Title

NMR-Based Tools for Structural Biology of Complex Systems: from Heterochromatin Gels to Mammalian Cells

Permalink

https://escholarship.org/uc/item/7nv2109m

Author

Ackermann, Bryce Eric

Publication Date

2022

UNIVERSITY OF CALIFORNIA SAN DIEGO

NMR-Based Tools for Structural Biology of Complex Systems: from Heterochromatin Gels to Mammalian Cells

A dissertation submitted in partial satisfaction of the
requirements for the degree of Doctor of Philosophy
in
Chemistry

by

Bryce Eric Ackermann

Committee in charge:

Professor Galia T. Debelouchina, Chair
Professor Kevin Corbett
Professor James Kadonaga
Professor Elizabeth Komives
Professor Joel Yuen Zhou

2022

The dissertation of Bryce Eric Ackermann is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

University of California San Diego

2022

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

CD	chromodomain
CE	cross effect
CP	cross-polarization
CSA	chemical shift anisotropy
CSD	chromoshadow domain
CTE	c-terminal extension
DARR	Dipolar Assisted Rotational Resonance
DNA	deoxyribose nucleic acid
DNP	dynamic nuclear polarization
EMSA	electrophoretic mobility shift assay
B_0	external magnetic field
FRAP	fluorescence recovery after photobleaching
GFP	green fluorescence protein
γ	gyromagnetic ratio
HP1	heterochromatin protein 1
HSQC	Heteronuclear Single Quantum Coherence
INEPT	Insensitive Nuclei Enhanced by Polarization Transfer
kDa	kilodalton
LLPS	liquid-liquid phase separation
K9me3	lysine 9 trimethylation
K20me3	lysine 20 trimethylation

K27me3	lysine 27 trimethylation
MAS	magic angle spinning
MRI	magnetic resonance imaging
MN	mononucleosome
NTE	n-terminal extension
NMR	nuclear magnetic resonance
pHP1	phosphorylated heterochromatin protein 1
PA	polarization agent
PTM	post-translational modification
PXVXL	proline/any amino acid/valine/any amino acid/leucine
RNA	ribose nucleic acid
S/N	signal to noise
SHL	superhelical location
TS	tetrasome
TOBSY	Total Through Bond Correlation Spectroscopy
1D	one-dimensional
2D	two-dimensional
3D	three-dimensional

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I acknowledge us, we really tried our best.

Chapter 3, in full, is a reprint of the material as it appears in Heterochromatin protein HP1α gelation dynamics revealed by solid-state NMR spectroscopy. 2019. Ackermann, Bryce E. and Debelouchina, Galia T. Angewandte Chemie International Edition. The dissertation author was the primary investigator and author of this material.

Chapter 4 is work to be contributed to a publication by Nesreen Elathram, Evan Clark, and Galia T. Debelouchina.

Chapter 5, in full, is a reprint of the material as it appears in Targetable tetrazine-based dynamic nuclear polarization agents for biological systems. 2020. Lim, Byung Joon*, Ackermann, Bryce E.*, Debelouchina, Galia T. ChemBioChem. The dissertation author was the co-primary investigator and co-author of this material.

Chapter 6, in full, is a reprint of the material submitted in A comparative study of nitroxide-based biradicals for dynamic nuclear polarization in cellular environments. 2022. Ackermann, Bryce E. Lim, Byung Joon, Narayanan, Sirish, Debelouchina, Galia T. The dissertation author was the primary investigator and author of this material.

VITA

2017 Bachelors of Science, Biochemistry and Molecular Biology, University of California, Davis

2018 Masters of Science, Chemistry, University of California San Diego

2022 Doctorate of Philosophy, Chemistry, University of California San Diego

PUBLICATIONS

Ackermann, Bryce E., Lim, Byung Joon, Narayanan, Sirish, Debelouchina, Galia T. "A comparative study of nitroxide-based biradicals for dynamic nuclear polarization in cellular environments." *In submission, 2022*

Her, C., Phan, T. M., Jovic, N., Kapoor, U., **Ackermann, B. E.**, Rizuan, A., . . . Debelouchina, G. T. "Molecular interactions underlying the phase separation of HP1α: Role of phosphorylation, ligand and nucleic acid binding." *bioRxiv*, 2022, doi:10.1101/2022.06.20.496886

Elathram, Nesreen, **Ackermann, Bryce E.** & Debelouchina, Galia T. "DNP-enhanced solid-state NMR spectroscopy of chromatin polymers." *Journal of Magnetic Resonance Open*, 2022, 10-11, 100057, doi:https://doi.org/10.1016/j.jmro.2022.100057

Ackermann, Bryce E. and Debelouchina, Galia T. "Emerging Contributions of Solid-State NMR Spectroscopy to Chromatin Structural Biology." *Front. Mol. Biosci*, 2021, 8:741581. doi: 10.3389/fmolb.2021.741581

Lim, Byung Joon*, **Ackermann, Bryce E.***, Debelouchina, Galia T. "Targetable tetrazine-based dynamic nuclear polarization agents for biological systems." *ChemBioChem,* 2020, doi: 10.1002/cbic.201900609. * signifies equal contributions

Ackermann, Bryce E. and Debelouchina, Galia T. "Heterochromatin protein HP1α gelation dynamics revealed by solid-state NMR spectroscopy." *Angewandte Chemie International Edition*, vol. 58, 6300-6305, 2019, doi:10.1002/anie.201901141.

