO_2$, (C) 1 wt% $CeO_2$ and 20 mM Pi, and (D) 1 wt% $Pd/CeO_2$ and 20 mM Pi are interpreted using a three-site exchange model. Ceria is an orange cube. Pd NPs are gray spheres. PhOH is a blue oblate. Solid and dashed red lines represent strong and weak interactions, respectively. Rate constants for each equilibrium (ksy, where x and y are the states in thermodynamic equilibrium; $k_{xy}^{app}$ indicates the apparent rate constant), populations of each state ($p_D$, $p_W$, and $p_T$ for the desorbed, weakly adsorbed, and tightly adsorbed states, respectively), the order parameter ($S^2$), and the rotational correlation time for local motion ($\tau_c^{loc}$) were fit as global parameters. Chemical shift changes associated with each equilibrium were fit as peak specific parameters (Supplementary Methods). Best fit parameters are shown. A full report of the fitted parameters is shown in Supplementary Tables S1 and S2.

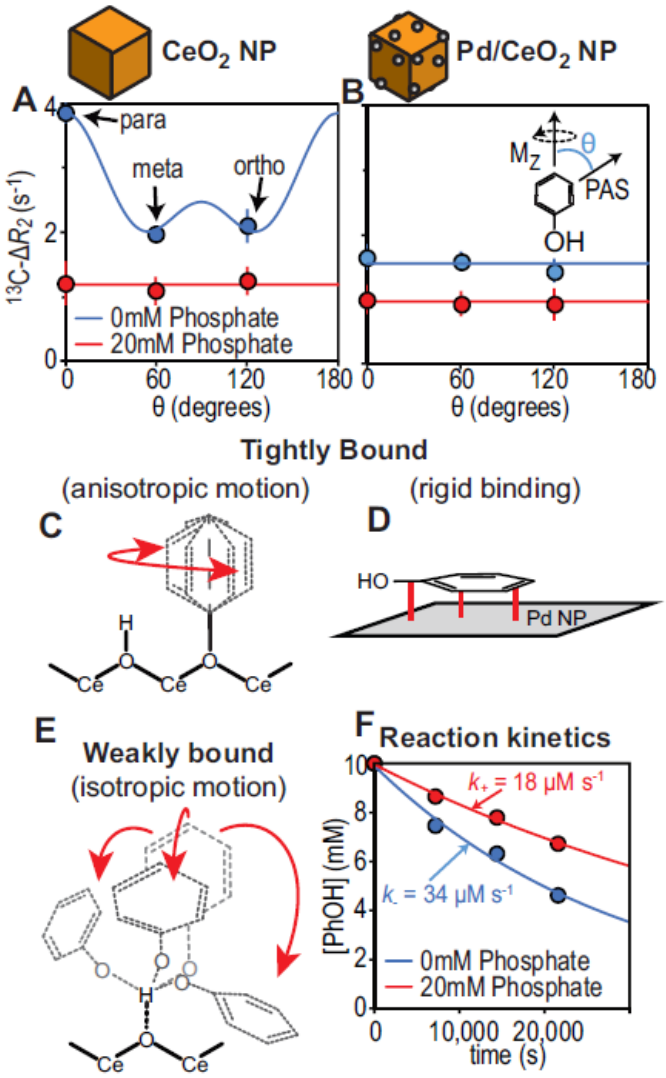


**Figure 5.3.** Structures of PhOH bound to $CeO_2$ and $Pd/CeO_2$ NPs. Angular dependency of the $^{13}C$ $\Delta R_2$ measured for PhOH in the presence of 1 wt% (A) $CeO_2$ and (B) $Pd/CeO_2$. θ is defined as the angle between the principle axis system of the relaxation vector (i.e. the C-H bond vector, PAS – see Supplementary Methods) and the axis of the restricted rotation (i.e. the axis of the C-O bond, Mz; see top-right corner in panel B). Experimental data are circles. Simulated data are solid curves. Blue and red data were measured in the presence of 0 and 20 mM phosphate, respectively. Structural models for tightly adsorbed PhOH on (C) $CeO_2$ (note that the oxygen of PhOH occupies an oxygen vacancy on ceria) and (D) $Pd/CeO_2$, and (E) of the weakly adsorbed state. (F) Kinetics of PhOH hydrogenation catalyzed by $Pd/CeO_2$ in the presence of 0 mM (blue) and 20 mM (red) phosphate. Experimental data are circles. Fitted exponential decays are solid curves. Pseudo first order rate constants in the absence and in the presence of phosphate are 34 ± 6 and 18 ± 2 μM $^{s-1}$, respectively.

## Supplementary Methods

***Synthesis of CeO2 and Pd/CeO2 nanoparticles.*** Cerium (III) nitrate hexahydrate ($Ce(NO_3)_3 \cdot 6H_2O$), Palladium (II) acetate ($Pd(O_2CCH_3)_2$) and phenol ($C_6H_5OH$), were purchased from Sigma-Aldrich. Sodium hydroxide (NaOH) in pellets, monosodium phosphate ($NaH_2PO_4$), disodium phosphate (Na2HPO4), and ethyl acetate were purchased from Fisher Scientific. Acetone was purchased from Barton Solvents. Ethanol was purchased from Decon labs. Deionized water (17.4 MΩ) was produced in house with a Thermo Scientific Barnstead E-pure system. All chemicals were used without further purification. Ceria cubes were synthesized following a literature procedure *(1)*. In brief $Ce(NO_3)_3 \cdot 6H_2O$ (6.96 g, 16 mmol) and sodium hydroxide (76.8 g, 1.92 mol) were dissolved in 40 and 280 mL of deionized water, respectively. Then, the two solutions were mixed in a teflon bottle, stirred at 550 rpm for 30 min to obtain a milky slurry. The teflon bottle containing the mixture was then set in a stainless steel autoclave vessel and heated at 180 °C for 24 h. After cooling down to room temperature the resulting solution was centrifuged to separate a white precipitate. The precipitate was washed several times with deionized water, and dried overnight at 60 °C in air. The obtained light yellow powders were then calcined at 450 °C for 4h with a ramp rate of 1°C $min^{-1}$ in air. Pd catalysts were prepared by impregnation with a 1 wt % Pd loading relative to the mass of the support. Palladium acetate (21.3 mg, 0.09 mmol) was dissolved in 5 mL acetone assisted by sonication. The support (1 g, ceria cubes) was placed in a mortar and impregnated with the Pd solution in 0.1 mL increments. After each impregnation step, the catalyst was mixed thoroughly with a pestle and dried in oven for 5 min (60 °C). The material was calcined at 350 °C for 2 h with a 2.5 °C $min^{-1}$ ramp rate and then reduced under flowing hydrogen (10 $cm^3$ $min^{-1}$) at 350 °C for 2 h with a ramp rate of 2.5 °C $min^{-1}$.

***Palladium dispersion and surface area*.** The palladium dispersion of the catalysts was calculated from chemisorption data on the basis of the equation:

$$D\% = 100\frac{S_f\, M\, V_{ad}}{m\, W\, V_m} \quad (1)$$

where $S_f$ = stoichiometry factor (Pd/$H_2$ molar ratio) = 2, M = atomic mass of Pd (106.42 g $mol^{-1}$), $V_{ad}$ = volume of chemisorbed $H_2$ (mL) under standard temperature and pressure (STP) conditions, $m$ = mass of the sample (g), $W$ = weight fraction of Pd in the sample as determined by ICP-OES, and $V_m$ = molar volume of $H_2$ (22,414 mL $mol^{-1}$) under STP conditions. The cubic crystallite size of palladium was calculated on the basis of the equation:

$$\text{Pd crystallite size (nm)} = \frac{6\, M}{m\, W\, d_{Pd}\, N_a\, SA_{Pd}} \quad (2)$$

$$\text{Pd surface area } (m^2 g^{-1} \text{ of Pd}) = \frac{S_f\, N_a\, SA_{cross}\, V_{ad}}{m\, W\, V_m} \quad (3)$$

where M = atomic mass of Pd (106.42 g $mol^{-1}$), m = mass of the sample (g), W = weight fraction of Pd in the sample as determined by ICP-OES; $d_{Pd}$ = density of palladium ($1.202 \times 10^{-20}$ g $nm^{-3}$), $N_a$= Avogadro's number ($6.023 \times 10^{23}$ $mol^{-1}$); $SA_{cross}$ = palladium cross sectional area ($7.87 \times 10^{-20}$ $m^2$), $SA_{Pd}$ = palladium surface area from the equation above ($nm^2$ $g^{-1}$ of Pd). The number 6 is derived from assuming a cubic geometry.

***Kinetics of phenol hydrogenation.*** The phenol solution was prepared in water and buffer solvent, respectively. For the buffered condition, a phosphate buffer (20 mM) was produced by mixing 0.11 g of monosodium phosphate and 0.33 g of disodium phosphate in 100 mL water. Phenol (18.8 mg, 0.2 mmol) was added to 20 mL of the above buffer solution to get 10 mM phenol concentration. Similarly, a 10 mM solution of phenol in water was prepared dissolving 94.1 mg (1 mmol) in 100 mL of water. The pH of the above solutions was adjusted to 7.0 ± 0.1 using dilute

HCl or NaOH if necessary. In a typical reaction the catalyst (20 mg) and phenol solution (4 mL, 0.010 M) were added to a 10 mL glass tube. The tube was sealed with a septum and purged with hydrogen for 10 min at rate of 20 $cm^3$ $min^{-1}$. After 10 min, the pressure relief needle was removed and the hydrogen supply was stopped. The reaction tube was then placed in an oil bath that was maintained at 35 °C. After the desired time, the septum was removed and the solution was extracted using ethyl acetate (1.0 mL 4 times). A 50 µL aliquot of the extracted solution was taken and added to 1.00 mL of ethanol along with a resorcinol internal standard (100 µL, 55 mM) and analyzed in an Agilent GC-MS instrument (7890A, 5975C) with a HP-5MS column. The run started at 60 °C, and the temperature was then ramped to 150 °C at 5 °C $min^{-1}$. Then the temperature was ramped to 300 °C at 20 °C $min^{-1}$ and then it was held at 300 °C for 3 min.

***NMR Spectroscopy.*** NMR data for all samples were collected at 25 °C on either a Bruker 600 or 800 MHz spectrometer equipped with triple resonance z-gradient cryoprobes. All experiments used 32 dummy scans and 64 scans with a recycle delay of 5 seconds between each scan. $^{13}C$ DEST experiments were performed using two saturation field strengths (500 and 1,000 Hz), applied for 1 second, and using the pulse program described in **Figure 5.S2**. A total of 32 different offsets between ±25kHz were recorded, including saturation power controls where the field was set to 0 Hz. $^{13}C$ CPMG experiments were performed using field strengths between 100 and 1,000 Hz by varying the delay time between 180° pulses in the π-pulse train. Relaxation delays were in the range of 100 to 500 ms. $^{13}C$ $R_1$ values were measured according to the pulse sequence in **Figure 5.S2**. $^{1}H$ CPMG were recorded using the PROJECT sequence *(2)*. NMR spectra were processed and analyzed using Mnova NMR software (http://mestrelab.com/software/mnova/nmr/). Experimental error was evaluated to be ≤ 5 % of

the measured value for DEST and ≤ 10 % of the measured value for RD and $R_1$ experiments by running repeats of the above experiments.

***Data Analysis and Fitting.*** The $^{13}$C RD, $^{13}$C $R_1$, $^{13}$C DEST, and $^{1}$H RD data measured at two spectrometer fields (for the RD and R1 experiments) and two saturation field strengths (for the DEST experiment) were fit globally using either a two- or three-site exchange model with characteristic motional dynamics at each site. Experimental data were compared with model based simulated data using the following objective function:

$$\chi^2 = (\chi^2_{C-RD} + \chi^2_{C-R1} + \chi^2_{C-DEST} + \chi^2_{H-RD})/n_{df} \quad (4)$$

Where $\chi^2_{C-RD}$, $\chi^2_{C-R1}$, $\chi^2_{C-DEST}$, and $\chi^2_{H-RD}$ are the objective functions for the $^{13}$C RD, $^{13}$C R1, $^{13}$C DEST, and $^{1}$H RD experiments, respectively, and $n_{df}$ is the number of degrees of freedom. The $\chi^2$ functions for each experiment are defined as following:

$$\chi^2_{C-RD} = \sum_i \sum_j \sum_k \left(R_2^{obs,i,j,k} - R_2^{calc,i,j,k}\right)^2 / \sigma_{C-RD}^{i,j,k} \quad (5)$$

$$\chi^2_{C-R1} = \sum_i \sum_j \left(R_1^{obs,i,j} - R_1^{calc,i,j}\right)^2 / \sigma_{C-R1}^{i,j} \quad (6)$$

$$\chi^2_{C-DEST} = \sum_i \sum_j \sum_l \sum_m \left(I^{obs,i,j,l,m} - I^{calc,i,j,l,m}\right)^2 / \sigma_{C-DEST}^{i,j,l,m} \quad (7)$$

$$\chi^2_{H-RD} = \sum_i \sum_j \sum_k \left(R_2^{obs,i,j,k} - R_2^{calc,i,k,j}\right)^2 / \sigma_{H-RD}^{i,j,k} \quad (8)$$

where the summation indices *i*, *j*, *k*, *l*, and *m* refer to the peak index, spectrometer frequency, CPMG field, saturation field, and frequency offset, respectively, the *calc* and *obs* superscripts denote calculated and experimental data, respectively, *I* indicates the normalized peak intensity, and $\sigma$ is the experimental error associated with each data point.

***Simulation of $^{1}$H and $^{13}$C RD experiments.*** $^{1}$H- and $^{13}$C-$R_2$ rates for any specific value of *i*, *j*, and *k* were calculated as:

$$R_2^{calc,i,j,k} = \frac{\ln(M(t_1)/M(t_2))}{t_2 - t_1} \quad (9)$$

where $t_2$ is the constant relaxation delay during the RD experiment and $t_1$ is 0 seconds. *M(t)* is

the transverse magnetization, which is given by:

$$M(t) = (AA^*A^*A)^n M(0) \quad (10)$$

$$M(0) = \begin{bmatrix} M^D \\ M^W \\ M^T \end{bmatrix} \quad (11)$$

*M(0)* is the initial magnetization for each site in the exchange model. Here we use the example of a three-site exchange, and $M^D$, $M^W$, and $M^T$ are the transverse magnetizations of sites D, W, and T (see main text, Figure 2), respectively. The parameter n is the number of CPMG cycles used, *A** is the complex-conjugate of *A*, and *A* is given by:

$$A = e^{-R\tau_{CP}/2} \quad (12)$$

$\tau_{CP}$ is the time delay between two π pulses during the CPMG pulse train. *R* is given by the summation:

$$R = iR^{CS} + R^{rel} + R^{ex} \quad (13)$$

$$R^{CS} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -\Delta\omega^{DW} & 0 \\ 0 & 0 & -\Delta\omega^{DT} \end{bmatrix} \quad (14)$$

$$R^{rel} = \begin{bmatrix} R_2^D & 0 & 0 \\ 0 & R_2^W & 0 \\ 0 & 0 & R_2^T \end{bmatrix} \quad (15)$$

$$R^{ex} = \begin{bmatrix} k_{DW}^{app} + k_{DT}^{app} & -k_{WD} & -k_{TD} \\ -k_{DW}^{app} & k_{WD} + k_{WT} & -k_{TW} \\ -k_{DT}^{app} & -k_{WT} & k_{TW} + k_{TD} \end{bmatrix} \quad (16)$$

for a three-site connected exchange model (matrices for the two-site exchange model are shown elsewhere (3)). $\Delta\omega^{pq}$ is the difference in frequency between sites p and q (in rad $s^{-1}$), $R_2^p$ is the intrinsic transverse relaxation rate of site p, and $k_{pq}$ is the exchange rate between sites p and q. The formulation of $R^{ex}$ changes depending on the connectivity of the exchange model. For a linear three site exchange, all instances of $k_{DT}^{app}$ and $k_{TD}$ will be set to zero. The intrinsic relaxation rate of the free ligand ($R_2^D$) was measured experimentally, the rate for other states ($R_2^W$ and $R_2^T$) were

fit phenomenologically (in case of $^{1}$H $R_2$ rates) or calculated based on motional models (in case of $^{13}$C $R_2$ rates) (see *$^{13}$C $R_1$/$R_2$ Relaxation Modeling*).

***Simulation of $^{13}$C $R_1$ experiments.*** $^{13}$C $R_1$ values from inversion recovery experiments are calculated in a very similar manner to the above procedure for $R_2$:

$$R_1^{calc,i,j} = \frac{\ln(M(t_1)/M(t_2))}{t_2 - t_1} \quad (17)$$

Here $t_2$ is the relaxation delay and $t_1$ is 0 seconds. *M(t)* is the longitudinal magnetization which is given by:

$$M(t) = (AA^*A^*A)^n M(0) \quad (18)$$

$$M(0) = \begin{bmatrix} M^D \\ M^W \\ M^T \end{bmatrix} \quad (19)$$

*M(0)* is the initial magnetization for each site in the exchange model (here we use the example of a three-site exchange). *A** is the complex-conjugate of *A*, and *A* is given by:

$$A = e^{-Rt} \quad (20)$$

*t* is the time delay during the relaxation evolution. *R* is given by the summation:

$$R = R^{rel} + R^{ex} \quad (21)$$

$$R^{rel} = \begin{bmatrix} R_1^D & 0 & 0 \\ 0 & R_1^W & 0 \\ 0 & 0 & R_1^T \end{bmatrix} \quad (22)$$

$$R^{ex} = \begin{bmatrix} k_{DW}^{app} + k_{DT}^{app} & -k_{WD} & -k_{TD} \\ -k_{DW}^{app} & k_{WD} + k_{WT} & -k_{TW} \\ -k_{DT}^{app} & -k_{WT} & k_{TW} + k_{TD} \end{bmatrix} \quad (23)$$

for a three-site connected exchange model. The definitions of the variables are the same as for the $R_2$ simulation with the addition of $R_1^p$, being the intrinsic longitudinal relaxation rate of site p. The relaxation rate of the free ligand ($R_1^D$) is measured experimentally but the bound state $^{13}$C relaxation rates ($R_1^W$ and $R_1^T$) are calculated from motional models (see *$^{13}$C $R_1$/$R_2$ Relaxation Modeling*).

***Simulation of $^{13}C$ DEST experiments.*** Signal intensities for the simulated DEST experiment are calculated with the following differential equation.

$$\frac{d}{dt}\begin{bmatrix} E/2 \\ I_x^D \\ I_y^D \\ I_z^D \\ I_x^W \\ I_y^W \\ I_z^W \\ I_x^T \\ I_y^T \\ I_z^T \end{bmatrix} = -\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & R_2^{D*} & \Omega^D & -\omega_y & -k_{WD} & 0 & 0 & -k_{TD} & 0 & 0 \\ 0 & -\Omega^D & R_2^{D*} & \omega_x & 0 & -k_{WD} & 0 & 0 & -k_{TD} & 0 \\ -2\Theta^D & \omega_y & -\omega_x & R_1^{D*} & 0 & 0 & -k_{WD} & 0 & 0 & -k_{TD} \\ 0 & -k_{DW}^{app} & 0 & 0 & R_2^{W*} & \Omega^W & -\omega_y & -k_{TW} & 0 & 0 \\ 0 & 0 & -k_{DW}^{app} & 0 & -\Omega^W & R_2^{W*} & \omega_x & 0 & -k_{TW} & 0 \\ -2\Theta^W & 0 & 0 & -k_{DW}^{app} & \omega_y & -\omega_x & R_1^{W*} & 0 & 0 & -k_{TW} \\ 0 & -k_{DT}^{app} & 0 & 0 & -k_{WT} & 0 & 0 & R_2^{T*} & \Omega^T & -\omega_y \\ 0 & 0 & -k_{DT}^{app} & 0 & 0 & -k_{WT} & 0 & -\Omega^T & R_2^{T*} & \omega_x \\ -2\Theta^T & 0 & 0 & -k_{DT}^{app} & 0 & 0 & -k_{WT} & \omega_y & -\omega_x & R_1^{T*} \end{bmatrix} \quad (24)$$

This formulation is for a connected three-site exchange model. Here, $\Theta^p = R_1^p I_z^p$ (where $I_z^p$ is the equilibrium longitudinal magnetization of state p), $R_2^{p*} = R_2^p + k_{pq} + k_{pr}$, $R_1^{p*} = R_1^p + k_{pq} + k_{pr}$, I represents the magnetization of a $^{13}C$ nucleus in the rotating frame, $\Omega^p$ is the difference between the resonance frequency of state *p* and the applied saturation field, $\omega_{y,x}$ is the continuous wave saturation field strength on the applied axis, and E is unity. The value of $I^{calc}$ is determined by the ratio of the solution to the above equation with and without saturation applied, respectively.