ABSTRACT OF THE DISSERTATION

NMR-Based Tools for Structural Biology of Complex Systems: from Heterochromatin Gels to Mammalian Cells

by

Bryce Eric Ackermann

Doctor of Philosophy in Chemistry

University of California San Diego, 2022

Professor Galia T. Debelouchina, Chair

Magic angle spinning (MAS) nuclear magnetic resonance (NMR) spectroscopy is a structural biology technique capable of characterizing complex and heterogenous samples. Here, we develop novel ways of performing MAS NMR to characterize two complex systems - the phase separated environment formed by heterochromatin protein

1α (HP1α), and the cellular interior of mammalian cells. HP1α is a protein fundamental to the organization and gene regulation of heterochromatic genome territories. HP1α can undergo a process called liquid-liquid phase separation, generating concentrated hubs of heterochromatin components. We applied MAS NMR to track the structural dynamics of phase separated HP1α through its development into a gel which resembles the matured state of heterochromatin in cells. Beyond the capabilities of other structural techniques, our methodology was able to identify the residues in HP1α that participated in the crosslinking interactions in the gel state. The gelation process slowed in the presence of chromatin, leading us to further probe the interactions between HP1 and nucleosomes. We investigated the interactions of HP1 with the nucleosome, and studied how HP1 phase separation might influence other proteins that act on the nucleosome, such as transcription factors and chromatin remodelers. We have also developed structural biology methodology for proteins in mammalian cells. Our efforts center around dynamic nuclear polarization (DNP), a tool to enhance the sensitivity of NMR, aiding the detection of biomolecules directly in their native cellular environment. To conduct DNP, samples must be doped with a biradical polarization agent (PA) that transfers polarization to nearby nuclei. We first designed and synthesized a novel bio-orthogonal PA, TTz, that employs a tetrazine-based reaction to target proteins in the cellular milieu. Our strategy was generalized to several proteins and was able to selectively enhance the NMR signal of our protein of interest over the background. We then developed the PA, POPAPOL, to address the susceptibility of PAs to radical reduction in the cellular environment. POPAPOL exhibited longer radical lifetimes than other popular PAs, promising a new design route for future PAs for applications in cells.

Chapter I

Heterochromatin Introduction

1.0 Synopsis

This thesis spans concepts from atomic-level protein motions to gene regulation networks to mechanisms of nuclear magnetic resonance. To cover such breadth, the introduction will be divided into a structural biology and a magnetic resonance section.

The structural biology section will cover the role of chromatin in gene regulation (1.1), the importance of heterochromatin protein 1 (1.2), the principles of liquid-liquid phase separation 1 (1.3), and the influence of heterochromatic components in liquid-liquid phase separation (1.4).

The nuclear magnetic resonance section will begin with the fundamentals of nuclear magnetic resonance (2.1), the application of solution (2.2) and solid-state methodologies to study protein structure and motion (2.3), magnetic resonance signal enhancement using dynamic nuclear polarization (2.4), and developments to conduct nuclear magnetic resonance spectroscopy in living cells (2.5).

1.1 Chromatin

A crucial principle and central theme for this thesis is that proteins, as we do, require community. The most immediate example of this principle is the formation of protein complexes, where distinct proteins group up non-covalently. One example is the nucleosome, a stable complex responsible for efficient packaging of the genome, composed of DNA and two copies of each histone protein H2A, H2B, H3, and H4 (**Figure 1.1**).[2]

When extended, the DNA contained in each human cell can reach six feet in length. DNA must therefore be tightly packaged to fit inside the micrometer sized