***$^{13}C$ $R_1$/$R_2$ Relaxation Modeling.*** Relaxation rates ($R_1$ and $R_2$) for $^{13}C$ spins that are bound to a nanoparticle surface are calculated based on rotational motion. This includes contributions from the tumbling of the nanoparticle as well as local motion of phenol. The general equation to calculate $R_1$ and $R_2$ considering dipolar and CSA relaxation mechanisms are:

$$R_1 = \frac{d^2}{4}(J(\omega_H - \omega_C) + 3J(\omega_C) + 6J(\omega_C + \omega_H)) + c^2 J(\omega_C) \quad (25)$$

$$R_2 = \frac{d^2}{8}(4J(0) + J(\omega_H - \omega_C) + 3J(\omega_C) + 6J(\omega_H) + 6J(\omega_C + \omega_H)) + \frac{c^2}{6}(4J(0) + 3J(\omega_C)) \quad (26)$$

Where:

$$d = \frac{\mu_0}{4\pi}\hbar\gamma_C\gamma_H r^{-3} \quad (27)$$

$$c = \frac{\gamma_C B_0 \Delta\sigma}{\sqrt{3}} \quad (28)$$

$$J(\omega) = \frac{2}{5}S^2\left[\frac{\tau_{NP}}{1+\tau_{NP}^2\omega^2}\right] + \frac{2}{5}(1-S^2)\left[\frac{\tau_e}{1+\tau_e^2\omega^2}\right] \quad (29)$$

$$\frac{1}{\tau_e} = \frac{1}{\tau_{NP}} + \frac{1}{\tau_{loc}} \quad (30)$$

$\mu_0$ is the magnetic permittivity, $\hbar$ is the reduced Planck's constant, $\gamma_u$ is the gyromagnetic ratio of the specified nuclei, $r$ is the internuclear bond distance (taken to be 1.09 Å here), $J(\omega)$ is the spectral density function, $\tau_{NP}$ and $\tau_{loc}$ are the rotational correlation time for the nanoparticle and the local motion of the ligand, respectively. $\Delta\sigma$ is the chemical shift anisotropy (for aromatic C-H is approximately 200 ppm) *(4, 5)*. $S^2$ is the generalized order parameter for use in model free analysis of motion. $S^2$ is a fittable parameter whose value is the same regardless of the C-H bond vector due to the rigid structure of phenol.

The Restricted Rotation (RR) model of motion allows for the calculation of $S^2$ instead of fitting it. Here we use a model that allows for a single rotational axis, as described by *(6)* :

$$J(\omega) = \frac{2}{5}[Y_2^0(\theta)]^2\left[\frac{\tau_{NP}}{1+\tau_{NP}^2\omega^2}\right] + \frac{2}{5}[Y_2^1(\theta)]^2\left[\frac{\tau_e}{1+\tau_e^2\omega^2}\right] \quad (31)$$

Here, θ is the angle between the axis of rotation and the principle axis system of the relaxation vector. For a proton attached $^{13}C$ in natural abundance PhOH, transverse relaxation is induced by $^{13}C$ CSA and dipolar coupling to the covalently attached hydrogen atom. As for a $^{13}C$ atom in a six-membered aromatic ring the least-shielded component of the CSA tensor ($\sigma_{11}$) is aligned with the $^{13}C$-$^{1}H$ bond *(4, 5)* and the dipolar coupling is also directed along the $^{13}C$-$^{1}H$ bond, θ is calculated as the angle between the axis of rotation and the C-H bond vector *(7)* (see Figure 3 in main text).

***Global fitting parameters.*** The optimization was run using $p_D$, $p_W$, $k_{WD}$, $k_{TW}$, $k_{DT}^{app}$, $\tau^{loc}$ of the weakly and tightly bound states ($\tau_W^{loc}$ and $\tau_T^{loc}$ , respectively), and the order parameters of the weakly and tightly bound states ($S_W^2$ and $S_T^2$, respectively) as global parameters. Remaining

values of populations and rate constants were calculated using the following equations:

$$p_T = 1 - p_D - p_W \quad (32)$$

$$k_{DW}^{app} = k_{WD}\, p_W / p_D \quad (33)$$

$$k_{WT}^{app} = k_{TW}\, p_T / p_W \quad (34)$$

$$k_{TD} = k_{DT}^{app}\, p_D / p_T \quad (35)$$

$\tau_{NP}$ was calculated to be 1.4 µs by using the average nanoparticle size (25 nm, Supplementary Figure S3) and the Stoke-Einstein equation. This value was held constant during the data fitting procedure.

Error on the fitted parameters was computed by Monte Carlo simulation of synthetic datasets from the estimated experimental error (see NMR spectroscopy section).

***Scripts and Pulse Programs*.** All Matlab scripts and pulse programs used here can be found at https://group.chem.iastate.edu/Venditti/downloads.html for download.

## Supplementary Figures

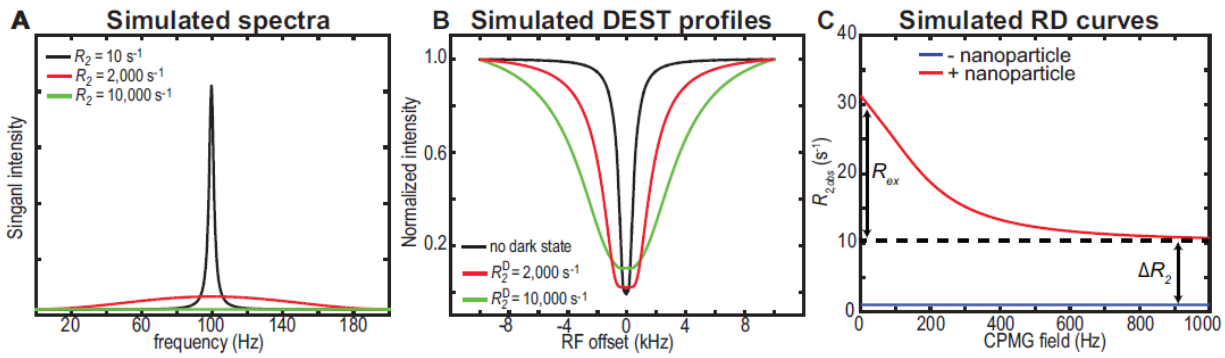


**Figure 5.1S.** Guide to interpretation of the DEST and RD experiments (see reference *(8)* for an exhaustive theoretical treatment). In the DEST experiment, the broad resonances of the small molecule bound to the slow-tumbling nanoparticle (see panel A) are selectively saturated by an off-resonance radiofrequency (RF) and saturation is transferred to the NMR-visible unbound ligand via chemical exchange. Simulated DEST profiles (saturation field strength = 150 Hz) for a dark state with transverse relaxation $R_2^D$ and population of 5 % in exchange ($k_{ex}$ = 600 s$^{-1}$) with a NMR visible state ($R_2$ = 10 s$^{-1}$) are shown in panel B. The presence of a high-molecular weight, NMR-invisible species in exchange with the visible NMR signal is detected as broadening in the DEST profile. If a small molecule is in exchange on the µs-ms timescale between the free state (characterized by small $R_2$) and the nanoparticle-bound state (characterized by large R2), and the free and bound state have different NMR chemical shifts, the shape of a typical RD profile measured for the small molecule NMR signals will be affected by the exchange as shown in panel C. At high CPMG field, the observed $R_2$ rate ($R_{2,obs}$) will be enhanced by lifetime line broadening ($R_{2,obs} = R_{2,free} + \Delta R_2$, in which $R_{2,free}$ is the $R_2$ value in the absence of nanoparticle and $\Delta R_2$ is lifetime line broadening). At a low CPMG field, $R_{2,obs}$ will be enhanced by both lifetime line broadening and chemical-exchange contribution to $R_2$ ($R_{2,obs} = R_{2,free} + \Delta R_2 + R_{ex}$, in which $R_{ex}$ is the exchange contribution to $R_2$). Increasing the CPMG field results in the progressive suppression of $R_{ex}$ and a decrease in $R_2$ (which value levels off when $R_{ex}$ is fully suppressed by the refocusing field), which introduces a curvature (dispersion) in the RD profile. Both RD and DEST profiles can be fit using the Bloch-McConnell theory to report on the kinetics ($k_{on}$ and $k_{off}$) and thermodynamics (populations of the free and bound states) of the small molecule-NP interactions (see Supplementary Methods). Note that DEST experiments are more sensitive to slow exchange process (ms-s timescale), while the curvature of an RD profile is mostly influenced by intermediate exchange (µs-ms timescale). $\Delta R_2$ values report on a much wider range of exchange processes extending from the slow to the fast exchange regime.

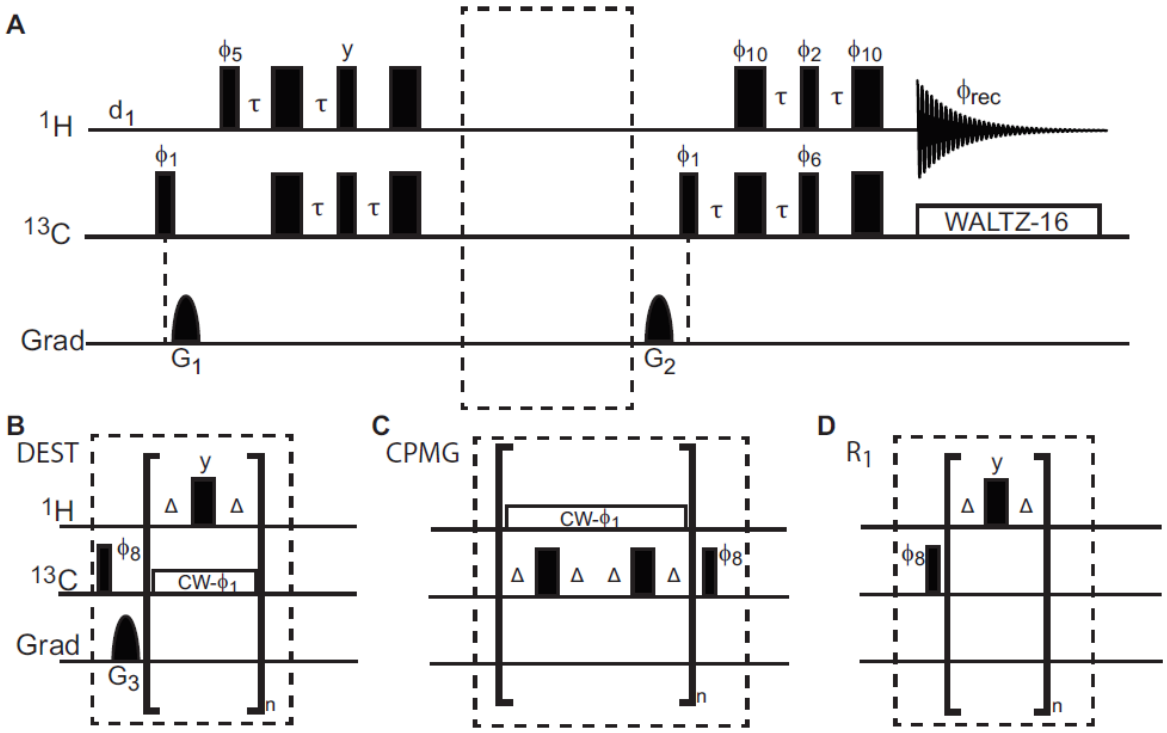


**Figure 5.2S.** Double refocused-INEPT based sequences used here for acquisition of all $^{13}C$ NMR data. $^{13}C$ DEST, CPMG, and $R_1$ were measured by inserting the blocks shown in B, C, and D, respectively, into the dashed box shown in A. Narrow and wide rectangular pulses are applied with flip angles of 90° and 180°, respectively, along the x-axis unless indicated otherwise. The delay τ is set to 1.56 ms. Δ is set to 50 ms and 25 ms for the DEST and R1 experiments, respectively. Δ is a variable parameter in the CPMG experiment that determines the CPMG field ($\upsilon_{cpmg}$ = 1/4/Δ Hz). CW is a continuous wave used for saturation of the dark-state in DEST and for proton decoupling in the CPMG experiment. The phase cycling employed is: $\phi_1$ = (x, −x); $\phi_2$ = 2(x), 2(−x); $\phi_5$ = 8(x), 8(-x); $\phi_6$ = 4(y), 4(−y); $\phi_8$ = (y, −y); $\phi_{10}$ = 16(x), 16(−x); $\phi_{rec}$ = 2(x), 4(y), 2(x), 2(y), 4(x), 2(y). The duration and strength of the gradients are as follows: $G_1$, 1 ms, 30 G/cm; $G_2$, 0.05 ms, 35 G/cm; $G_3$, 0.1 ms, 40 G/cm.

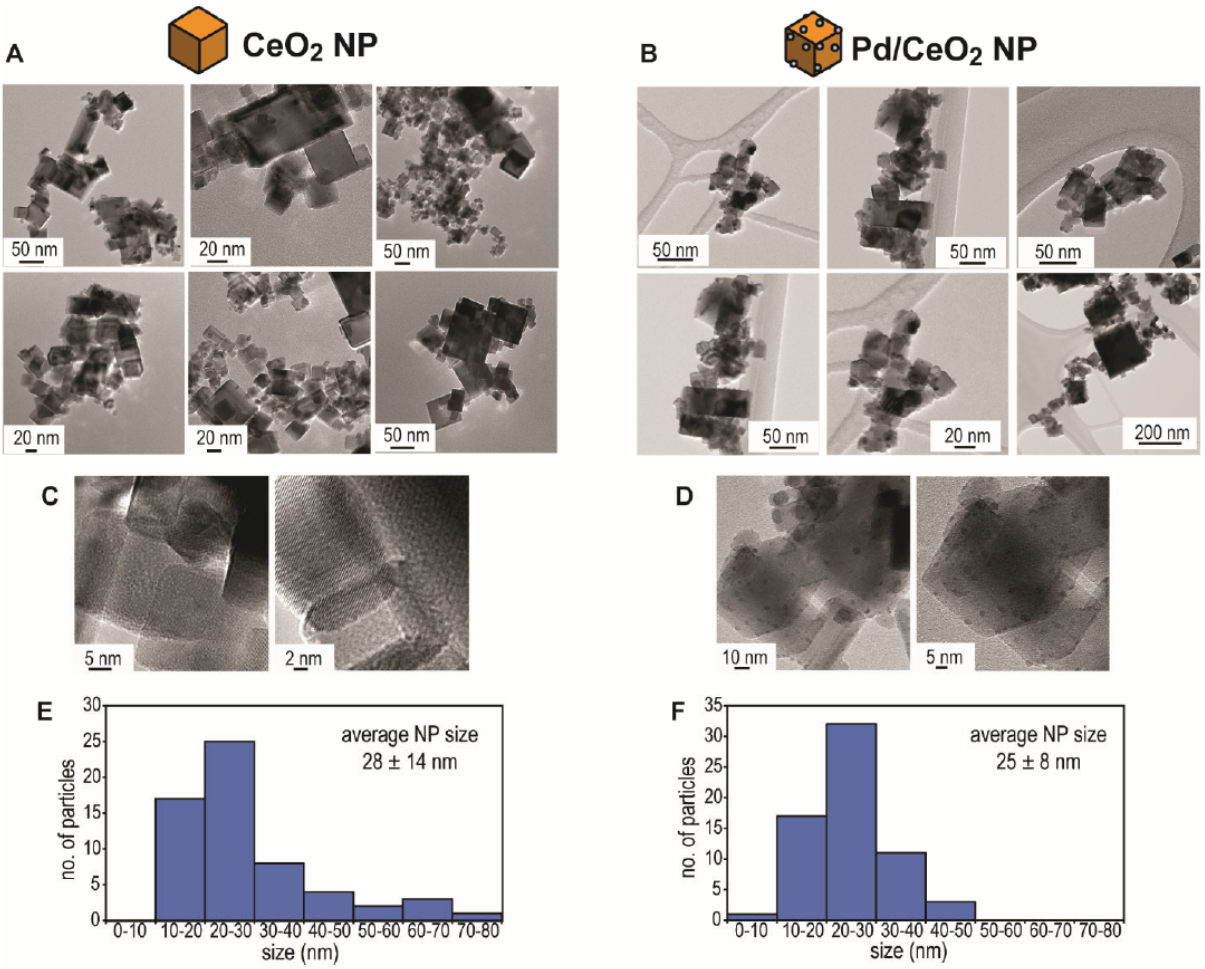


**Figure 5.3S.** TEM images of (A) $CeO_2$ and (B) $Pd/CeO_2$. High resolution (HR)-TEM images of (C) $CeO_2$ and (D) $Pd/CeO_2$. Histogram of ceria particle size distribution in (E) $CeO_2$ and (F) $Pd/CeO_2$. The size of Pd NPs in $Pd/CeO_2$ is ≤ 2 nm (note the darker spots on the surface of the ceria cube in panel D). Size distribution analysis was performed based on the TEM and HR- TEM images using the Digital micrograph software (see Supplementary Methods).

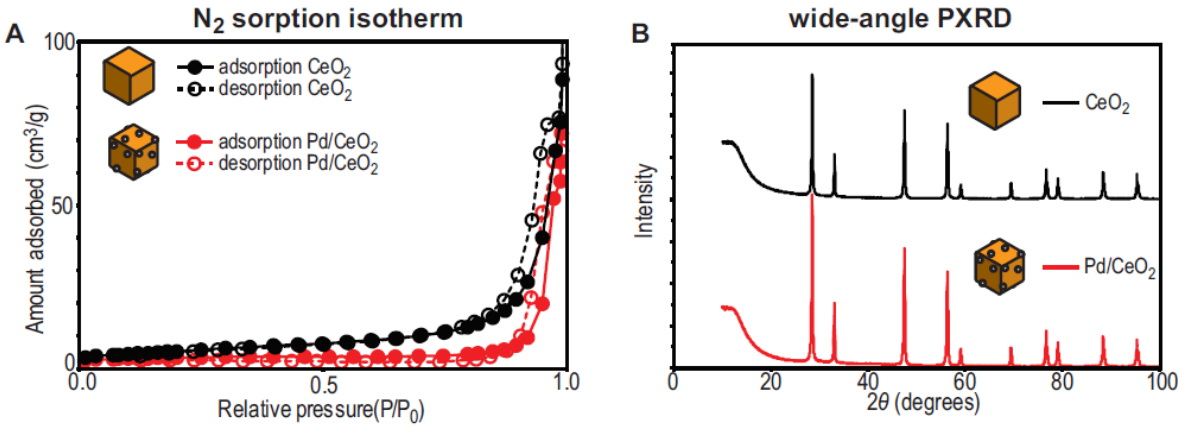


**Figure 5.4S.** (A) $N_2$ sorption isotherms of $CeO_2$ (black) and $Pd/CeO_2$ (red). (B) Wide-angle PXRD patterns of $CeO_2$ support (black) and $Pd/CeO_2$ catalyst (red).

The PXRD pattern of $CeO_2$ was indexed to the fluorite cubic structure of ceria (space group $F_{m3m}$ (No. 225), JCPDS 34-0394). Pd species could not be detected by PXRD analysis, suggesting that the dispersity of Pd nanoparticles over the support is high *(9)* and/or the crystallite size of Pd is small (≤ 2nm) *(10)*. These conclusions are in good agreement with the TEM data shown in Figure S3. Given the small size of the Pd NPs and an approximate molecular size for PhOH of ~ 0.8 nm, we expect a Pd NP to host 1 to 2 PhOH molecules.

Nitrogen sorption analysis gave surface area of 16.5 ± 1.0 m2/g for $CeO_2$ (see Supplementary Methods for details on data analysis). The additional thermal aging in reducing environment for incorporation of Pd (see Supplementary Methods) led to a lowering of the overall surface area (9.4 ± 2.3 $m^2/g$) of $Pd/CeO_2$ relative to $CeO_2$. Based on the metallic surface area of 2.9 $m^2/g$ obtained by H2-chemisorption measurements (see Supplementary Methods), the actual ceria surface area on the $Pd/CeO_2$ catalyst can be calculated to be 6.5 $m^2/g$ (i.e. 9.4 – 2.9 m2/g). In addition, $H_2$-chemisorption experiments return a Pd dispersion on the catalyst of 65 %, which is consistent with the TEM and PXRD data indicating that the metal is highly dispersed on the support.

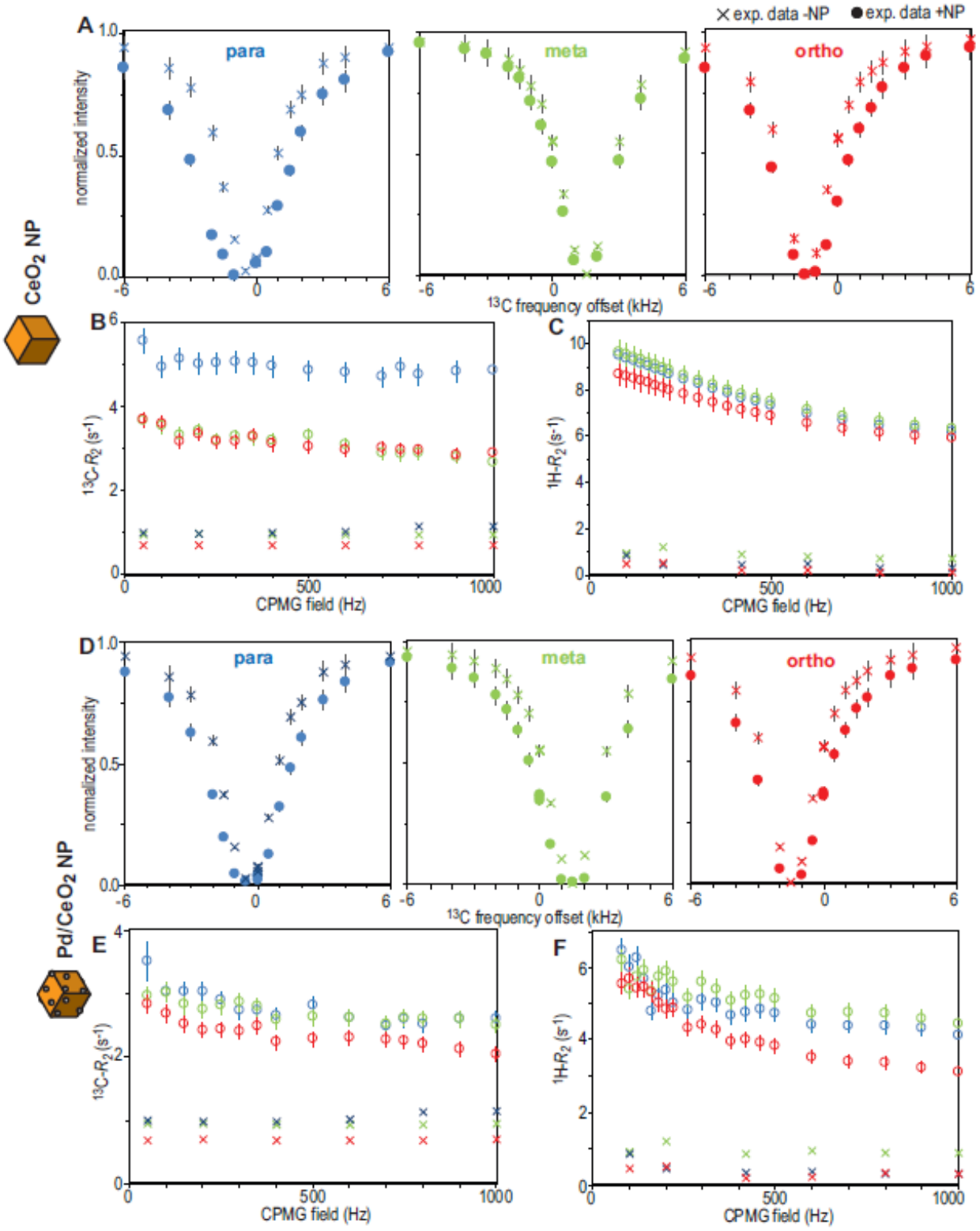


**Figure 5.5S.** Example of $^{13}C$ DEST profiles (saturation field = 1 kHz) measured for PhOH in the absence (cross) and in the presence (circle) of 1 wt% (A) $CeO_2$ and (D) $Pd/CeO_2$. Example of $^{13}C$ RD profiles measured for PhOH in the absence (cross) and in the presence (circle) of 1 wt%

(B) $CeO_2$ and (E) $Pd/CeO_2$. Example of $^1H$ RD profiles measured for PhOH in the absence (cross) and in the presence (circle) of 1 wt% (C) $CeO_2$ and (F) $Pd/CeO_2$. All data were acquired on an 800 MHz spectrometer. Data for the para, meta, and ortho positions are colored blue, green, and red, respectively.

These data indicate that addition of $CeO_2$ and $Pd/CeO_2$ introduce different perturbations in the experimental NMR data. In particular, fast, local rotation of PhOH about the C-O bond results in an increased $^{13}C$ $R_2$ for the para carbon in the tightly adsorbed state. The higher transverse relaxation of the para position results in a more pronounced broadening of its DEST profile and an increased $^{13}C$ $\Delta R_2$ if compared to the meta and ortho positions. Addition of $Pd/CeO_2$ produces similar effects in the DEST and RD profiles of the para, meta, and ortho position, indicating that PhOH undergoes isotropic tumbling when bound to $Pd/CeO_2$. The fitted order parameter, S2, of ~0.9 suggests that PhOH is rigidly attached to the $Pd/CeO_2$ nanoparticle in the tightly associated state.

It is also interesting to note that the DEST profile measured for the ortho position in the presence of $CeO_2$ is slightly broader than the one measured for the meta position. This experimental trend is not captured by the Restricted Rotation model, which implies identical $^{13}C$ $R_2$ values for the meta and ortho positions in the tightly associated state (Figure 1A in main text). As the ortho position is in close contact with the slightly paramagnetic surface of $CeO_2$ (Figure 3C in main text), additional relaxation mechanisms (other than the evolution of the C-H bond vector under local rotation of the C-O bond) may increase the 13C R2 of the ortho carbon in the tightly associated state. The increased transverse relaxation of the ortho position would result in a broader DEST profile and a larger $\Delta R_2$ than predicted by the Restricted Rotation model, therefore explaining the difference between modelled and experimental data in Figure 1 and Supplementary Figure S6.

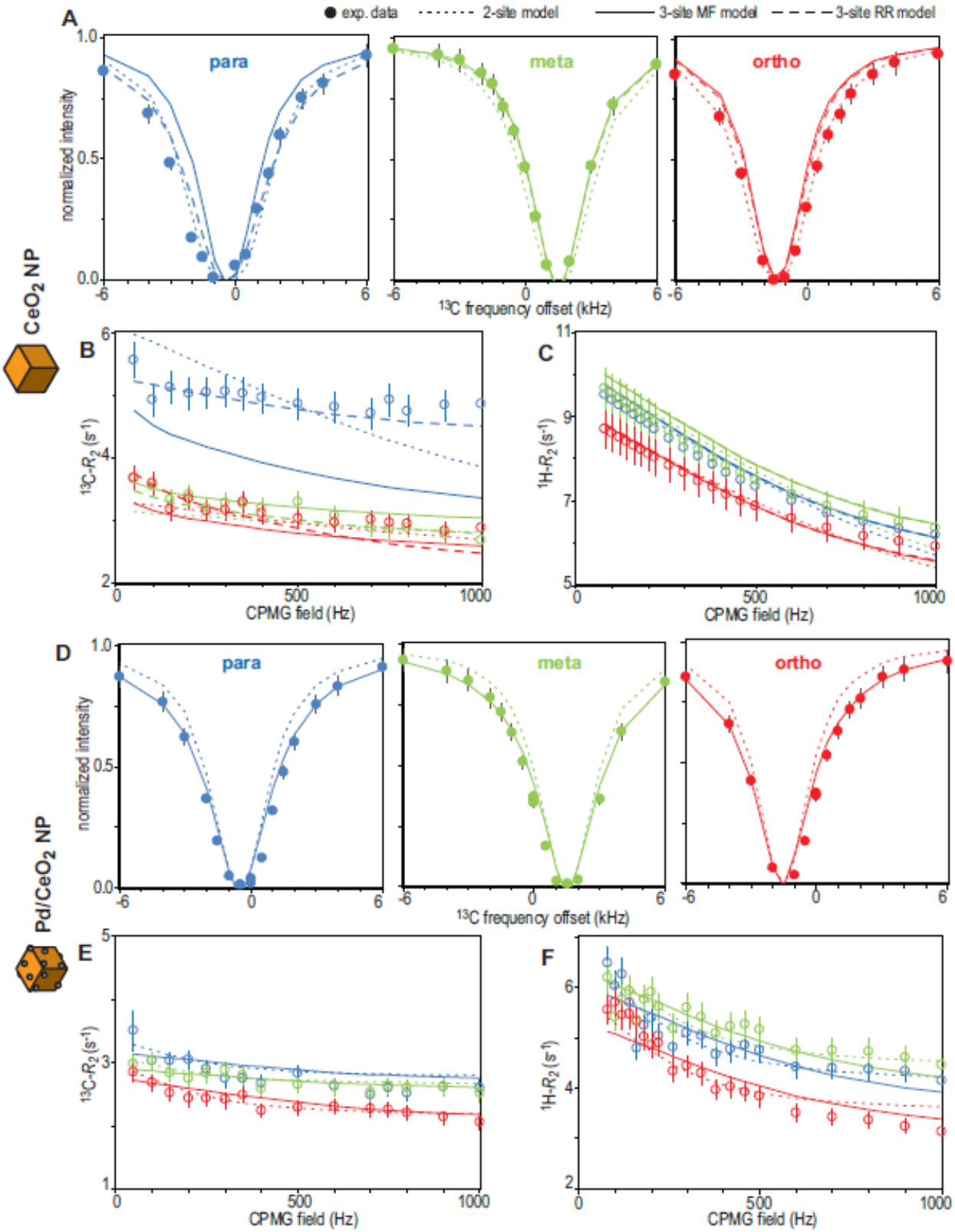


**Figure 5.6S.** This is an expanded version of Figure 1 in main text. Example $^{13}$C DEST profiles (saturation field = 1 kHz) measured for PhOH in the presence of (A) $CeO_2$ and (D) Pd/$CeO_2$.

Example $^{13}$C RD profiles measured for PhOH in the presence of (B) $CeO_2$ and (E) $Pd/CeO_2$.

Example $^{1}$H RD profiles measured for PhOH in the presence of (C) $CeO_2$ and (F) $Pd/CeO_2$. All data were acquired on an 800 MHz spectrometer. Experimental data for the para, meta, and ortho positions are shown as blue, green, and red circles, respectively (errors were measured by repeated measurements, see Supplementary Methods). The dotted line shows best fit to a two-site exchange model. The solid line shows best fit to a three-site exchange model in which local dynamics of the tightly adsorbed state are described using Model Free theory (with a global order parameter, $S_T^2$). The dashed line shows best fit to a three-site exchange model in which local dynamics of the tightly adsorbed state are described using a Restricted Rotation model in which PhOH is allowed fast rotation about the C-O bond (see Figure 3 and Supplementary Methods).

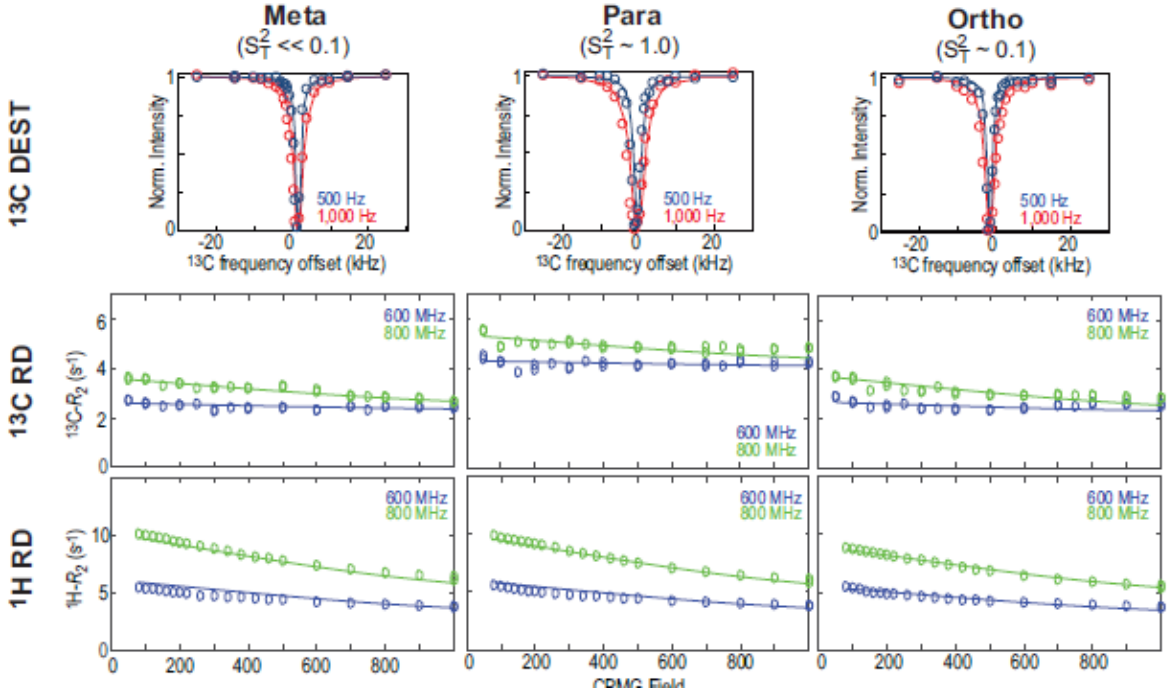


**Figure 5.7S.** $^{13}$C DEST (top panel), $^{13}$C RD (middle panel), and $^{1}$H RD (bottom panel) profiles measured for the meta (left), para (center), and ortho (right) position of PhOH in the presence of 1 wt% $CeO_2$. Experimental data are shown as blue and red circles for the DEST experiments measured at saturation fields of 500 and 1,000 Hz, respectively. For the RD profiles, experimental R2 rates acquired at 800 and 600 MHz are shown as green and blue circles, respectively. Solid curves show best fit to a three-site exchange model in which the tightly adsorbed state is treated using Model Free theory and the meta, para, and ortho positions are allowed to adopt different order parameters. Best fit order parameters are shown.

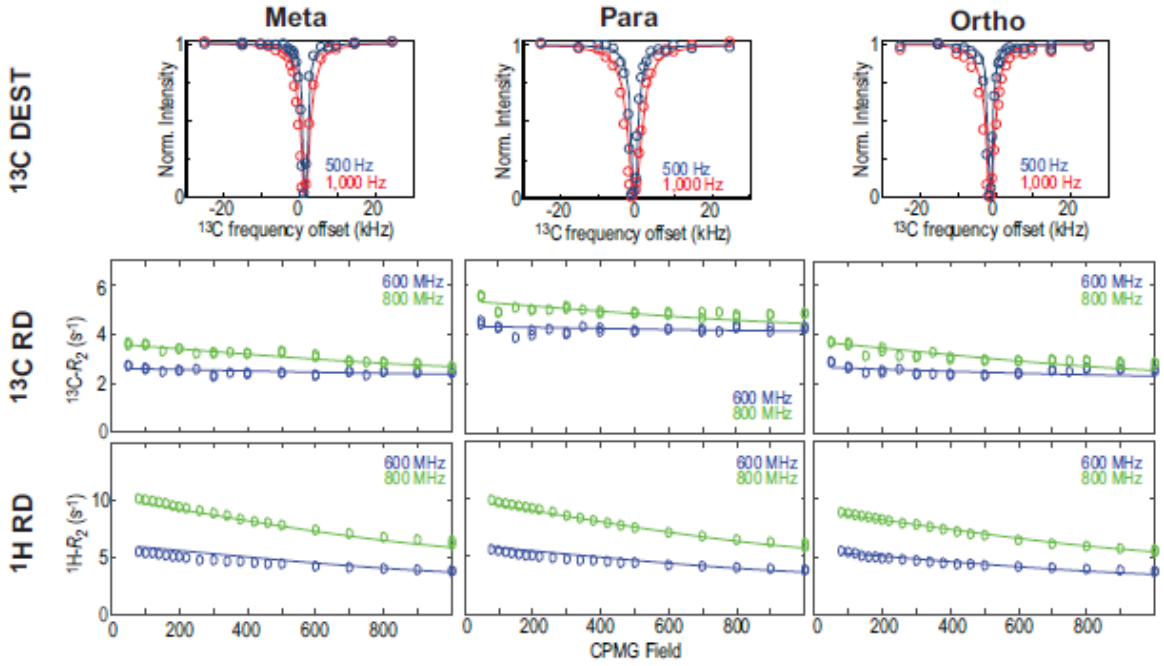


**Figure 5.8S.** $^{13}$C DEST (top panel), $^{13}$C RD (middle panel), and $^{1}$H RD (bottom panel) profiles measured for the meta (left), para (center), and ortho (right) position of PhOH in the presence of 1 wt% $CeO_2$. Experimental data are shown as blue and red circles for the DEST experiments measured at saturation fields of 500 and 1,000 Hz, respectively. For the RD profiles, experimental $R_2$ rates acquired at 800 and 600 MHz are shown as green and blue circles, respectively. Solid curves show best fit to a three-site exchange model in which the tightly adsorbed state is treated using Restricted Rotation theory *(6).*

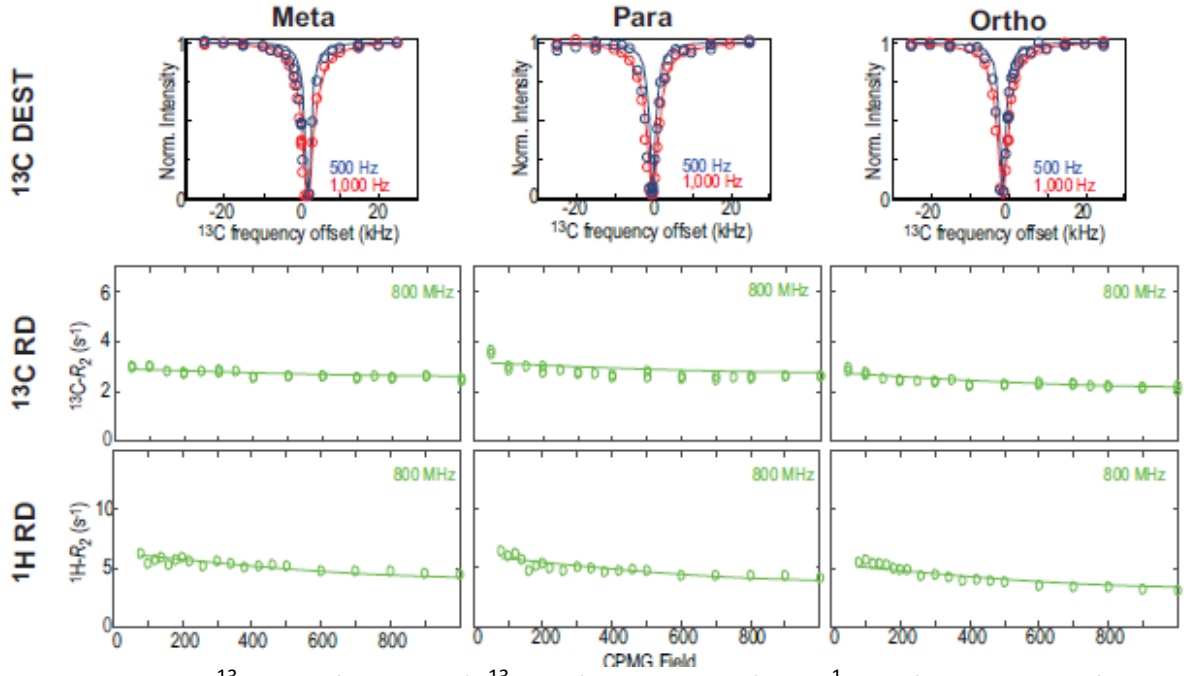


**Figure 5.9S.** $^{13}C$ DEST (top panel), $^{13}C$ RD (middle panel), and $^{1}H$ RD (bottom panel) profiles measured for the meta (left), para (center), and ortho (right) position of PhOH in the presence of 1 wt% Pd/CeO2. Experimental data are shown as blue and red circles for the DEST experiments measured at saturation fields of 500 and 1,000 Hz, respectively. For the RD profiles, experimental R2 rates are shown as green circles. Solid curves show best fit to a three-site exchange model in which the tightly adsorbed state is treated using Model Free theory and a global order parameter. Samples containing $Pd/CeO_2$ are less stable than samples containing $CeO_2$ and slowly degrade over time. To reduce the overall acquisition time and avoid to introduce systematic errors in the analysis, NMR data were acquired at 800 MHz only. The $\Delta\omega$ parameters describing the change in $^{1}H$ and $^{13}C$ chemical shift between desorbed and weakly adsorbed states were set to the values reported in Table S1 (i.e. we assume that dispersion of Pd on ceria does not affect the chemical shift of the weakly adsorbed state). This assumption reduces the number of fittable parameters and facilitates convergence of the fitting procedure.