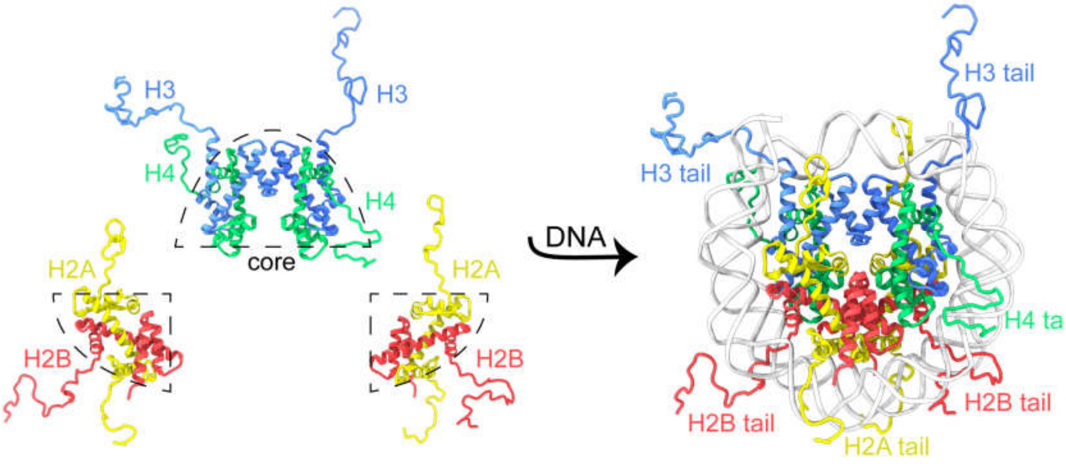


Figure 1.1. Architecture of the nucleosome. (Left) One H3-H4 tetramer and two H2A-H2B dimers are required for octamer formation. The folded alpha helical histone core is outlined. (Right) DNA is wrapped around the histone octamer to form the nucleosome. Disordered tails extending from the nucleosome are labeled.

nucleus, while retaining the dynamics and organization necessary to support life, all while avoiding entangling throughout the many drastic morphological changes during a cell's life. This is done in part by the nucleosome, which is repeated lengthwise throughout the genome like beads on a string, forming a polymer called chromatin.[3]

The nucleosome structure is possible because of the histone octamer design. It can be helpful to think of histones as having two structural domains, the core and the tail (**Figure 1.1**). The core of a histone is an alpha helical domain sandwiched between other histones and wrapped by DNA. The extensive electrostatic interactions between DNA and the histone core allows the DNA to bend in a spiral around the surface of the histone octamer. The histone tail on the other hand, is a dynamic disordered peptide that extends from the nucleosome surface. The tail can hold onto the outside of the DNA spiral to stabilize the nucleosome or it can recruit many of the chromatin effector proteins that regulate genes. An important feature of the nucleosome is that the bound

DNA is sterically excluded from DNA-binding proteins that may initiate transcription, therefore the nucleosome is a basal repressor of aberrant gene expression.

Zooming out from the nucleosome, chromatin has its own folding patterns.[4] Neighboring nucleosomes can stack in a zig-zag pattern or can open and adopt flexible conformations, and these structures may relate to the accessibility of nucleosomal DNA.[5] The detection of stacked chromatin in cells has evaded researchers,[6] popularizing the view that *in vivo* chromatin is flexible and heterogenous.[7] The juxtaposition between *in vitro* ordered structures and *in vivo* heterogeneity can likely be attributed to several factors, including, but not limited to: the length of DNA between nucleosomes, histone variants and post-translational modifications, the ionic strength of the intracellular environment, the DNA sequence, the presence of chromatin effectors, and active enzymatic processes.

Importantly, the differing nucleosome packing structures may indicate a division of chromatin type, most broadly separated into heterochromatin and euchromatin **(Figure 1.2A)**.[8] Heterochromatin makes up >25% of the genome,[9] has a dense physical structure, encapsulates genes from the center and edges of chromosomes, and reduces gene expression.[10] Euchromatin forms the opposing force with active gene expression and an open structure. These chromatin territories were first found by their different staining densities under the microscope.[8] Soon followed the discovery of the territorial association to gene expression. Now, high resolution sequencing methods can directly map out how the genome is compacted and expressed as a result of its biochemical profile.[11]