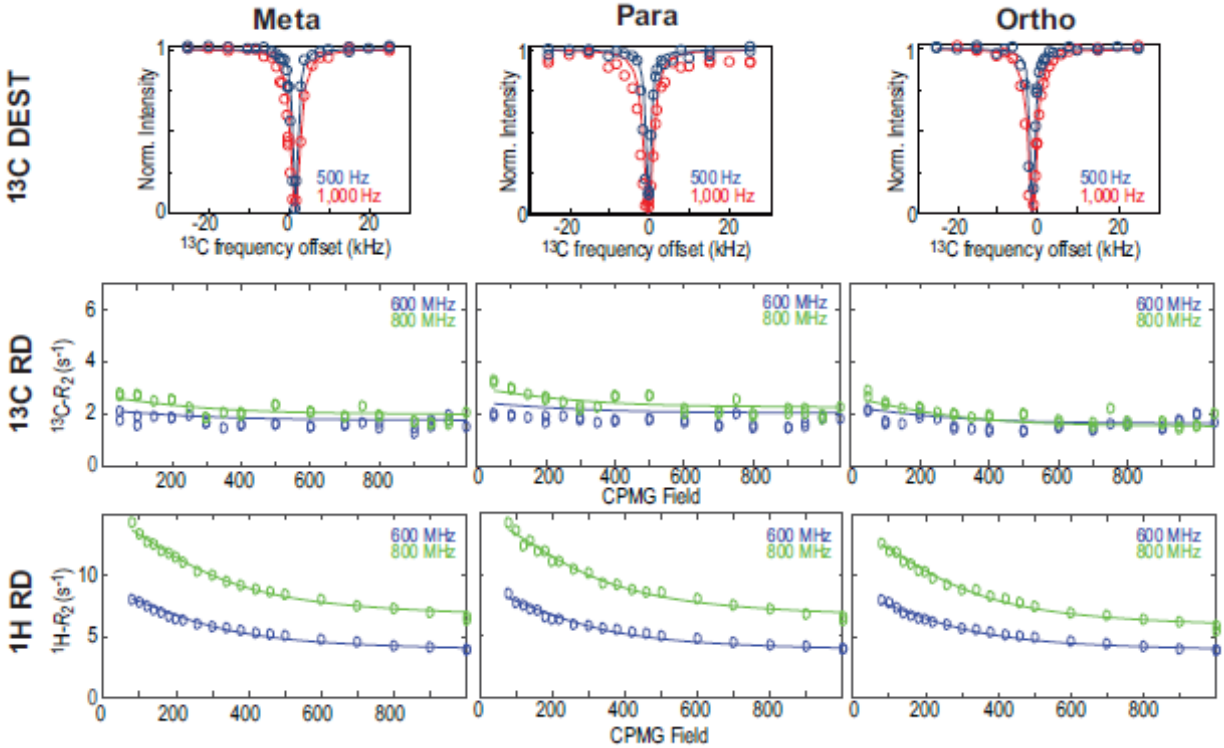


**Figure 5.10S.** $^{13}C$ DEST (top panel), $^{13}C$ RD (middle panel), and $^{1}H$ RD (bottom panel) profiles measured for the meta (left), para (center), and ortho (right) position of PhOH in the presence of 1 wt% $CeO_2$ and 20 mM phosphate. Experimental data are shown as blue and red circles for the DEST experiments measured at saturation fields of 500 and 1,000 Hz, respectively. For the RD profiles, experimental $R_2$ rates acquired at 800 and 600 MHz are shown as green and blue circles, respectively. Solid curves show best fit to a three-site exchange model in which the tightly adsorbed state is treated using Restricted Rotation theory. Note that the population of tightly adsorbed state ($p_T$) is < 0.01 %, and the fitted parameters for this state are ill defined.

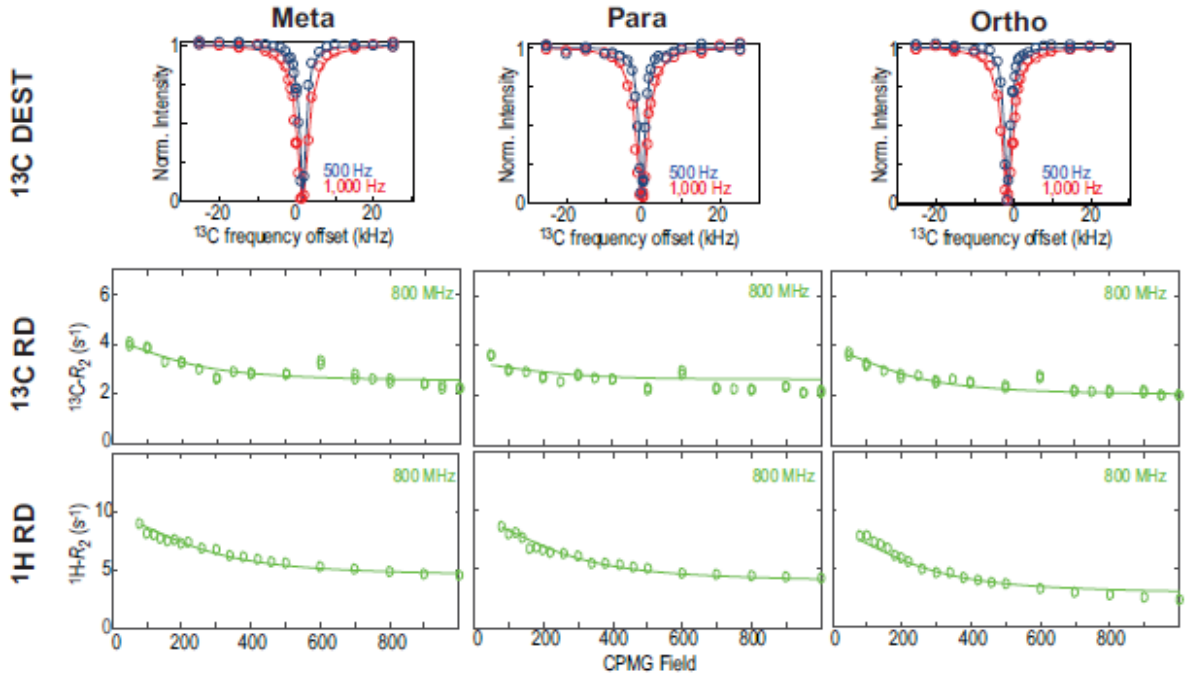


**Figure 5.11S.** $^{13}C$ DEST (top panel), $^{13}C$ RD (middle panel), and $^{1}H$ RD (bottom panel) profiles measured for the meta (left), para (center), and ortho (right) position of PhOH in the presence of 1 wt% $Pd/CeO_2$ and 20 mM phosphate. Experimental data are shown as blue and red circles for the DEST experiments measured at saturation fields of 500 and 1,000 Hz, respectively. For the RD profiles, experimental $R_2$ rates are shown as green circles. Solid curves show best fit to a three-site exchange model in which the tightly adsorbed state is treated using Model Free theory and a global order parameter. Samples containing $Pd/CeO_2$ are less stable than samples containing $CeO_2$ and slowly degrade over time. To reduce the overall acquisition time and avoid to introduce systematic errors in the analysis, NMR data were acquired at 800 MHz only. The $\Delta\omega$ parameters describing the change in $^{1}H$ and $^{13}C$ chemical shift between desorbed and weakly adsorbed states were set to the values reported in Table S1 (i.e. we assume that dispersion of Pd on ceria does not affect the chemical shift of the weakly adsorbed state). This assumption reduces the number of fittable parameters and facilitates convergence of the fitting procedure.

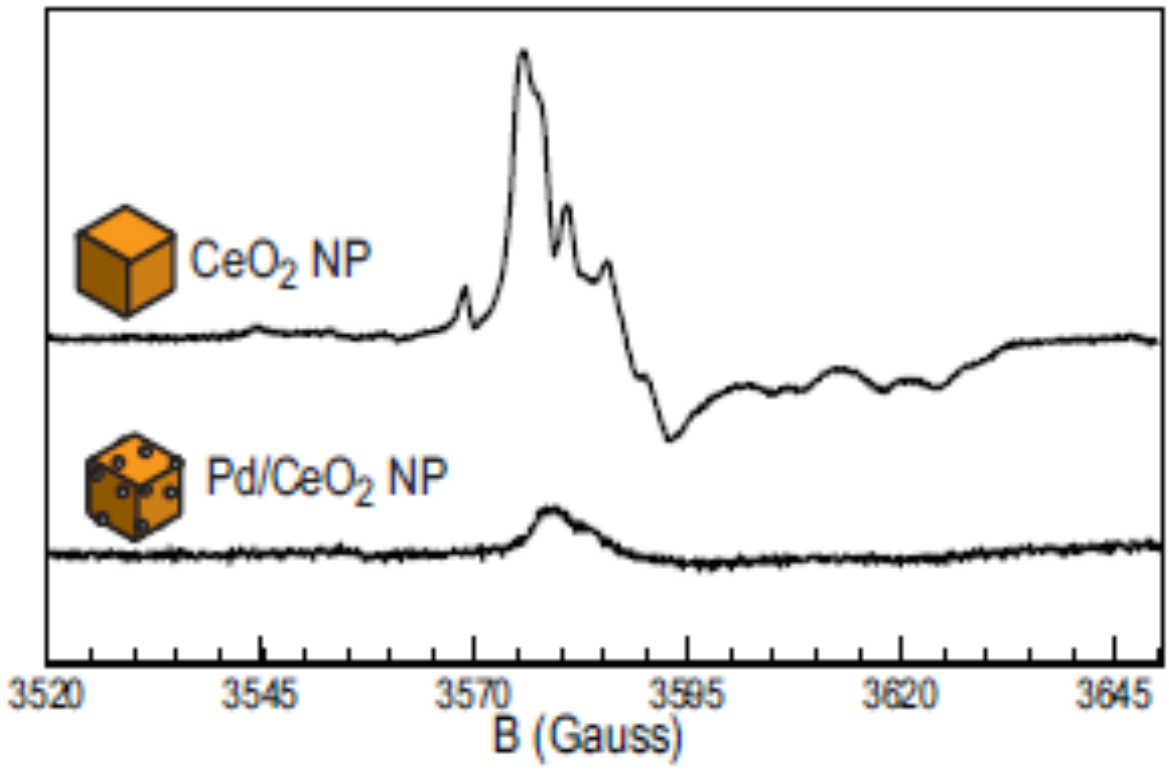


**Figure 5.12S.** X-band EPR spectra measured on powder samples of $CeO_2$ (top) and $Pd/CeO_2$ (bottom) nanoparticles. The signal is proportional to the amount of $Ce^{3+}$, which is directly proportional to the number of oxygen vacancies in the material (each oxygen vacancy is surrounded by two $Ce^{3+}$ atoms) (11). Dispersion of Pd on the ceria support quenches the intensity of the EPR signal by one order of magnitude, which suggests a drastic reduction in oxygen vacancies upon impregnation of the support with Pd.

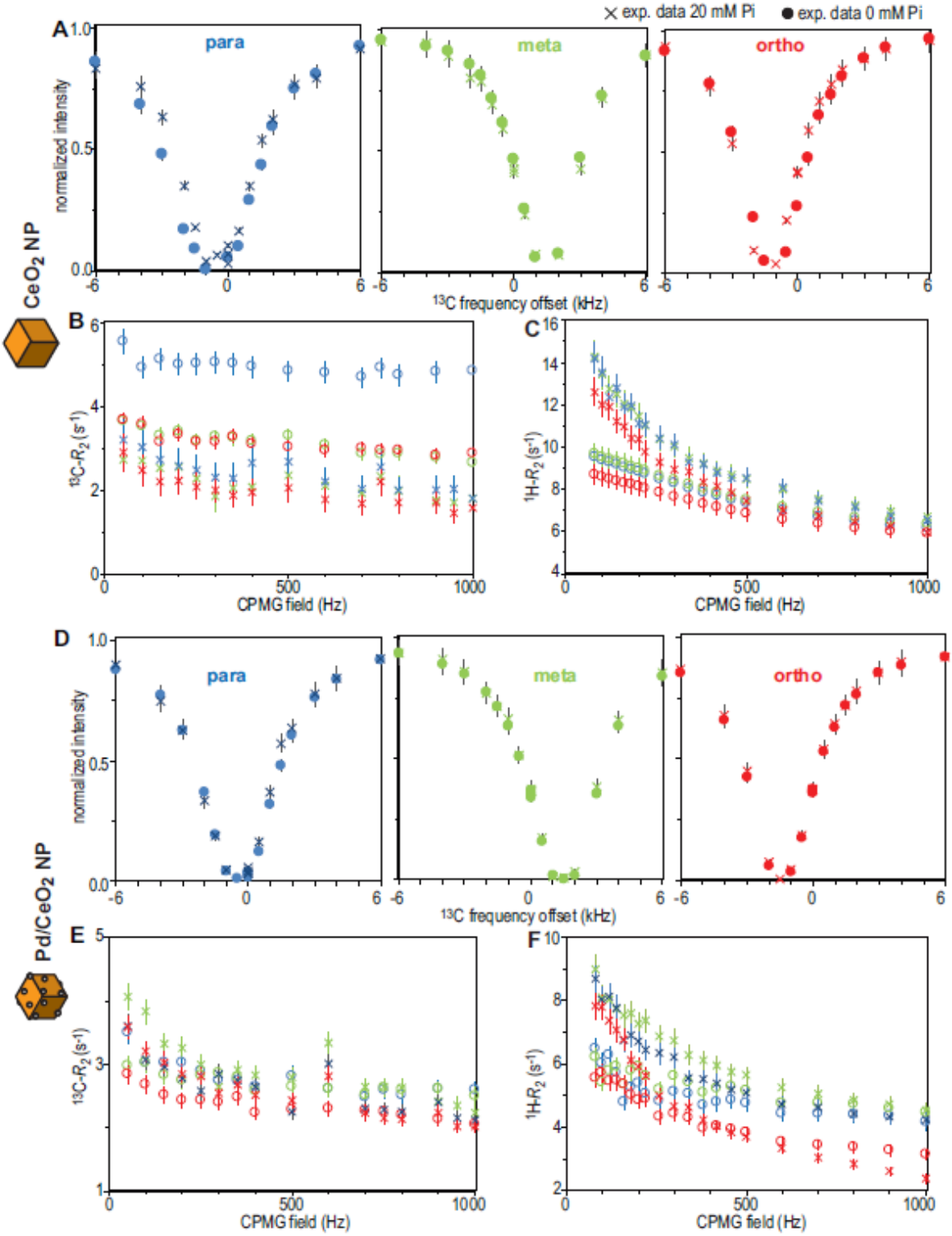


**Figure 5.13S.** Example of $^{13}C$ DEST profiles (saturation field = 1 kHz) measured for PhOH in the presence of 1 wt% (A) $CeO_2$ and (D) $Pd/CeO_2$. Example of $^{13}C$ RD profiles measured for PhOH in the presence of 1 wt% (B) $CeO_2$ and (E) $Pd/CeO_2$. Example of $^{1}H$ RD profiles measured

for PhOH in the presence of 1 wt% (C) $CeO_2$ and (F) $Pd/CeO_2$. All data were acquired on an 800 MHz spectrometer. Crosses and circles indicate data measured in the presence of 0 mM and 20 mM sodium phosphate, respectively. Data for the para, meta, and ortho positions are colored blue, green, and red, respectively.

These plots show that addition of phosphate affects the overall shape of the RD profiles (which report on µs-ms exchange processes), indicating that weak adsorption of PhOH to $CeO_2$ and $Pd/CeO_2$ is perturbed by the inorganic anion. Also, addition of phosphate results in an overall reduction of the $^{13}C$ $R_2$ rate and in a narrowing of the $^{13}C$ DEST profile measured for the para position in the presence of $CeO_2$, which reflects the destabilization of the tightly bound state observed upon addition of phosphate (Figure 2 in main text). Finally, it is interesting to note that the DEST profiles acquired in the presence of $Pd/CeO_2$ are unaffected by phosphate, suggesting that tight association between PhOH and $Pd/CeO_2$ is unperturbed by the inorganic ion.

## Supplementary Tables

**Table 5.1S.** Kinetic, population, chemical shift, rotational correlation time, and order parameters for PhOH binding to $CeO_2$ nanoparticles derived from global fitting of $^{13}C$ DEST, $^{13}C$ RD, $^{13}C$ R1, and $^{1}H$ RD data using a three-site exchange model. Definitions for the global parameters are given in Figure 2. 13C and $^{1}H$ $\Delta\omega^{DW}$ are the absolute values of the 13C and 1H chemical shift difference in ppm between desorbed and weakly adsorbed state, respectively. Note that $^{13}C$ and $^{1}H$ chemical shift difference between weakly and tightly adsorbed states were fixed to 0 in the fit because the RD experiments are insensitive to the chemical shift of the tightly adsorbed state (which is in slow exchange with the weakly adsorbed state), and the DEST is a low-resolution experiment that does not provide chemical shift information *(8)*.

| **Global Parameters** | **0 mM phosphate** (3-site connected) | **0 mM phosphate** (3-site linear) | **20 mM phosphate** (3-site linear) |
|---|---|---|---|
| $p_D$ (%) | 95.0 ± 0.5 | 94.7 ± 0.5 | 97.5 ± 1.0 |
| $p_W$ (%) | 4.9 ± 0.5 | 5.2 ± 0.5 | 2.5 ± 1.0 |
| $p_T$ (%) | 0.06 ± 0.01[a] | 0.06 ± 0.01[a] | < 0.01 |
| $\tau_C^{loc,W}$ (ps) | < 100 | < 100 | < 100 |
| $\tau_C^{loc,T}$ (ps) | < 100 | < 100 | < 100 |
| $S_W^2$ | < 0.01 | < 0.01 | < 0.01 |
| $S_T^2$ [b] | - | - | - |
| $k_{WD}$ ($s^{-1}$) | 3,753 ± 166 | 3,613 ± 115 | 1,706 ± 30 |
| $k_{DW}^{app}$ ($s^{-1}$) | 194 ± 10 | 198 ± 14 | 59 ± 16 |
| $k_{TW}$ ($s^{-1}$) | 1,830 ± 583 | 2,358 ± 365 | - |
| $k_{WT}^{app}$ ($s^{-1}$) | 20 ± 6 | 21 ± 5 | - |
| $k_{DT}^{app}$ ($s^{-1}$) | 0 ± 1 | - | - |
| $k_{TD}$ ($s^{-1}$) | 0 ± 1 | - | - |
| **Peak-specific parameter** | | | |
| $^{13}C$ $\Delta\omega^{DW}$ (ppm) o/m/p[c] | 0.19 / 0.13 / 0.18 | 0.19 / 0.13 / 0.17 | 0.15 / 0.11 / 0.10 |
| $^{1}H$ $\Delta\omega^{DW}$ (ppm) o/m/p[c] | 0.09 / 0.09 / 0.10 | 0.09 / 0.09 / 0.09 | 0.11 / 0.11 / 0.11 |

[a] note that $p^T$=0.06 % corresponds to 6 µM PhOH

[b] data were fit using a Restricted Rotation model for the tightly adsorbed state

[c] Error on chemical shift parameters is < 0.01 ppm in all cases

**Table 5.2S**. Kinetic, population, chemical shift, rotational correlation time, and order parameters for PhOH binding to Pd/CeO2 nanoparticles derived from global fitting of $^{13}$C DEST, $^{13}$C RD, $^{13}$C R1, and $^{1}$H RD data using a three-site exchange model. Definitions for the global parameters are given in Figure 2. 13C and $^{1}$H $\Delta\omega^{DW}$ are the absolute values of the $^{13}$C and $^{1}$H chemical shift difference in ppm between desorbed and weakly adsorbed state, respectively. Note that $^{13}$C and $^{1}$H chemical shift difference between weakly and tightly adsorbed states were fixed to 0 in the fit because the RD experiments are insensitive to the chemical shift of the tightly adsorbed state (which is in slow exchange with the weakly adsorbed state), and the DEST is a low-resolution experiment that does not provide chemical shift information (8).

| **Global Parameters** | **0 mM phosphate** (3-site connected) | **0 mM phosphate** (3-site linear) | **20 mM phosphate** (3-site linear) |
|---|---|---|---|
| $p_D$ (%) | 97.7 ± 0.2 | 97.6 ± 0.2 | 98.3 ± 0.1 |
| $p_W$ (%) | 2.2 ± 0.2 | 2.3 ± 0.2 | 1.6 ± 0.1 |
| $p_T$ (%) | 0.08 ± 0.01[a] | 0.09 ± 0.01[a] | 0.10 ± 0.01[a] |
| $\tau_C^{loc,W}$ (ps) | < 100 | < 100 | < 100 |
| $\tau_C^{loc,T}$ (ps) | < 100 | < 100 | < 100 |
| $S_W^2$ | < 0.01 | < 0.01 | < 0.01 |
| $S_T^2$ | 0.9 ± 0.1 | 0.9 ± 0.1 | 0.8 ± 0.1 |
| $k_{WD}$ (s$^{-1}$) | 2,768 ± 210 | 2,878 ± 235 | 1,364 ± 54 |
| $k_{DW}^{app}$ (s$^{-1}$) | 63 ± 10 | 62 ± 12 | 22 ± 1 |
| $k_{TW}$ (s$^{-1}$) | 271 ± 21 | 273 ± 23 | 290 ± 24 |
| $k_{WT}^{app}$ (s$^{-1}$) | 9 ± 2 | 11 ± 5 | 15 ± 5 |
| $k_{DT}^{app}$ (s$^{-1}$) | 0 ± 1 | - | - |
| $k_{TD}$ (s$^{-1}$) | 0 ± 1 | - | - |
| **Peak-specific parameter** | | | |
| $^{13}$C $\Delta\omega^{DW}$ (ppm) o/m/p[b] | 0.19 / 0.13 / 0.17 | 0.19 / 0.13 / 0.17 | 0.15 / 0.11 / 0.10 |
| $^{1}$H $\Delta\omega^{DW}$ (ppm) o/m/p[b] | 0.09 / 0.09 / 0.09 | 0.09 / 0.09 / 0.09 | 0.11 / 0.11 / 0.11 |

[a] note that pT = 0.10 % corresponds to 10 µM PhOH

[b] Chemical shift parameters are set to the values reported in **Supplementary Table 5.S1** – see **Supplementary Figures 5.S9** and **5.S11**

**References**

1. H.-X. Mai et al., J Phys Chem B 109, 24380-24385 (2005).

2. J. A. Aguilar, M. Nilsson, G. Bodenhausen, G. A. Morris, Chem Commun (Camb) 48, 811-813 (2012).

3. T. K. Egner, P. Naik, N. C. Nelson, Slowing, II, V. Venditti, Angew Chem Int Ed Engl 56, 9802-9806 (2017).

4. K. Pervushin, R. Riek, G. Wider, K. Wuthrich, J Am Chem Soc 120, 6394-6400 (1998).

5. W. S. Veeman, Prog NMR Spectrosc 16, 193-235 (1984).

6. J. R. Sachleben, V. Colvin, L. Emsley, E. W. Wooten, A. P. Alivisatos, J Phys Chem B

102, 10117-10128 (1998).

7. A. Ceccon, V. Tugarinov, A. Bax, G. M. Clore, J Am Chem Soc 138, 5789-5792 (2016).

8. N. J. Anthis, G. M. Clore, Q Rev Biophys 48, 35-116 (2015).

9. N. C. Nelson, J. S. Manzano, A. D. Sadow, S. H. Overbury, I. I. Slowing, ACS Catal 5, 2051-2061 (2015).

10. Y. Lei et al., Fuel 233, 10-20 (2018).

11. F. Esch et al., Science 309, 752-755 (2005).

# CHAPTER 6

## DEVELOPMENT AND UTILIZATION OF ORGANIC GELS FOR ANALYSIS OF NANOPARTICLE SYSTEMS BY NMR

Published in Chemical Reviews in 2016

Timothy Egner[1], Vincenzo Venditti[1,2]

[1]Department of Chemistry, Iowa State University
[2]Roy J. Carver Department of Biochemistry, Biophysics, and Molecular Biology, Iowa State University

### Introduction

In chapter 2, I describe a nanoparticle system wherein we utilized DEST and RD experiments on nanoparticles suspended in a gel to determine kinetics and thermodynamic properties of the interaction between small molecules and ceria nanoparticles. These interactions were able to be studied because methods were developed to suspend nanoparticles within an agarose gel[1]. Nanoparticles in suspension tend to settle too quickly for accurate NMR measurements to be obtained, however, a gel can be used to overcome this issue. Although the agarose system is useful, the systems that can be utilized are limited because agarose will only gel with water and not organic solvents. Even small amounts of organic solvents can perturb the structural integrity and prevent the gel from forming. In order to overcome this issue, gels that are compatible with organic solvents need to be developed. The two major types of gels that could be utilized to study nanoparticle catalytic systems are derived from covalent or non-covalent gels [2].

Non-covalent gels typically consist of a molecule that has heterogeneous patches of hydrophobic and hydrophilic moieties. If the solvent is hydrophilic, the hydrophobic patches can

aggregate through intermolecular forces (IMF), while the hydrophilic patches interact with the solvent, resulting in a mesh-like structure with areas that exclude solvent molecules [3]. Likewise, if the solvent is hydrophobic, the hydrophilic patches on the gel molecules can aggregate together through IMF while the hydrophilic patches interact with the solvent. Due to the nature of these gels, it may be possible to flip the structure of the gel based on the hydrophobicity or hydrophilicity of the solvent used. Non-covalent gels tend to be sensitive to solvent conditions, as they will only form a gel with the use of a single solvent. Use of a single solvent makes the applications of the gel limited; an organic gel compatible with many organic solvents would expand its utility. Examples of non-covalent gels include agarose, methylcellulose, and glucopyranosides.

Covalent gels usually consist of a large polymer with crosslinked chains within the structure [4]. Very commonly, the polymerization of these gels consists of a free radical process linking consecutive alkenes together to form an alkane based backbone. Covalent gels required the use of crosslinking agents to hold individual strands together. For example, in free radical polymerization, a crosslinking agent would have two or more alkene moieties in its structure to allow two or more polymer chains to be joined. Due to the fact that these gels are held together by covalent bonds, they are significantly less sensitive to variations in the solvent and could thus be used with a wider variety of solvents for a single gel material. Common examples of covalent gels are polystyrene, ABS, acrylates, and polyethylene.

In the process of developing gels for NMR studies, it is important to determine whether a gel would be useful in its application to suspend nanomaterial and allow for acquisition of NMR

data. I have formulated six qualitative criteria to aid in the process of determining the aptitude of a gel system:

(1) Low residual NMR signal: Gels that have minimal NMR signal that are far away from ligand and nanoparticle signals are ideal.

(2) Low internal viscosity: The gel structure should be relatively rigid and should not significantly affect the bulk solvent conditions within the pores.

(3) Large pore size: The gel pore diameters should be large relative to the nanomaterial of interest as to not restrict solution tumbling of the nanoparticles.

(4) Macroscopic integrity: The gel should be easy to handle in a dry and swollen state, and should be able to withstand gentle agitation.

(5) Solvent compatibility: The gel can be swollen with a wide variety of conditions.

(6) Simple preparatory conditions: The gel needs to be able to be prepared quickly, easily, and be relatively inexpensive.

Each of these criteria are important to the overall quality of the gel, and while failure in one of them may still result in a usable gel, failure in three or more suggests that the gel is not optimal for NMR experiments. In this chapter, I will give an overview of six different gels that could be suitable for use in NMR studies of nanomaterial and their characteristics for each of these six criteria (**Table 6.1**).

**Methods and Results**

*1. Agarose*

Agarose has previously been described as a useful gel for NMR nanoparticle studies. Agarose gels are simple to make, as they require brief heating of a mixture of solid agarose and

water to 80˚C to solubilize [5]. Additionally, the gel is strong on its own at 1 wt% and can be handled easily. Formation of the gel can be done within an NMR tube and it only requires cooling to below 35˚C which can take less than 15 minutes at room temperature. The NMR spectrum of an agarose gel is quite sparse with only methoxy groups detected at around 3-4 ppm. All other areas of the spectra are clear and available for analysis of other signals. Previous studies show that at 1 wt%, agarose gels have a pore size of 100-500 nm [6], which is large enough to allow free isotropic tumbling of nanoparticles (>100) within the pores. Additionally, because the pore sizes are large relative to the nanoparticles, there is relatively little surface area for ligands to interact with the gel. For example, in a spectrum of agarose and phenol compared to phenol alone, the line widths for phenol are nearly identical, indicating that there are no significant interactions between the phenol and agarose gel (**Figure 6.1**). Agarose has mostly ideal characteristics, however it is only compatible with water as a solvent. Identification and characterization of gels that are compatible with organic solvents can expand the versatility of gel systems for NMR analysis.

*2. Polystyrene*

Polystyrene (PS) is one type of organic gel that could be used for NMR studies. This gel is compatible with a wide variety of organic solvents including aromatics, aliphatics, and chloromethanes [7]. Polystyrene gels are made by heating styrene monomers with an initiator for approximately 1 hr [8]. Since this is a relatively long time period, using this gel can be problematic if the nanomaterial of interest does not stay suspended during this time, and will make it difficult to attain a homogenous distribution of nanomaterial throughout the gel. Additionally, the PS gel must be washed to remove unreacted monomers. This process can take

up to two days to achieve by washing with 20 mL of toluene or benzene for every 1 mL of gel with fresh solvent introduced every 6-12 hours. The gel then needs to be dried for 2 days at room temperature and then reswollen with a deuterated solvent in an NMR tube for 1-2 weeks. The long swelling time is required for homogenous distribution of the solvent within the gel. If the gel is not swollen for a long enough time, major issues arise with shimming the sample in the NMR magnet during data acquisition. In testing this gel, I have found that polymerizing within a 3 mm inner diameter glass tube works well to achieve the proper dimensions; however stirring is not possible in this case. The 3mm diameter stick will expand to 5mm when swollen with $CDCl_3$. The NMR spectrum of a PS gel is shown in **Figure 6.2**. Large signals originating from the gel can be seen in the aromatic and aliphatic region. Significant signal overlap between the gel and ligand of interest is expected, which restricts the versatility of this gel. This could be overcome using deuterated styrene but would make production much more expensive.

*3. Polydimethylsilicone*

Polydimethylsilicone (PDMS) has previously been used in NMR studies as an alignment media for measuring residual dipolar couplings [9]. PDMS consists of long chains of -O-Si$(CH_3)_2$- and thus the only 1H signal for this gel appears at approximately 0.1ppm, which is out of the way of most other signals that would be in associated with a NP/ligand system. There are many advantages to using a PDMS gel for NP studies. There are a few previously described processes for preparing PDMS gels, including one described by Zhang et al (2010), whereby the gel is made by heating TEOS, hydroxyl-terminated PDMS, and ammonia in ethanol [10]. These gels take approximately 1 hour to form, with 24 hour periods for washing, drying, and gel swelling, for a total of >72 hours of total preparatory time (swelling time may need to be extended to prevent

inhomogeneity issues). Although shorter preparatory times would be ideal, it does not prevent the use of this gel for NP studies. The NMR spectrum of PDMS swollen with $CDCl_3$ is shown in **Figure 6.3**. The only signal from this gel appears at 0.1 ppm and is quite broad. When phenol is added to the gel, the signals are broad, indicating that the internal viscosity or ligand-gel interations is high. High internal viscosity poses an issue because there may not be enough relaxation contract between free and NP states to be usable. This gel may be useful for higher temperature applications as this would decrease viscosity.

*4. Glucopyranoside ((4,6-O-Benzylidene)-Methyl-α-D-Glucopyranoside)*

Glucopyranoside (GP) is another material that could be utilized as an organic gel. GP is considered a 1-dimension gelator which forms aggregates with itself through hydrogen bonding. Organogelators such as GP are thought to be useful for jellifying oil spills to aid in the clean up processes [11]. GP gels can be made by adding 1 wt% GP to toluene and mixed with gentle heating (35°C), and upon cooling the gel will form. GP gels form quickly and easily, although the gelling process works predominantly with aromatic solvents such as benzene, toluene, and xylenes. If aliphatic solvents are used, the solution will not gel. Deuterated aromatic solvents can be costly and limit its practical applicability. These gels can form within an NMR tube, however there are issues with its macroscopic integrity. If the sample is agitated or inverted, the gel will break into smaller pieces. This is likely due to the fact that the gel is not cross linked, with the only long range structural component consisting of 1-D aggregates [12].

To test this gel, an NMR sample was prepared using 60:40 toluene: CDCl3 and 1wt% GP, with the gel directly formed within the NMR tube. **Figure 6.3** shows that large signals from toluene (not deuterated) are present at 1.6 and 6.6 ppm. Low intensity signals from GP are

observed but are quite sharp. These residual sharp signals may be GP monomers that are not aggregate. Due to limited solvent compatibility (of which are more expensive) and poor stability, this material seems to have limited application for our purposes of ligand/NP NMR studies.

*5. Methylcellulose*

Methylcellulose (MeC) is a heterogeneously methylated carbohydrate that is similar to agarose. Patches of hydrophobicity from the methylated regions can aggregate in an aqueous solvent to form a gel. Similarly, the unmethylated hydrophilic regions can aggregate to form a gel in a hydrophobic solvent. MeC appears to have limited solvent options for gelling with DMF being the only reported organic solvent [13]. To make this gel, 1.5 wt% MeC and DMF were added to a tube and heated to 80°C. The DMF immediately begins to "hydrate" through the MeC powder, and with moderate stirring and agitation with a stir rod the mixture will become a viscous fluid within 10 minutes. Once the mixture is viscous, the ligands of interest and nanomaterial can be added to the preparation. The suspension is quickly transferred to an NMR tube before cooling to room temperature, which typically takes less than 20 minutes. **Figure 5.5** shows an NMR spectrum of this gel. There are significant signals from the DMF-$d_7$ present in this spectrum, but they are sharp and therefore easy to distinguish from gel and ligand signals. Phenol signals can be observed around 7.25-6.8 ppm in the spectrum, and unlike in water, the hydroxyl proton is visible at 9.55 ppm. Upon addition of ceria nanoparticles, the aromatic signals become broadened and the hydroxyl signal is almost completely broadened into the baseline. Spectral data of nanoparticles within MeC gels lends credence to the hypotheses that phenol hydrogen bonds with the ceria surface (**Figure 5.5 Inset**). Since the motion of the hydroxyl proton is very restricted in this binding mode, the signal should be disproportionaly broaden as compared to

the aromatic proton signals. Overall this gel is fast and easy to prepare, has good macroscopic integrity, low internal viscosity, and minimal residual signal. However, the major drawback with MeC gels is that it is only compatible with DMF.

*6. Polystearylacrylate*

Polystearylacrylate (PSA) is similar in structure to polyacrylamide, but it consists of an $C_{18}$ chain instead of a hydrophilic amide group. PSA is very hydrophobic and is therefore a good candidate for non-polar organic solvent compatibility [14]. PSA can be crosslinked with ethylene glycol dimethylacrylate (EGDMA), and because of this crosslinking, it has good macroscopic integrity. PSA gels have been formed by heating a mixture of SA, EGDMA, TEMED, and AIBN (1:0.01:0.001:0.001 wt%) at 80°C for approximately 5 minutes (based on methods from Darvishi et al (2013) [15]. Nanomaterial can be added to the reaction mixture once all the stearylacrylate has melted.

One difficulty that arises with PSA gels is that in the process of preparation for NMR studies, the gel must be cast in cylinders that when swollen are the same diameter as an NMR tube. To achieve this, 2 mm inner diameter glass tubes with open ends were placed directly into the reaction mixture, and any liquid that flows into the glass tube will form a gel with the proper dimensions. When the gel has cooled and is solid, the glass rods can be broken out of the plastic mass and the material inside the tubes can be gently pushed out. The gels are then washed with toluene, benzene, or $CDCl_3$ for two days, with solvent changes every 6-12 hours. After washing, the gels are dried at room temperature for two days on aluminum foil with care taken to prevent the gels from becoming stuck to the foil and cracking. This procedure could be optimized by utilizing an oven, vacuum, or both to decrease the total drying time. Once dried, the PSA stick

can be placed in an NMR tube and is then swollen with a solvent and ligand solution of choice. The swelling process takes approximately one week to achieve uniform swelling. If the swell time is too short, issues will arise when attempting to shim the sample.

A PSA gel was swollen with $CDCl_3$ with or without nanoparticles. **Figure 6.6** shows the PSA swollen with $CDCl_3$ and shows that there is significant signal from the gel around 2-3 ppm from the $C_{18}$ side chain, with little signal elsewhere. When there are no nanoparticles within the sample, the phenol signals are quite sharp, comparable to linewidths of phenol in $CDCl_3$. Upon the addition of ceria, the phenol signals become broadened; only 0.25 wt% ceria was used here, but broadening is comparable to 1wt% ceria in water (See Chapter 4). This finding suggests that the phenol-ceria interaction is more populated (higher affinity) in a non-polar solvent. Indeed, the hypothesized binding mode is hydrogen bonding to the ceria surface (see Chapters 4 and 5). In aqueous solvent, the water molecules can competitively bind to surface exposed oxygens, but in this system those sites are free as $CDCl_3$ will not hydrogen bond, allowing more phenol to bind. Of all gels systems reported on in this chapter, PSA seems to be the most likely candidate gel for application ligand/NP NMR studies using a variety of solvents.