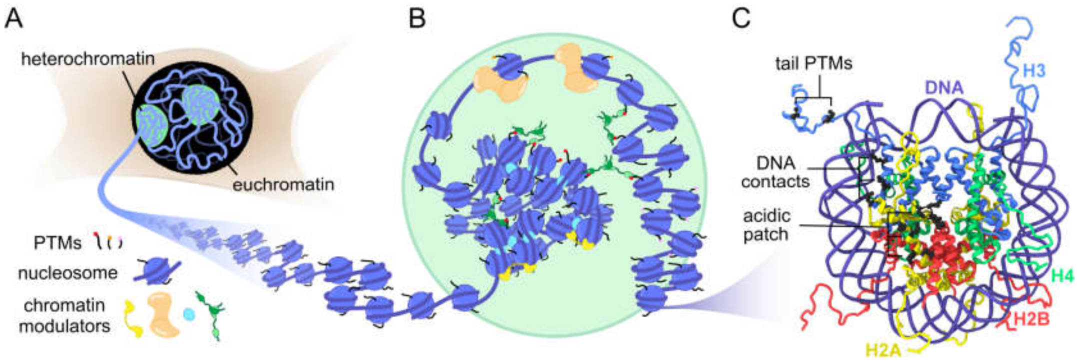


Figure 1.2. Heterochromatin organization from the nucleus to the nucleosome. (A) The nucleus generally contains two distinct territories, heterochromatin and euchromatin. Heterochromatin often shows as dense puncta. (B) The heterochromatin territory is distinguished by specific post-translational modifications (PTMs) and chromatin modulators, such as HP1. These features dictate the compaction of adjacent nucleosomes and higher order chromatin folding. (C) The nucleosome is a platform for PTMs which often target the histone tails, sites of DNA contact, and the acidic patch of H2A/H2B. Adapted from *Ackermann and Debelouchina.*[1]

Chromatin polymers occupy the entire area of the nucleus, which enables atomic level modification of the nucleosome to have micron-scale organizational.[12] In particular, the histone tails are primary targets for post-translational modifications (PTMs) and protein binding. The modification landscape of histone tails spurred the histone code hypothesis,[13] which posits that the organization and function of chromatin stems from the assortment of tail modifications which can include methylation, acetylation, phosphorylation, ubiquitylation, and many more **(Figure 1.2C)**. The common modifications in heterochromatin regions are the trimethylation of the lysine 9 in histone H3 (H3 K9me3), trimethylation of lysine 20 in histone H4 (H4 K20me3), and trimethylation of lysine 27 in histone H3 (H3 K27me3).[14] Associated with these modifications are the methyltransferases that install such modifications (SUV39H, SUV420H, etc.), and the proteins that recognize these marks, such as heterochromatin protein 1 (HP1) **(Figure 1.2B)**.

The histone tail modification H3 K9me3 was foundational to the genesis of the histone code hypothesis and is imperative to the work in this thesis. The histone code predicts that a set of protein readers can decode histone marks and signal gene regulatory events.[13] This was supported by discovery of the specific affinity of HP1 for the H3 K9me3 mark,[15] and the enrichment of H3 K9me3 in heterochromatin.[16] Histone marks may change the dynamics of the tail to adjust its accessibility to binding partners. For example, H3 K9me3 has been shown to help release the tail from DNA for better chromodomain binding.[17] Histone marks may also have their own structural effects on chromatin. H4 K20me3, a mark also associated with heterochromatin, is sufficient to condense chromatin,[18] whereas H4 acetylation marks associated with euchromatin are sufficient to open up chromatin.[19] Additionally, the code must be written and erased, which involves the concerted action of distinct methyltransferases and demethylases.[20] Combinatorial code complexity can be added through adjacent residues or incompatible marks. For example H3 serine 10 phosphorylation can prevent chromodomain binding to H3 lysine 9,[21] and H3 lysine 9 can be acetylated rather than methylated,[22] both of these cases are associated with more active gene expression.[23] Altogether there exists a system of heterochromatin construction, maintenance and disassembly.

At the heart of this thesis is determining how a heterochromatin protein such as HP1 can influence chromatin and use nucleosome modifications to create these separate domains and govern their level of gene expression.

1.2 Heterochromatin protein 1

HP1 was aptly named, being one of the first proteins discovered to be associated directly with heterochromatin territories.[24] Evidence that HP1 was critical to positional gene silencing solidified the significance of HP1 to heterochromatin.[25] HP1 is particularly enriched in constitutive heterochromatin regions marked by H3 K9me2/3. These genomic regions are regularly silenced in most cell types, bearing "selfish" genes and sequence repeats that can disrupt a healthy genome if left uncontrolled.

Beyond this early characterization, work during the past several decades has been aimed at determining how the HP1 protein family promotes gene repression. In humans, HP1 has three paralogs, α, β, and γ, each with a conserved overall structure and overlapping functions. The work in this thesis is primarily concerned with the α paralog. HP1α is a small ~25 kDa multi-domain protein, bearing a chromodomain (CD) for recognition of H3K9me2/3, and a chromoshadow domain (CSD) responsible for dimerization and binding of protein partners (**Figure 1.3)**. The CD is responsible for the primary heterochromatic localization of HP1α, using H3K9me2/3 to discriminate from the sea of differentially modified nucleosomes, while the CSD dimer functionally adds the ability to link two separate nucleosomes through space. The CSD and CD each have disordered tails and are interspersed by a flexible hinge region capable of interacting