**Conclusion**

Gels have been shown to be useful for trapping NPs in homogeneous distribution for the purpose of NMR investigations [1]. However, limited solvent compatibility means less overall NP systems can be experimented on in their native solvent. Using six basic criteria for qualifying gels, namely: low residual NMR signal, low internal viscosity, large pore size, macroscopic integrity, solvent compatibility, and simple preparatory conditions, I have tested and reported on six of different gels and their properties. Overall, PSA appears to be the most suitable for general

purpose with various organic solvents for NMR studies. There are other possibilities depending on the specific criteria required for a given NP system.

**References**


1. Egner TK, Naik P, Nelson NC, Slowing II, Venditti V. Mechanistic Insight into Nanoparticle Surface Adsorption by Solution NMR Spectroscopy in an Aqueous Gel. Angew Chem Int Ed Engl. 2017;56: 9802–9806. doi:10.1002/anie.201704471

2. Picchioni F, Muljana H. Hydrogels Based on Dynamic Covalent and Non Covalent Bonds: A Chemistry Perspective. Gels (Basel, Switzerland). 2018;4: 21. doi:10.3390/gels4010021

3. Singh RP, Kundu PP. Thermogelation of methylcellulose from solution in N,N-dimethylformamide and characterization of the transparent gels. J Appl Polym Sci. 2008;110: 3031–3037. doi:10.1002/app.28527

4. Storey BT. Copolymerization of styrene and p-divinylbenzene. Initial rates and gel points. J Polym Sci Part A Gen Pap. 1965;3: 265–282. doi:10.1002/pol.1965.100030128

5. Zimm BH, Levene SD. Problems and prospects in the theory of gel electrophoresis of DNA. Q Rev Biophys. 1992;25: 171–204.

6. Viovy J-L. Electrophoresis of DNA and other polyelectrolytes: Physical mechanisms. Rev Mod Phys. 2000;72: 813–872. doi:10.1103/RevModPhys.72.813

7. Luy B, Kobzar K, Knör S, Furrer J, Heckmann D, Kessler H. Orientational Properties of Stretched Polystyrene Gels in Organic Solvents and the Suppression of Their Residual 1H NMR Signals. J Am Chem Soc. 2005;127: 6459–6465. doi:10.1021/ja043344o

8. Stover HDH, Frechet JMJ. NMR characterization of crosslinked polystyrene gels. Macromolecules. 1991;24: 883–888. doi:10.1021/ma00004a012

9. Freudenberger JC, Spiteller P, Bauer R, Kessler H, Luy B. Stretched Poly(dimethylsiloxane) Gels as NMR Alignment Media for Apolar and Weakly Polar Organic Solvents: An Ideal Tool for Measuring RDCs at Low Molecular Concentrations. J Am Chem Soc. 2004;126: 14690–14691. doi:10.1021/ja046155e

10. Zhang X, Ye H, Xiao B, Yan L, Lv H, Jiang B. Sol−Gel Preparation of PDMS/Silica Hybrid Antireflective Coatings with Controlled Thickness and Durable Antireflective Performance. J Phys Chem C. 2010;114: 19979–19983. doi:10.1021/jp106192z

11. Basu N, Chakraborty A, Ghosh R. Carbohydrate Derived Organogelators and the Corresponding Functional Gels Developed in Recent Time. Gels (Basel, Switzerland). 2018;4. doi:10.3390/gels4020052

12. Cui J, Zheng J, Qiao W, Wan X. Solvent-tuned multiple self-assembly of a new sugar-appended gelator. J Colloid Interface Sci. 2008;326: 267–274. doi:https://doi.org/10.1016/j.jcis.2008.07.008

13. Singh RP, Kundu PP. DSC and micro structural studies of methylcellulose gels in N, N dimethylformamide. J Polym Res. 2013;20: 226. doi:10.1007/s10965-013-0226-1

14. Seida Y, Suzuki T. Adsorption Properties of Stearyl Acrylate Gel for VOC Observed by QCM-A. Macromol Symp. 2017;372: 132–139. doi:10.1002/masy.201600161

15. Darvishi A, Zohuriaan Mehr MJ, Marandi GB, Kabiri K, Bouhendi H, Bakhshi H. Copolymers of glycidyl methacrylate and octadecyl acrylate: synthesis, characterization, swelling properties, and reactivity ratios. Des Monomers Polym. 2013;16: 79–88. doi:10.1080/15685551.2012.705493

**Figures**

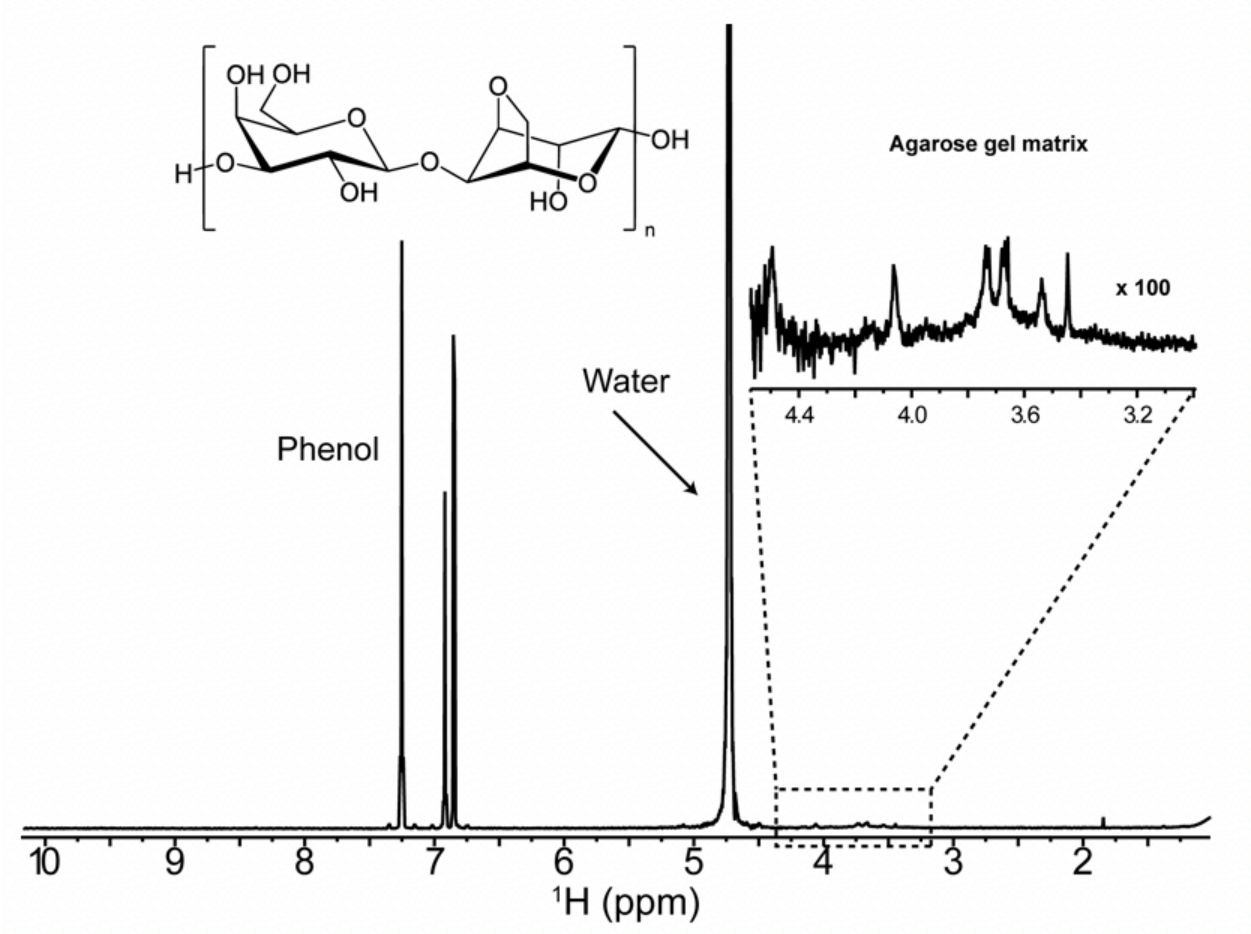


**Figure 6.1:** $^{1}$H NMR spectrum of 10mM phenol and 1wt% agarose gel in $D_2O$. The structure of agarose without methylation is shown. The only visible signals from the gel are from the methylated hydroxyl groups.

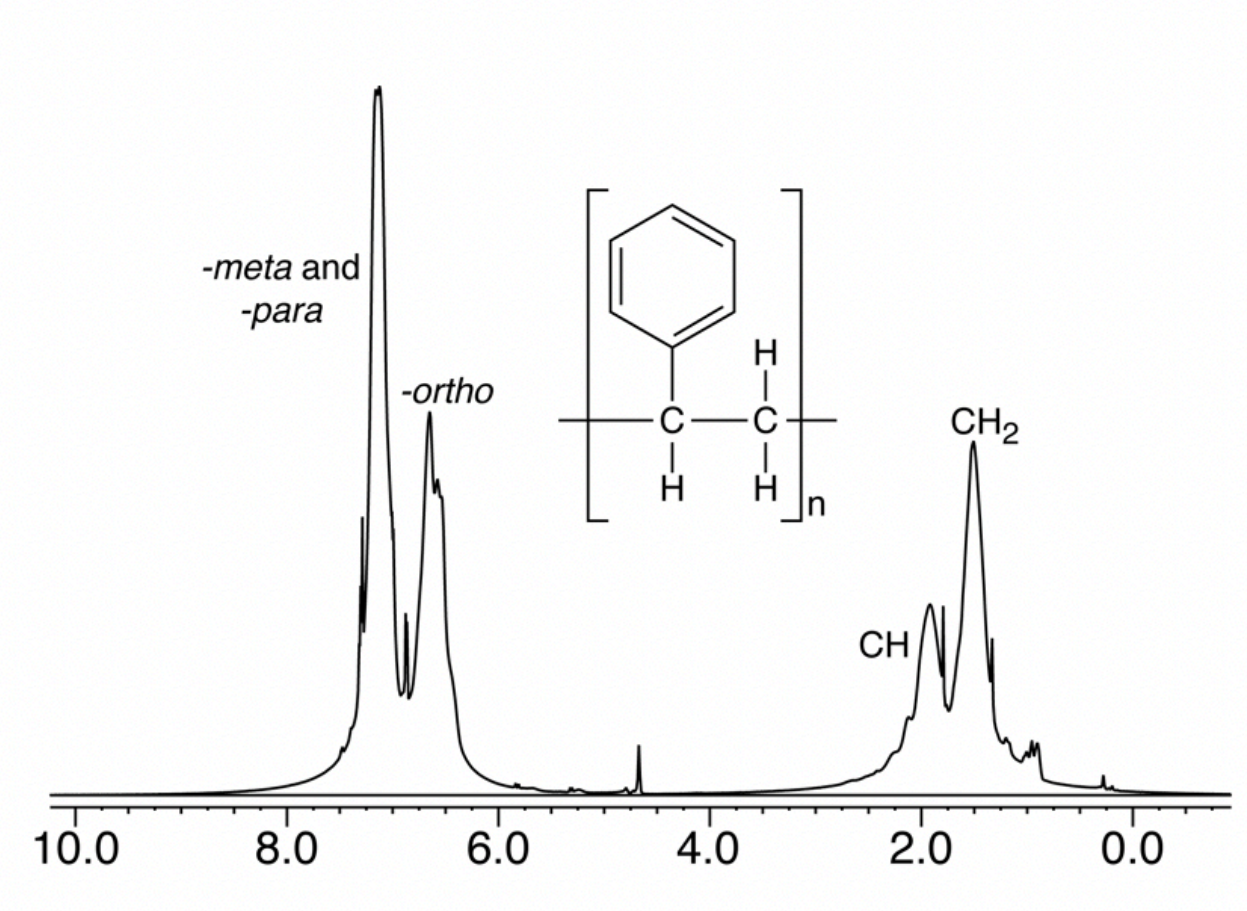


**Figure 6.2:** $^{1}$H NMR spectrum of 10mM phenol and PS in $CDCl_3$. The polymer structure of PS is shown with annotations of the spectral resonances. The concentration of PS is not known but is much greater than that of phenol. The phenol signals are not visible due to significant overlap from PS aromatic signals.

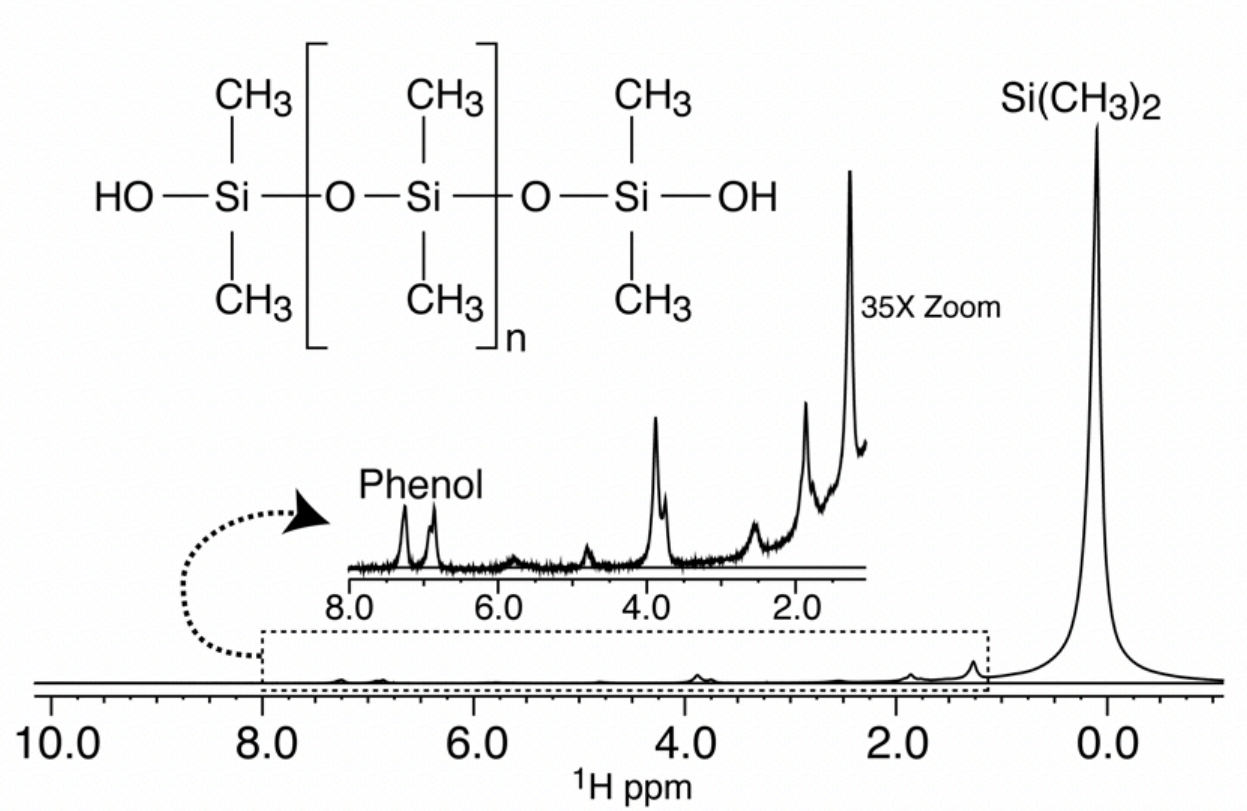


**Figure 6.3:** $^{1}$H NMR spectrum of 10mM phenol and PDMS in $CDCl_3$. The structure of PDMS is shown and the terminal hydroxyl groups are linked to other chains using tetraethyl ortho-silicate. Phenol signals are observable but significantly broadened, likely due to high internal viscosity of the swollen gel. A few impurities are present in this spectrum in the range of 1.0 to 6.0 ppm.

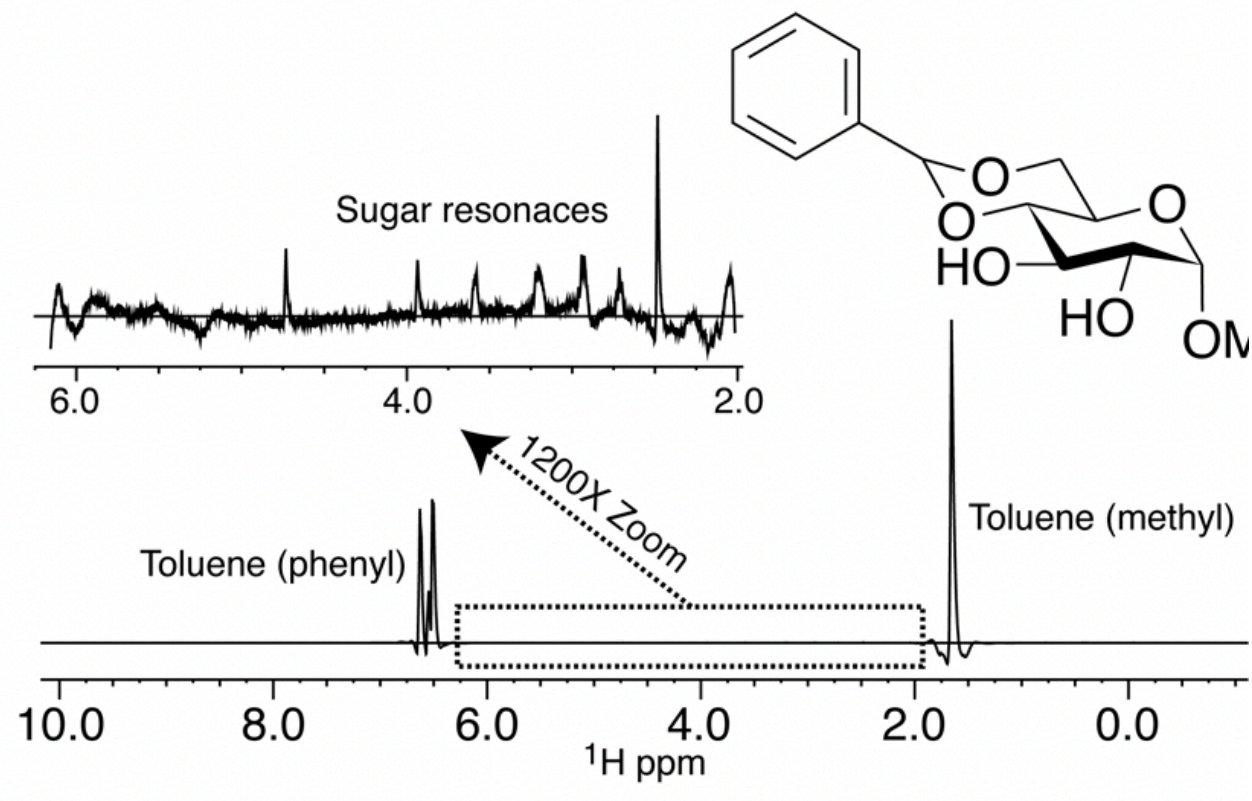


**Figure 6.4:** $^1$H NMR spectrum of GP gel in 60:40 v/v toluene:$CDCl_3$. The structure of GP is shown and resonances are shown in the inset spectrum.

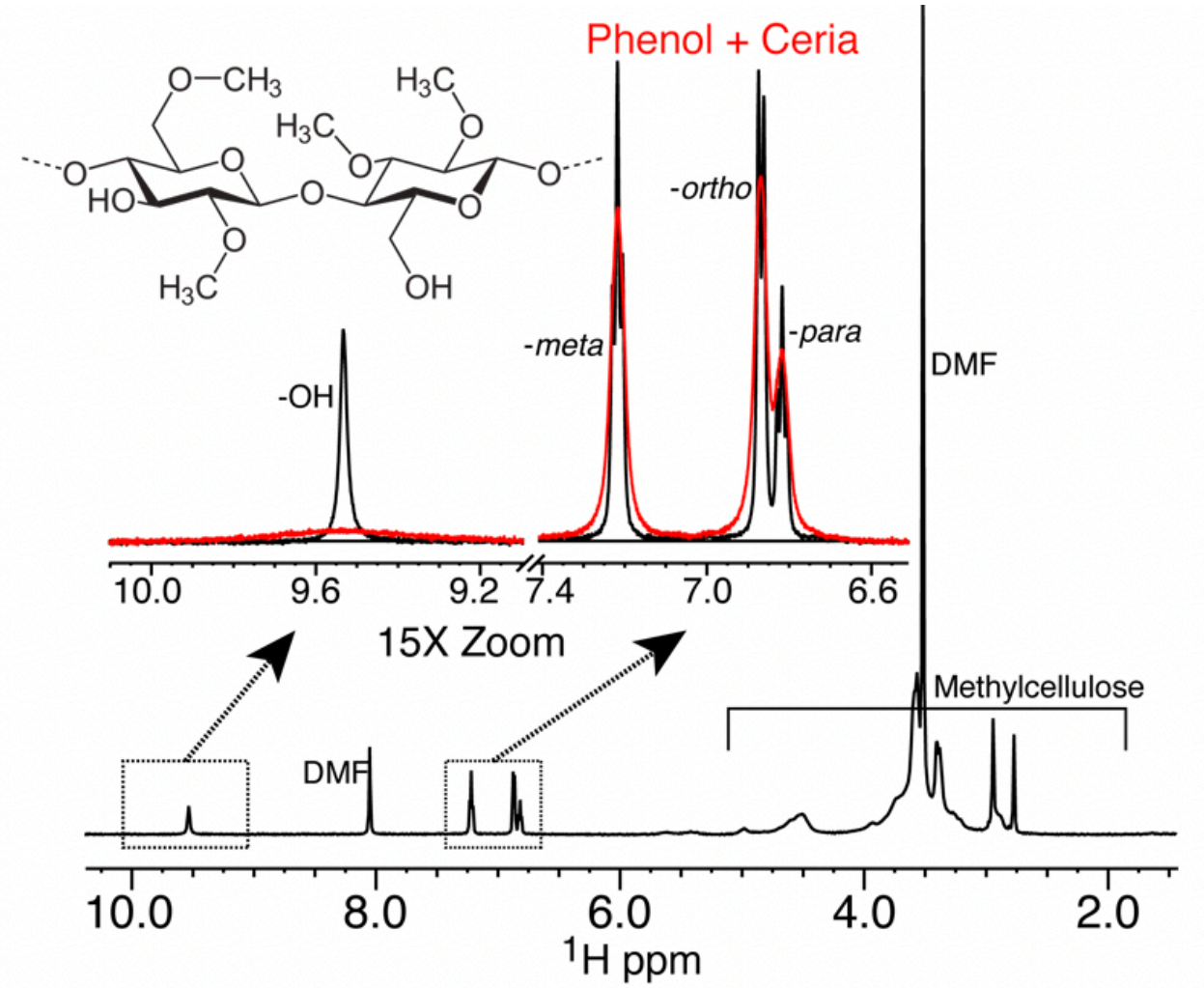


**Figure 6.5:** $^{1}$H NMR spectrum of 10mM phenol in the absence and presence of 1 wt% ceria nanocubes in a 1.5wt% MeC gel in DMF-$d_7$. The structure of MeC is shown and resonances indicated with annotations. The inset spectrum shows a comparison of the effect of phenol linewidths with (red) and without (black) NPs. Transverse relaxation rates for phenol increases in the presences of NPs.

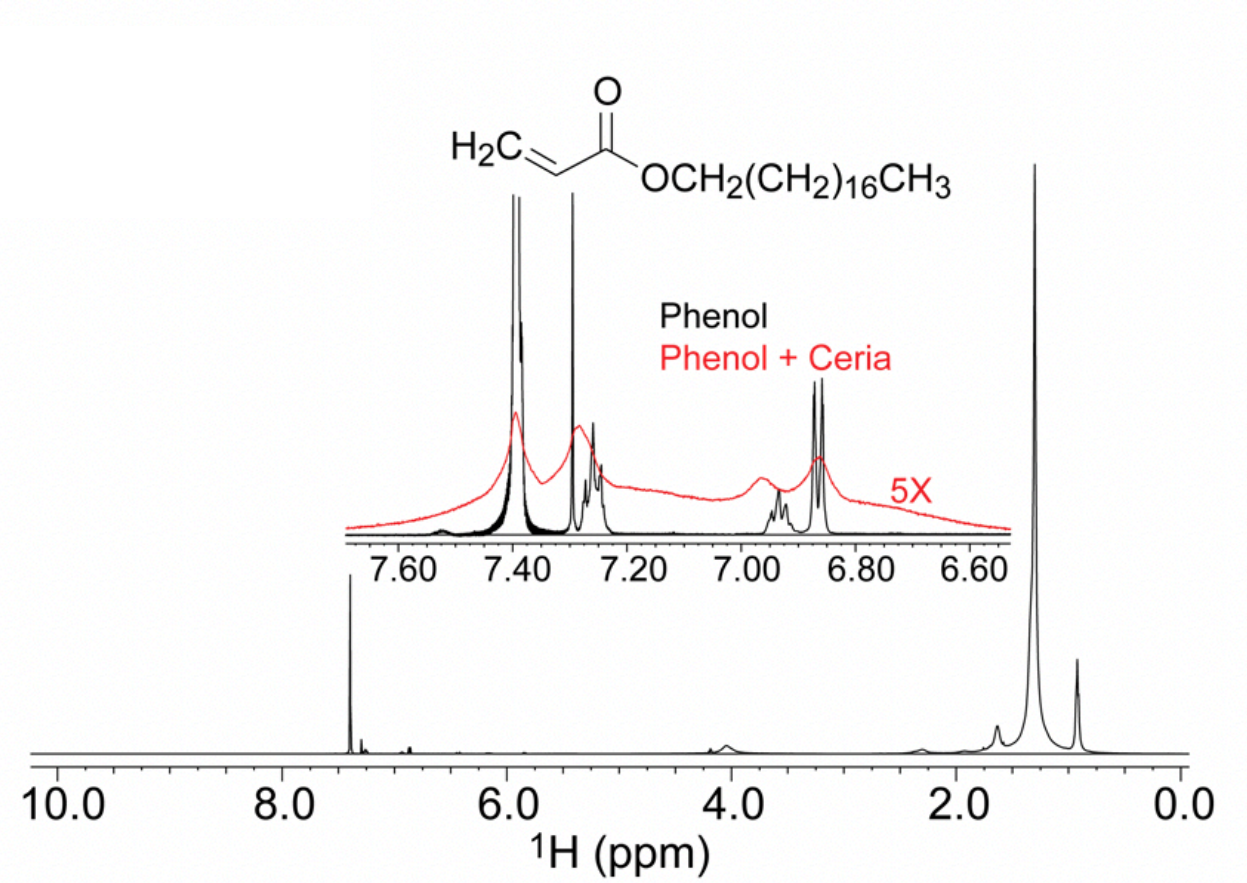


**Figure 6.6:** $^{1}$H NMR spectrum of 10mM phenol in the absence and presence of 0.25 wt% ceria nanocubes in a 3wt% PSA gel swollen with $CDCl_3$. The PSA monomer is shown and gel signals are confined to the range of 0.8 to 2.0 ppm. Residual benzene signal (from washing step during preparation) is seen at 7.4 ppm and residual $CHCl_3$ at 7.28 ppm. The inset spectrum shows a comparison of the effect of phenol linewidths with (red) and without (black) NPs. Line widths are very significantly increased upon the addition of NPs into the PSA gel matrix.

## Tables

**Table 6.1: A** summary of the six criteria to define whether a gel is suitable for use in ligand/NP NMR studies. There are three denotation to indicate ideal (✓), intermediate/unknown (•), or poor (X) qualities for each of the criteria. Preparatory conditions are the most subjective category but ideal to poor is mostly ranked by how much time is required to prepare the gel and complexity of the process / materials required. In some cases, slow gel setting prevents the use of NPs that settle too quickly.

| | Low Residual Signal | Low Internal Viscosity | Large Pore Size | Macroscopic Integrity | Solvent Compatibility | Preparatory Conditions |
|---|---|---|---|---|---|---|
| **Agarose** | ✓ | ✓ | ✓ | ✓ | X | ✓ |
| **PS** | X | • | • | ✓ | ✓ | X |
| **PDMS** | ✓ | X | • | ✓ | ✓ | X |
| **GP** | ✓ | ✓ | • | X | X | ✓ |
| **MeC** | ✓ | ✓ | • | ✓ | X | ✓ |
| **PSA** | • | ✓ | ✓ | ✓ | ✓ | X |

# CHAPTER 7

## COLLABORATIVE PROJECTS INVOLVING NMR ANALYSIS

*These works are those of collaborative projects performed during my studies of which a substantial amount of time and effort was dedicated. Each collaborative project has its own section below with a brief summary as well as an in-depth description of my contributions.*

## PART 1

### SUSTAINABLE SCALABLE SYNTHESIS OF SULFIDE NANOCRYSTALS AT LOW COST WITH AN IONIC LIQUID SULFUR PRECURSOR

Published in *Nature Communications* in October 2018.

Bin Yuan[1,2], Timothy Egner[3], Vincenzo Venditti[3,4], and Ludovico Cademartiri[1,2,5]

---

[1]Department of Materials Science & Engineering, Iowa State University
[2]Department of Chemical & Biological Engineering, Iowa State University
[3]Department of Chemistry, Iowa State University
[4]Roy J. Carver Department of Biochemistry, Biophysics, and Molecular Biology, Iowa State University
[5]Ames Laboratory, U.S. Department of Energy

### Abstract

Increasing the sustainability of nanocrystals is crucial to their application and the protection of the environment. Sulfur precursors for their synthesis are commonly obtained through multiple steps from $H_2S$, only to be converted back to $H_2S$ during the synthesis of the nanocrystals. This convoluted process requires energy, reduces yields, increases waste and auxiliaries, and complicates recycling. Using $H_2S$ directly could drastically improve sustainability, but is prevented by toxicity and handling. We here show that $H_2S$ is stabilized by reaction with oleylamine (OLA) (the most common and versatile ligand in nanoparticle synthesis) to form an

ionic liquid precursor that addresses all major principles of green chemistry: it is made in one exothermic step, it leaves the reaction yielding a safer product and allowing the separate recycling of the precursors, and it produces high quality nanocrystals with high yields (sulfur yield > 70%) and concentrations (90 g $L^{-1}$) in ambient conditions [1].

**Contributions to project**

*Confirmation of ionic liquid formation with a mixture of $H_2S$ and oleylamine by NMR*

One of the critical steps in producing this green product was to characterize the use of $H_2S$ with OLA to produce nanoparticles. The mechanism of this procedure was unknown, however similar studies had been performed previously that suggested that $H_2S$ and OLA form an ionic liquid composed of OLA-$NH_3^+$ + $^-$SH salt [2]. NMR and IR were used to study this ionic liquid. The NMR spectrum of OLA and OLA + H2S in $CDCl_3$ is shown in **Figure 7.1**. The $H_2N$ signal appears at 0.98 ppm as a broad singlet. It is expected that this signal would disappear if an ionic liquid is formed with $H_2S$ due to exchange broadening or significant changes in electronic environment. With the addition of $H_2S$, the signal at 0.98 ppm is no longer present (**Figure 7.1c**). Additionally, we confirmed this finding with IR, and similar trends were observed with the shifting of vibrational frequencies (**Figure 7.1a**). In all, we were able to confirm with NMR analysis that an ionic liquid was indeed being formed between OLA and $H_2S$.

*Quantification of ligand concentration in the metal sulfide nanoparticles dispersion*

Another important feature in characterizing this system was to determine if it out-performs other, non-green methods. The yield of this process needs to be determined to properly quantify its performance, which must consider that the product mixture contains ligands as well as the metal sulfide nanoparticles. NMR was utilized to determine the concentration of ligands

within the sample. Ligand signals are present, but most are relatively broad, which is due to the slow tumbling of the nanoparticles. The ligand is expected to bind to the nanoparticle through the amino group, which would leave the terminal methyl group farthest from the nanoparticle surface, with exposure to the bulk solvent. In this binding mode, NMR signals originating from nuclei near the surface will appear broader due to their relatively lower rotational freedom. At the terminal methyl group, rotational freedom is comparable to OLA in CDCl3 without nanoparticles, and thus can be used for quantification. The intensity of OLAs methyl signal was compared to an internal standard with a known concentration. The ligand wt% of the original nanoparticle sample was determined to be 25(± 0.02)%, and thus the overall yield is scaled down by 75%. Uncertainty in OLA wt% is based on uncertainty of the NMR measurement. This yield of 90.2g of product per liter of reaction mixture is 31% higher than the current highest reported yield [3]. Thus, this metal sulfide nanoparticle synthetic process with OLA and $H_2S$ is quite efficient even compared to non-green methods.

## Figures

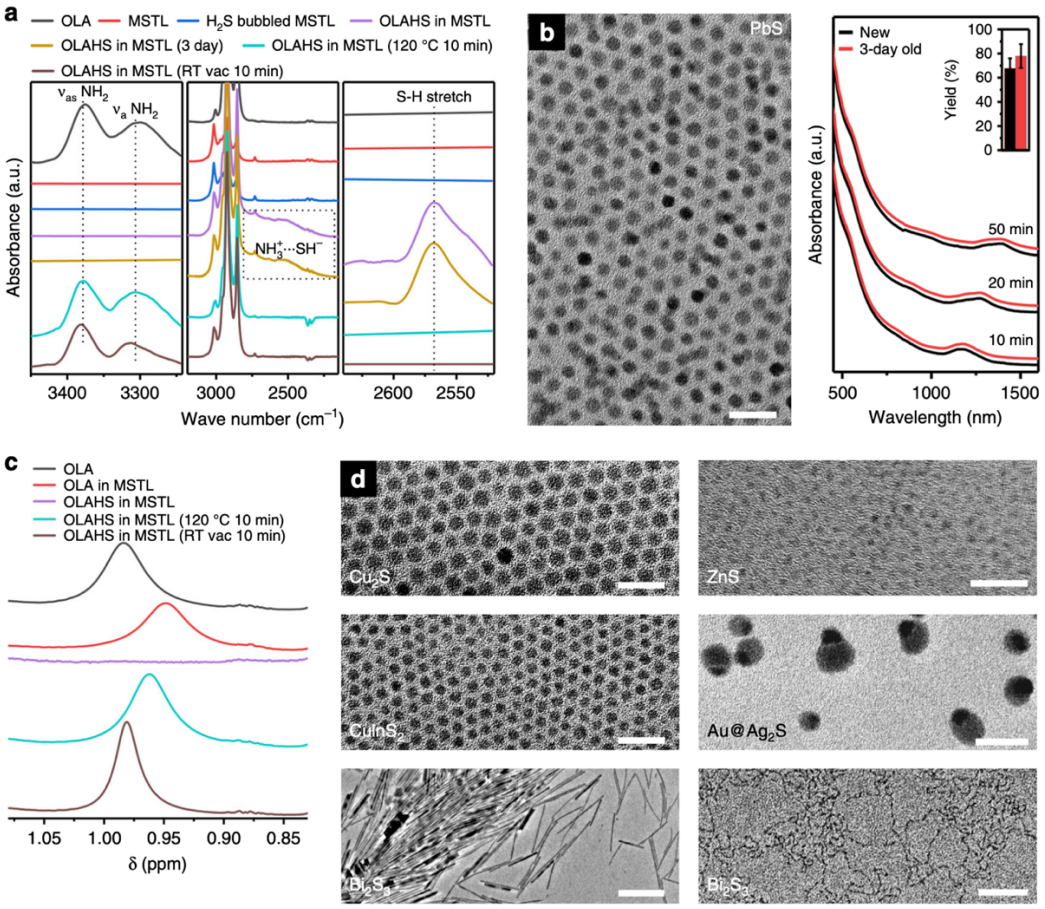


**Figure 7.1:** Characterization of the OLAHS ionic liquid precursor and its application in the synthesis of chalcogenides nanocrystals. **a** Background-subtracted attenuated total reflectance Fourier-transform infrared (ATR-FTIR) absorption spectrum of oleylamine (black), mesitylene (red), mesitylene after being bubbled with $H_2S$ (blue), OLAHS in mesitylene (purple), OLAHS in mesitylene (3-day old) (yellow), OLAHS in mesitylene after being kept at 120 °C for 10 min (green), and OLAHS in mesitylene after being kept under vacuum ($10^{-3}$ Torr) at room temperature for 10 min (brown). 'MSTL', 'RT', and 'vac' is short for 'mesitylene', 'room temperature', and 'vacuum', respectively. **b** TEM image of the as-synthesized PbS nanocrystals using OLAHS (left) (scale bar: 20 nm) and the UV–Vis–NIR absorption spectra for samples collected at different growth times (right) using a fresh (black) and 3-day old (red) OLAHS precursor. Inset: comparison of sulfur yield of reaction using fresh or 3-day old OLAHS as precursor. The error bars depict the standard deviation from more than five samples collected from the same batch of reaction. **c** $^1$H NMR spectra of the amine protons. **d** TEM images of $Cu_2S$, ZnS, $CuInS_2$, Au@$Ag_2S$ (janus nanoparticles), and $Bi_2S_3$ nanocrystals synthesized using OLAHS as precursor (scale bar is 20 nm except for $Bi_2S_3$ nanorods (scale bar: 180 nm) and $Bi_2S_3$ nanowires (scale bar: 40 nm)). All the spectra have been offset for clarity.

**PART 2**

# HYDROGEL-BASED TRANSPARENT SOILS FOR ROOT PHENOTYPING IN VIVO

Published in *PNAS* in May, 2019.

Lin Ma[1], Yichao Shi[1], Oskar Siemianowski[1], Bin Yuan[2], Timothy Egner[3], Seyed Vahid Mirnezami[4], Kara Lind[1], Baskar Ganapathysubramian[4], Vincenzo Venditti[3], and Ludovico Cademartiri[1,2,5]

---

[1]Department of Materials Science & Engineering, Iowa State University
[2]Department of Chemical & Biological Engineering, Iowa State University
[3]Department of Chemistry, Iowa State University
[4]Department of Mechanical Engineering, Iowa State University
[5]Ames Laboratory, U.S. Department of Energy

## Abstract

Root phenotypes are increasingly explored as predictors of crop performance but are still challenging to characterize. Media that mimic field conditions (e.g., soil, sand) are opaque to most forms of radiation, while transparent media do not provide field-relevant growing conditions and phenotypes. We describe here a “transparent soil” formed by the spherification of hydrogels of biopolymers. It is specifically designed to support root growth in the presence of air, water, and nutrients, and allows the time-resolved phenotyping of roots in vivo by both photography and microscopy. The roots developed by soybean plants in this medium are significantly more similar to those developed in real soil than those developed in hydroponic conditions and do not show signs of hypoxia. Lastly, we show that the granular nature and tunable properties of these hydrogel beads can be leveraged to investigate the response of roots to gradients in water availability and soil stiffness [4].

**Contributions to project**

*Determination of the sugar composition in hydrogel transparent soil*

A critical step in the characterization of the hydrogel beads was to determine the source of optical quality. In the process of creating different batches of the beads, there were differences in coloration and clarity depending on the source material for the batches. The variability between batches made the process unreliable and difficult to repeat to obtain high quality transparent soil.

The hydrogels are formed from gellum gum and alginate, a polymers of sugars mannuronic acid (M) and guluronic acid (G), which can form a hydrogel in the presence of $Ca^{2+}$. One hypothesis to explain why there was variability between batches was that there were differences in the composition of G and M sugars between the batches.

In order to address this question, we utilized NMR to determine the composition of sugars in the samples of alginate. In order to perform high resolution NMR, the carbohydrates were hydrolyzed to obtain shorter length sugars so that sharper signals could be acquired. Based on recommendations from a previous study of this material, NMR data were acquired at 85°C to allow for sharper spectral lines [5]. **Figure 7.2a** shows the $^{13}C$ NMR spectrum of the alginate in water at 85°C. Due to the different sugar connectivites, there are eight signals expected in the anomeric carbon region (MMM, MMG, GMM, GMG, GGG, GGM, MGG, GMG) where the center letter represents the observed sugar. The relative intensities of these signals represent the relative fraction of each combination in the overall composition of the polymer. Extraneous signals in this region are likely sugars at the end of a polymer chain. **Figure 7.2a** shows spectra

from five separate batches of alginate. Upon comparison with light transmittance data, there appears to be a correlation between the G/M sugar ratio and percent transmittance of light (**Figure 7.2c**). We found that alginate with higher G sugars composition tend to have better optical clarity. This information can now be used to accurately predict the quality of transparent soils produced from a given alginate batch and additionally offers a large improvement to high quality reproduction.