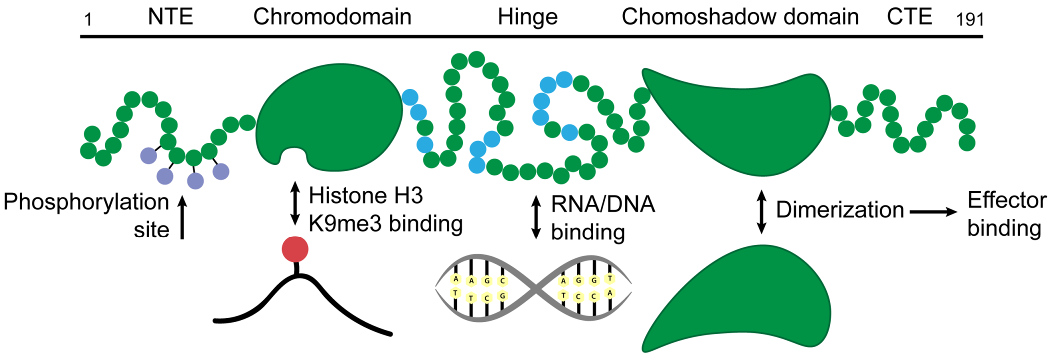


Figure 1.3. Domain map of HP1α. Known domain functions are shown. The NTE is a site of constitutive phosphorylation, the chromodomain binds H3 K9me2/3, the hinge interacts with nucleic acids, and the CSD leads to dimerization and effector binding.

with nucleic acids. The N-terminal tail of HP1α is a constitutively phosphorylated site that helps add specificity for H3 K9me2/3 binding.[26]

The first proposed role of HP1 was physical compaction of chromatin. In those studies, the mutation of the HP1 gene caused decompaction of *Drosophila* heterochromatin.[27] Then followed evidence that heterochromatin domains had more ordered, regularly spaced nucleosomes than euchromatin. This finding suggested that HP1 can function like a clamp that stacks adjacent nucleosomes.[28] This was supported by studies with a construct that contained a mutation that prevented dimerization. This construct gained diffusion mobility, was unable to compact chromatin, and no longer linked separate heterochromatin regions through space.[29] Clearly, the ability of HP1 to bind two H3 K9me2/3 marks simultaneously was critical to its function. However, original beliefs of a stable clamped structure are complicated by more recent studies showing HP1 to rapidly associate and dissociate from nucleosomes.[29c, 30] However, the H3 K9me2/3 binding does not tell the complete story. In fact, when the CSD of HP1 is mutated to prevent binding to partner proteins, HP1 dissociates faster from nucleosomes and the functions of partner proteins in heterochromatin regulation are perturbed.[29c, 30a, 31]

The CSD-CSD dimer has a large number of interacting partners that contain a distinct PXVXL binding motif (P is proline, X is any amino acid, V is valine, and L is leucine).[31a, 32] HP1 can interact with over 100 distinct proteins, including SUV39H1/2 which installs its localization mark, thus generating a feedback loop for heterochromatin maintenance.[31a, 33] Of this protein partner library, many members have important functions of their own and also disrupt heterochromatin if deleted.[34] For example, HP1 often complexes with transcription regulation factors such as zinc-finger proteins which target specific DNA sequences.[33, 35] HP1 also complexes with chromatin remodelers spanning all families (CHD3/CHD4,[36] ACF1,[31b] FACT,[37] ATRX,[38] Brg1,[39] SNF2[40]). Chromatin remodelers enzymatically slide nucleosomes to expose or hide sequences of DNA that would elicit transcription. With such a breadth of binding partners, HP1 appears to be a hub for other DNA or chromatin modifying proteins.

Finally, these interaction landscapes may all exist within the physical process of liquid-liquid phase separation (LLPS). HP1α was recently found to undergo LLPS upon phosphorylation of its N-terminal tail or upon the addition of nucleic acids. Such a discovery carries the weight of explaining many features of heterochromatin. Therefore, much of the work in this thesis was done with LLPS in mind and will be considered in detail below.

1.3 Liquid-liquid phase separation

To rewrite equilibrium, cells require the means to compartmentalize reactions in time and space, and to control the crossing of molecules from one compartment to another. This fact is most obviously highlighted by the lipid membrane, a structure shared by all organisms. The lipid membrane is a structural wall where entry is enforced by proteins embedded in the membrane. Interestingly, completely different forms of barriers can take place. Barriers can form between two liquids of different composition similarly to the way oil separates from water **(Figure 1.4)**. These barriers, formed by the process of liquid-liquid phase separation, can create membrane-less compartments that decide their molecular composition, retain liquidity throughout, adopt spherical shapes, and fuse together upon contact. Without the obvious wall of lipids, this phenomenon evaded microscopes for decades.[41] Only recently have scientists scraped the surface of how LLPS intricately organizes the interior of a cell by the phase separation properties of proteins and other biomolecules.[42] Phase separation is a thermodynamic process where the chemical potential of the mixture is minimized by creating entirely separate phases. Proteins contribute to the solvent composition and can therefore nucleate and join new phases based on their chemical properties. This ultimately leads to the popular model of cellular organization where the cell is divided into many separate liquid compartments that sculpt their composition based on molecular properties. These new phases have the potential to completely change the context of a biomolecule, controlling reaction kinetics, material properties, and macromolecular structure. Examples include substrates being prohibited from entering, compartment transformation into a gel, and protein restructuring because folding is dependent on solvent composition. The incredible result of such a process is that