**Figures**

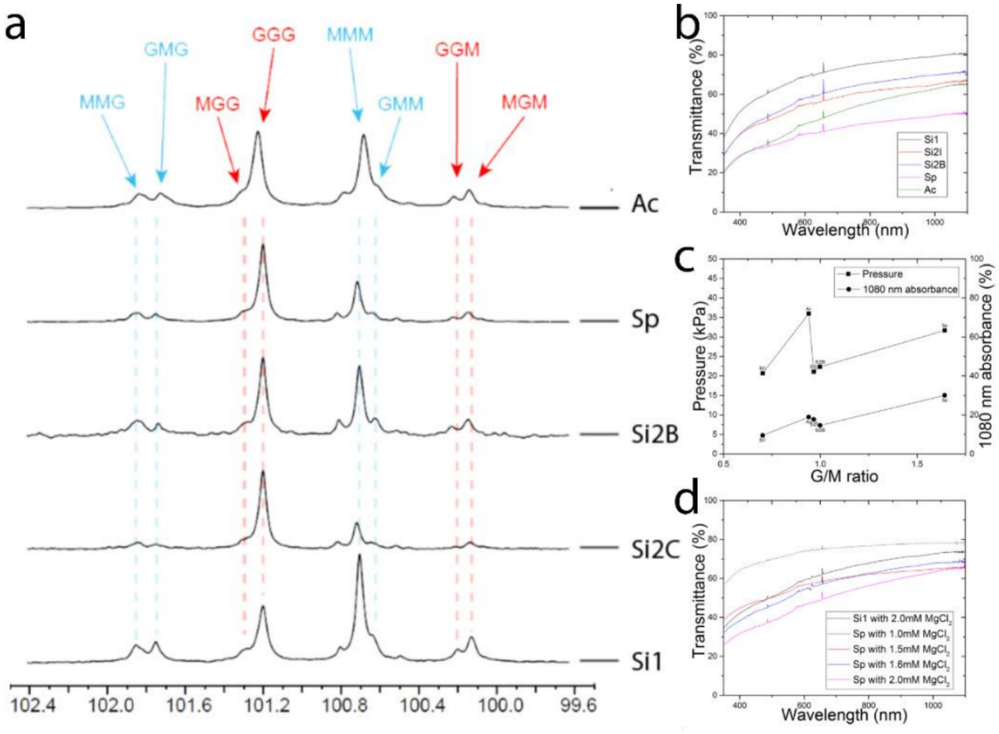


**Figure 7.2: (a)** $^{13}$C-NMR spectrum of five different sodium alginate products (Si1, Si2I, Si2B, Sp, Ac). **(b)** transmittance of visible light through 1 X MS TS beads made by 1.2% gel and 10mM $MgCl_2$ and transmittance is tested by UV-Vis with DI water as background. **(C)** relationship between mechanical property and absorbance of TS beads, and the G/M ratio of sodium alginate. The absorbance is calculated from UV-Vis results by the Beer-Lambert Law $T=10^{-A}$. **(d)** transmittance of various TS beads made by mixing gel with Sp and $MgCl_2$ of various concentration. Transmittance is tested by UV-Vis with DI water as background.

**PART 3**

# ACTIVE SITE BREATHING OF HUMAN ALKBH5 REVEALED BY SOLUTION NMR AND ACCELERATED MOLECULAR DYNAMICS

Published in *Biophysics Journal* in November 2018.

Jeffrey Purslow[1], Trang Nguyen[1], Timothy Egner[1], Rochelle Dotas[1], Balabhadra Khatiwada[1], Vincenzo Venditti[1,2]

[1]Department of Chemistry, Biophysics, and Molecular Biology, Iowa State University
[2]Roy J. Carver Department of Biochemistry, Biophysics, and Molecular Biology, Iowa State University.

## Abstract

AlkB homolog 5 (Alkbh5) is one of nine members of the AlkB family, which are nonheme Fe2+/α-ketoglutarate-dependent dioxygenases that catalyze the oxidative demethylation of modified nucleotides and amino acids. Alkbh5 is highly selective for the $N^6$-methyladenosine modification, an epigenetic mark that has spawned significant biological and pharmacological interest because of its involvement in important physiological processes, such as carcinogenesis and stem cell differentiation. Herein, we investigate the structure and dynamics of human Alkbh5 in solution. By using $^{15}N$ and $^{13}C$methyl relaxation dispersion and $^{15}N$-$R_1$ and $R_{1\rho}$ NMR experiments, we show that the active site of apo Alkbh5 experiences conformational dynamics on multiple timescales. Consistent with this observation, backbone amide residual dipolar couplings measured for Alkbh5 in phage pf1 are inconsistent with the static crystal structure of the enzyme. We developed a simple approach that combines residual dipolar coupling data and accelerated molecular dynamics simulations to calculate a conformational ensemble of Alkbh5 that is fully consistent with the experimental NMR data. Our structural model reveals that Alkbh5

is more disordered in solution than what is observed in the crystal state and undergoes breathing motions that expand the active site and allow access to α-ketoglutarate. Disordered-to-ordered conformational changes induced by sequential substrate/cofactor binding events have been often invoked to interpret biochemical data on the activity and specificity of AlkB proteins. The structural ensemble reported in this work provides the first atomic-resolution model of an AlkB protein in its disordered conformational state to our knowledge [6].

**Contributions to Project**

The main goal of this project was to determine the structural and dynamic qualities of AlkBH5 to gain insight towards understanding its catalytic function and selectivity. We were able to utilize a variety of NMR experiments and computational MD simulations, to address these questions. A 200ns accelerated molecular dynamics (aMD) simulation was performed using an Amber ff14sb forcefield. In conjunction with this simulation, RDCs were collected on AlkBH5 using pf1 phage as an alignment medium. The aMD trajectory was binned into representative structures by separating into *n* pieces and then averaging the atomic coordinates for each atom (excluding hydrogen, and adding back at after averaging utilizing ideal geometry). These sets of structures comprised the structural ensemble that were used to compare with the experimental RDCs.

My contribution involved writing a program (available at https://group.chem.iastate.edu/Venditti/downloads.html) that allows optimization of the alignment of tensor parameters of an ensemble of structures against experimental RDCs. This program is based on the theory described by Venditti et al. [7] and is used directly within MatLab. This program iteratively optimizes five alignment tensor parameters for each structure in the

ensemble. Therefore, if a three-member ensemble is desired, there must be at least 15 measured RDCs for a non-degenerate fit. The programs operation was optimized for fast computation and is suitable for general application of this structural validation methodology. The quality of the agreement between MD representative structures and RDC data is given by $R_{free}$ [8] and reaches a plateau at a 20 member ensemble for AlkBH5 (**Figure 7.3b**). There were a total of 188 measured RDCs, which is a sufficient quantity of data to satisfy the necessary fitable parameters. **Figure 7.4a (left panel)** shows a representation of the structure and relative dynamics of AlkBH5 based on the experimentation and simulations in this study.

**Figures**

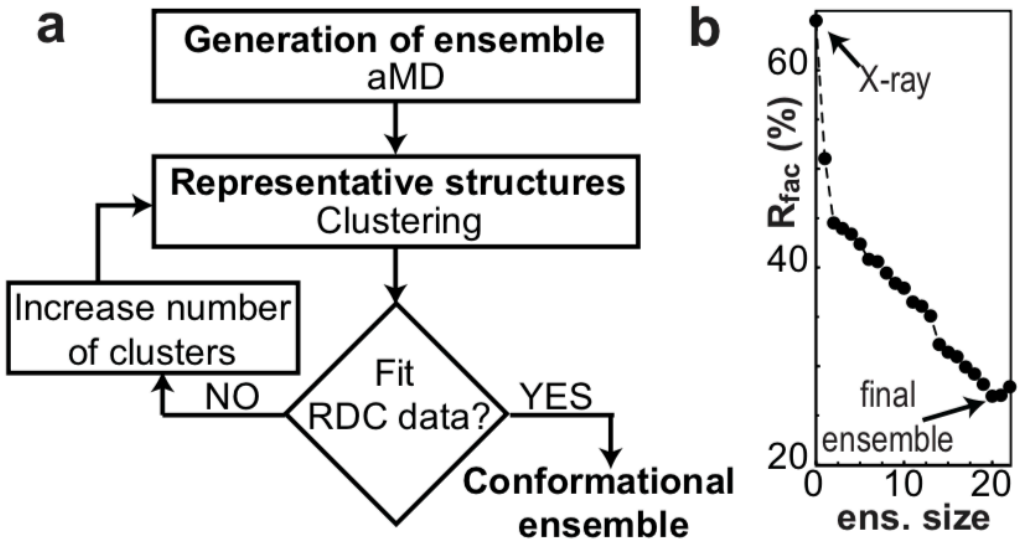


**Figure 7.3:** aMD/RDC refinement of protein conformational ensemble. (**a**) Schematic of the aMD/RDC ensemble calculation protocol developed in this work. A preliminary pool of structures is generated by aMD and subsequently filtered to maximize the agreement between experimental and back-calculated RDCs. The ensemble size (i.e., the number of conformers in the ensemble) is increased until the minimum R-factor is reached. (**b**) R-factor versus ensemble size for the aMD/RDC ensemble refinement of apo Alkbh5. A 20-member ensemble is required to fit the experimental RDCs.

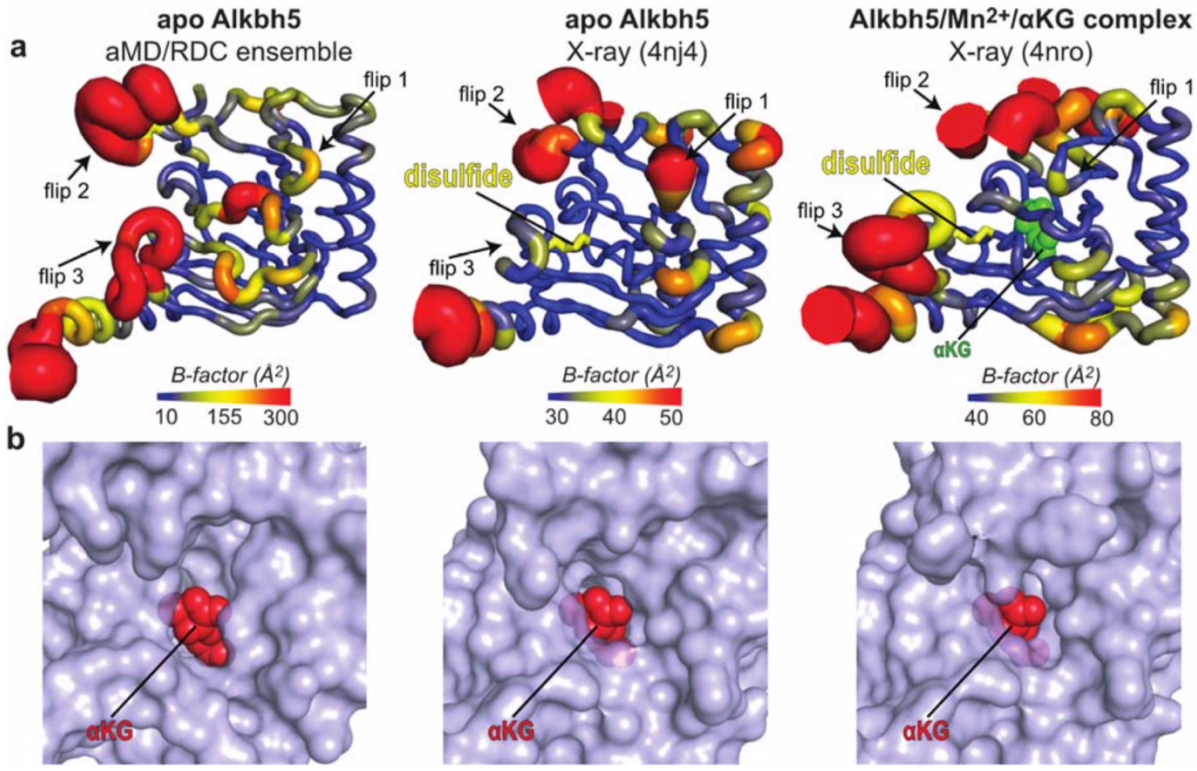


**Figure 7.4:** Comparison between solution and crystal structure of humal Alkbh5. (**a**) Sausage representation of the aMD/RDC ensemble (left), X-ray structure of apo Alkbh5 (center), and X-ray structure of Alkbh5 complexed with $Mn^{2+}$ and αKG (right). Cartoons are colored according to the B-factor, as indicated by the color bar. B-factors for the aMD/RDC ensemble were calculated using the formula $B^i=8\pi^2U_i^2$, where $B_i$ and $U_i$ are the B-factor and mean- square displacement of atom *i*, respectively. The $Cys^{230}$-$Cys^{267}$ disulfide bridge is shown as yellow stick. αKG is shown as green spheres. (**b**) Close-up view of the αKG binding pocket in the aMD/RDC ensemble (left), X-ray structure of apo Alkbh5 (center), and X-ray structure of Alkbh5 complexed with $Mn^{2+}$ and αKG (right). For the aMD/RDC ensemble, the conformer with the widest αKG binding pocket is displayed. One αKG molecule has been modelled in the binding site of the apo protein (left and center panels). Alkbh5 is shown as transparent light blue surface. αKG is shown as red spheres.

**References**


1. Yuan B, Egner TK, Venditti V, Cademartiri L. Sustainable scalable synthesis of sulfide nanocrystals at low cost with an ionic liquid sulfur precursor. Nat Commun. 2018;9: 4078. doi:10.1038/s41467-018-06549-8

2. Thomson JW, Nagashima K, Macdonald PM, Ozin GA. From Sulfur−Amine Solutions to Metal Sulfide Nanocrystals: Peering into the Oleylamine−Sulfur Black Box. J Am Chem Soc. 2011;133: 5036–5041. doi:10.1021/ja1109997

3. Zhang J, Gao J, Miller EM, Luther JM, Beard MC. Diffusion-Controlled Synthesis of PbS and PbSe Quantum Dots with in Situ Halide Passivation for Quantum Dot Solar Cells. ACS Nano. 2014;8: 614–622. doi:10.1021/nn405236k

4. Ma L, Shi Y, Siemianowski O, Yuan B, Egner TK, Mirnezami SV, et al. Hydrogel-based transparent soils for root phenotyping in vivo. Proc Natl Acad Sci U S A. 2019;116: 11063–11068. doi:10.1073/pnas.1820334116

5. Grasdalen H, Larsen B, Smisrod O. 13C-n.m.r. studies of monomeric composition and sequence in alginate. Carbohydr Res. 1981;89: 179–191. doi:https://doi.org/10.1016/S0008-6215(00)85243-X

6. Purslow JA, Nguyen TT, Egner TK, Dotas RR, Khatiwada B, Venditti V. Active Site Breathing of Human Alkbh5 Revealed by Solution NMR and Accelerated Molecular Dynamics. Biophys J. 2018;115: 1895–1905. doi:https://doi.org/10.1016/j.bpj.2018.10.004

7. Venditti V, Egner TK, Clore GM. Hybrid Approaches to Structural Characterization of Conformational Ensembles of Complex Macromolecular Systems Combining NMR Residual Dipolar Couplings and Solution X-ray Scattering. Chem Rev. 2016;116: 6305–6322. doi:10.1021/acs.chemrev.5b00592

8. Clore GM, Garrett DS. R-factor, Free R, and Complete Cross-Validation for Dipolar Coupling Refinement of NMR Structures. J Am Chem Soc. 1999;121: 9008–9012. doi:10.1021/ja991789k

# CHAPTER 8

## GENERAL SUMMARY AND CONCLUSION

NPs are being widely applied to many different types of chemical, biological, and physical problems but certain aspects of their chemical properties still require investigation. NMR is a powerful tool for studying molecular interactions and exchange processes. Within this work, I have described methodological progress towards studying ligand-nanoparticle interactions by solution NMR. As a recap, I will reiterate the important aspects of each chapter.

Within Chapter 4, I describe methodologies for utilizing $^{1}$H DEST and RD NMR measurements to investigate ligand interactions on nanoparticle surfaces. NMR is an important tool for investigating chemical exchange and some of the information gained by this method is inaccessible by other instrumental techniques. Study of these ligand-NP interactions are enabled by the use of an agarose gel to prevent nanoparticle settling. We also show that agarose is a suitable gel system for this type of study as it does not interfere with data measurement or quality. We found that phenol takes part in a three-site exchange on a ceria nanoparticle consisting of unbound, weakly associated, and tightly associated states.

In Chapter 5, I expand upon the findings described in Chapter 4 through the use of more sophisticated analysis modelling and more comprehensive data sampling. $^{13}$C NMR experiments were added to the repertoire of surface contrast NMR (scNMR) experiments, which now includes $^{1}$H DEST, $^{1}$H RD, $^{13}$C DEST, $^{13}$C RD, and $^{13}$C $R_1$ measurements. Again, we found that phenol interacted with ceria through two different surface binding modes: a weakly associated state and a tightly associated state as with our previous paper. Upon further analysis, the we found that the weakly associated state was consistent with hydrogen bonding, while the tightly associated

state was consistent with an interaction with an oxygen vacant site. Together, these findings agree with our conclusions described in Chapter 4 but we now have enhanced the level of detail. In addition, through these analyses, we found that phenol appears to bind flat onto Pd supported on ceria, while is the active nanoparticle catalyst for the hydrogenation reaction of phenol to cyclohexanone. The methodologies laid out in Chapter 4 and 5 can be used to investigate other nanoparticle and ligand system to understand mechanisms to improve catalytic processes.

The previous chapters demonstrated the utility of the scNMR experiments and in Chapter 6 I initiated the investigation into other gel systems that are compatible with organic solvents, which would increase the applicability of these methods. Agarose had been used in the previous chapters as a gel system to prevent the fast settling of NPs to allow for NMR analysis. However, agarose is only compatible with aqueous solvent systems and thus dramatically limits general applications. Here, I described several organic compatible gel systems that can be used for NMR studies under various conditions. In the process of characterizing these gels, I formulated six criteria that predict the overall usability of the gels: low residual gel signal, low internal viscosity, large pore size, solvent compatibility, and simple preparatory conditions. Using these criteria, I found that the most promising gel system was PSA for its solvent compatibility, relatively quick gelation time, low NMR background signal, and favorable mechanical properties.

In Chapters 3 and 7 I describe my projects outside of scNMR methodologies. In Chapter 3, I covered an emerging method for protein ensemble refinement using RDC and SAXS simultaneously. I was able to use information gained from studying this field to develop analysis software for the use of RDCs and MD simulations for protein structure refinement. These tools

enabled our lab to identify dynamic motions in AlkBH5 that may be necessary for ligand binding and specificity.

In other projects, I was able to apply NMR techniques to investigate the composition of alginate for use in transparent soil. The quality of *transparent* soil is heavily impacted by its optical properties. An issue had arisen that different batches of alginate had different physical transparent properties. We aimed to determine what was causing the differences in transparency between the batches by using NMR methodologies. I contributed to identifying the solution to issues with optical clarity of transparent soils by determining the sugar type composition of alginate. We found great variability in different batches of alginate which leads to certain origins having superior properties for use in transparent soils. In another project, I helped to characterize a sulfur source ionic liquid that was used for the green synthesis of metal sulfide nanoparticles. Additionally, I determined compositions of reaction products (ligands verses NPs) by NMR to allow for accurate determination of yields. With the proper NP composition, the yields were compared with recent reports in the literature and showed OLASH out performs other less green sulfur precursors.