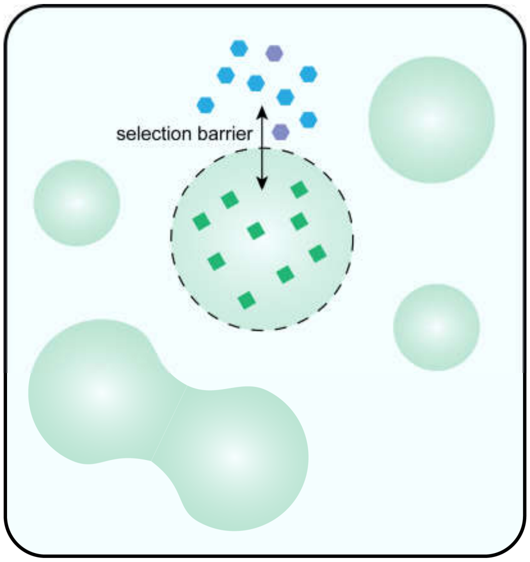


Figure 1.4. LLPS overview. LLPS generates compartments with selective barriers. These liquid compartments are spherical and fuse on contact.

biomolecules can rapidly organize to direct critical cell functions (that may never happen outside of the new phase context!) such as manufacturing proteins, signaling cell death, degrading toxins, and of course, regulating genes.[43]

There are several general properties that drive biomolecules to create and shuttle between phase separated compartments.[44] The first requirement is for weak and transient interactions with neighboring biomolecules. Such a condition is routinely possible in the cell interior where molecules regularly bump into each other.[45] However, to generate any specificity for compartment formation, the second requirement is a directional force by specific interactions. This is commonly done by short range interactions between: positively and negatively charged amino acids, positively charged amino acids and aromatic amino acids, or between two aromatic amino acids **(Figure 1.5)**.[46] Finally, these interactions must happen at multiple sites on a single molecule, a mechanism termed multivalency. This favors intrinsically disordered proteins which have no singular stable protein fold, permitting them to

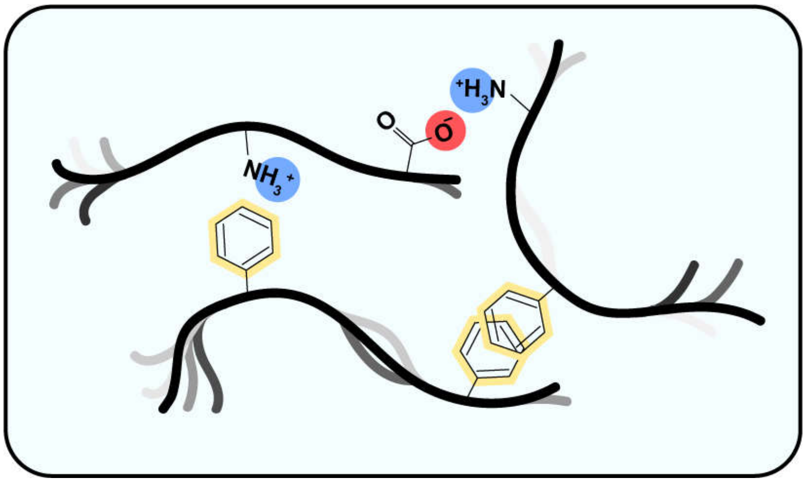


Figure 1.5. Molecular principles of LLPS. LLPS is favored by multivalent transient interactions among disordered proteins. These interactions are typically driven by electrostatics and aromatic side-chains.

contort through space and satisfy many weak interactions simultaneously (becoming more solvent-like). Once these conditions are satisfied - weak and directed interactions, multivalency, flexibility, and high concentration - a new liquid phase is formed separated from molecules in an incompatible interaction landscape.[47]

Unsurprisingly, the LLPS mechanism has become an attractive model for the organization of the nucleus. Chromatin itself,[48] and many important chromatin effectors, have displayed the tendency to phase separate,[49] including HP1α,[50] polycomb proteins,[51] histone H1,[52] DNA and histone methyltransferases,[53] and transcription factors.[48a, 54] The LLPS model is attractive for heterochromatin (and chromatin in general) because it can explain the unexplained hallmarks of heterochromatin: selective exclusion, dense condensation, responsive architecture, enzymatic precision, and epigenetic spreading.[14, 55] What remains to be seen is how the structure and dynamics of chromatin and heterochromatin proteins, like HP1, generate and maintain heterochromatin territories.

1.4 HP1 and chromatin in the LLPS context

HP1α was initially found to undergo liquid-liquid phase separation *in vitro* upon N-terminal phosphorylation or in the presence of DNA.[50b] The negatively charged phosphates or DNA were proposed to engage transiently with the positively charged hinge region, leading to phase separation conditions. These HP1α droplets transformed into gels over the course of days, suggesting that the HP1α- HP1α interactions settle into a more rigid cross-linked network. These results were supported by a study of *Drosophila* and mammalian nuclei that showed the heterochromatin domains to exhibit characteristics of phase separation like fusion and the presence of a barrier.[50a] These nuclear domains developed into less liquid-like compartments over time as HP1 and chromatin began immobilizing. In the nucleus, the biophysical properties of chromatin likely have a major influence on the material properties of phase separating proteins such as HP1.[56]

Under certain buffer conditions *in vitro,* small arrays of nucleosomes can undergo their own phase separation.[48a] The addition of a number of chromatin effectors, including HP1, further enhances this phase separation.[48b, 57] In fact, while only HP1α can phase separate on its own, both HP1β, and HP1γ can phase separate readily in the presence of nucleosome arrays with the H3 K9me3 modification.[58] The ability to link two separate H3 tails proves essential to promoting this chromatin condensation and may be central to the solidifying effect of heterochromatin.[29b, 48c, 50a, 57] These results support a model of heterochromatin formation where HP1 serves as a dynamic linker between H3 K9me3 nucleosomes to promote a new phase.[48c, 59]

A deeper look into the interaction landscape between HP1 and the nucleosome suggests that HP1 has extensive interactions with DNA and the histone cores. These

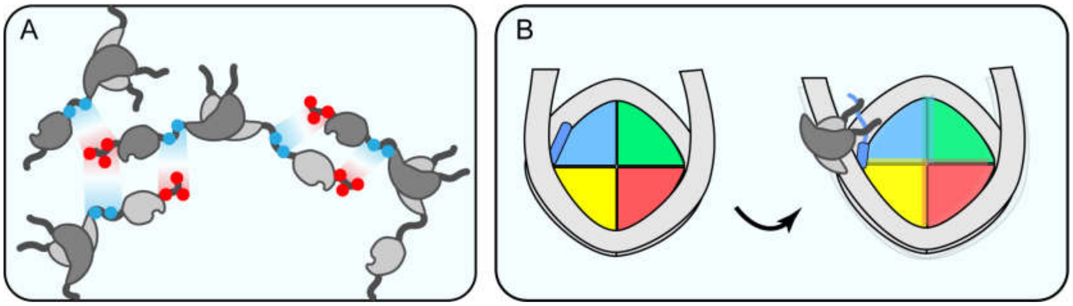


Figure 1.6. HP1α LLPS and nucleosomal interactions. (A) The phosphorylated NTE of HP1α interacts with the hinge region of neighboring HP1α dimers. Over time, these interactions reduce in dynamics and the system evolves toward a gel-like state. (B) The chromoshadow domain binds to the histone H3 αN-helix and this may require extensive histone dynamics to access the helix in the nucleosomal context.

interactions may induce chromatin condensation by loosening the DNA around nucleosomes and increasing the histone dynamics within the nucleosome core.[48b, 60] These results were shown only with the yeast homolog Swi6 which may have unique function. However, there is literature suggesting that HP1α can interact with the histone core as well.[32b, 61] For example, peptide binding studies have shown that the CSD can bind to the H3 αN-helix,[32b, 62] the last folded helix of the H3 histone core. This helix, which contains a PXVXL-like binding motif, is at the DNA entry/exit site and may become exposed upon DNA unwrapping. However, this interaction has never been shown in the context of chromatin. The proposed HP1-nucleosome core interaction may elucidate how HP1 can transmit nucleosome binding information to the formation of mesoscale gene regulatory domains.

This thesis is driven to uncover the role and mechanism of HP1α LLPS in the formation and gene regulation of heterochromatin domains. In Chapter 3, we elucidate dynamic changes of the HP1α N-terminal domain within phase separated droplets while following the droplets as they solidify into gels **(Figure 1.6A)**. We also show that the gelation effect persists in droplets which contain chromatin arrays, but to a lesser extent. In Chapter 4, we further characterize the interaction landscape between HP1 and the nucleosome, focusing on the CSD and the proposed docking site at the H3 αN-helix **(Figure 1.6B)**. Lastly, we probe how HP1-nucleosome interactions prevent the binding of transcription factors and chromatin remodelers. Together, this thesis offers insight into the phase separation and gelation mechanism of HP1α, the scope of multivalent interactions between HP1 and chromatin, and the competitive landscape of transcription factors and HP1 for chromatin binding sites.

References

[1] B. E. Ackermann, G. T. Debelouchina, *Front. Mol. Biosci.* **2021**, *8*.

[2] K. Luger, **1997**, *389*, 10.

[3] C. L. F. Woodcock, J. P. Safer, J. E. Stanchfield, *Experimental Cell Research* **1976**, *97*, 101-110.

[4] aF. Song, P. Chen, D. Sun, M. Wang, L. Dong, D. Liang, R. M. Xu, P. Zhu, G. Li, *Science* **2014**, *344*, 376-380; bS. Kilic, S. Felekyan, O. Doroshenko, I. Boichenko, M. Dimura, H. Vardanyan, L. C. Bryan, G. Arya, C. A. M. Seidel, B. Fierz, *Nature Communications* **2018**, *9*; cT. Schalch, S. Duda, D. F. Sargent, T. J. Richmond, *Nature* **2005**, *436*, 138-141; dMaria A. Ricci, C. Manzo, M. F. García-Parajo, M. Lakadamyali, Maria P. Cosma, *Cell* **2015**, *160*, 1145-1158.

[5] V. I. Risca, S. K. Denny, A. F. Straight, W. J. Greenleaf, *Nature* **2017**, *541*, 237-241.

[6] K. Maeshima, S. Ide, M. Babokhov, *Current Opinion in Cell Biology* **2019**, *58*, 95-104.

[7] aK. Maeshima, S. Tamura, J. C. Hansen, Y. Itoh, *Current Opinion in Cell Biology* **2020**, *64*, 77-89; bH. D. Ou, S. Phan, T. J. Deerinck, A. Thor, M. H. Ellisman, C. C. O'Shea, *Science* **2017**, *357*, eaag0025; cS. Cai, Y. Song, C. Chen, J. Shi, L. Gan, *Molecular Biology of the Cell* **2018**, *29*, 1652-1663.

[8] E. Heitz, *Das heterochromatin der moose*, Bornträger, **1928**.

[9] C. International Human Genome Sequencing, C. f. G. R. Whitehead Institute for Biomedical Research, E. S. Lander, L. M. Linton, B. Birren, C. Nusbaum, M. C. Zody, J. Baldwin, K. Devon, K. Dewar, M. Doyle, W. FitzHugh, R. Funke, D. Gage, K. Harris, A. Heaford, J. Howland, L. Kann, J. Lehoczky, R. LeVine, P. McEwan, K. McKernan, J. Meldrim, J. P. Mesirov, C. Miranda, W. Morris, J. Naylor, C. Raymond, M. Rosetti, R. Santos, A. Sheridan, C. Sougnez, N. Stange-Thomann, N. Stojanovic, A. Subramanian, D. Wyman, C. The Sanger, J. Rogers, J. Sulston, R. Ainscough, S. Beck, D. Bentley, J. Burton, C. Clee, N. Carter, A. Coulson, R. Deadman, P. Deloukas, A. Dunham, I. Dunham, R. Durbin, L. French, D. Grafham, S. Gregory, T. Hubbard, S. Humphray, A. Hunt, M. Jones, C. Lloyd, A. McMurray, L. Matthews, S. Mercer, S. Milne, J. C. Mullikin, A. Mungall, R. Plumb, M. Ross, R. Shownkeen, S. Sims, C. Washington University Genome Sequencing, R. H. Waterston, R. K. Wilson, L. W. Hillier, J. D. McPherson, M. A. Marra, E. R. Mardis, L. A. Fulton, A. T. Chinwalla, K. H. Pepin, W. R. Gish, S. L. Chissoe, M. C. Wendl, K. D. Delehaunty, T. L. Miner, A. Delehaunty, J. B. Kramer, L. L. Cook, R. S. Fulton, D. L. Johnson, P. J. Minx, S. W. Clifton, U. D. J. G. Institute, T. Hawkins, E. Branscomb, P. Predki, P. Richardson, S. Wenning, T. Slezak, N. Doggett, J.-F. Cheng, et al., *Nature* **2001**, *409*, 860-921.

[10] S. C. R. Elgin, S. I. S. Grewal, *Current Biology* **2003**, *13*, R895-R898.

[11] E. Lieberman-Aiden, N. L. van Berkum, L. Williams, M. Imakaev, T. Ragoczy, A. Telling, I. Amit, B. R. Lajoie, P. J. Sabo, M. O. Dorschner, R. Sandstrom, B. Bernstein, M. A. Bender, M. Groudine, A. Gnirke, J. Stamatoyannopoulos, L. A. Mirny, E. S. Lander, J. Dekker, *Science* **2009**, *326*, 289-293.

[12] J. Xu, H. Ma, J. Jin, S. Uttam, R. Fu, Y. Huang, Y. Liu, *Cell Reports* **2018**, *24*, 873-882.

[13] T. Jenuwein, *Science* **2001**, *293*, 1074-1080.

[14] A. Janssen, S. U. Colmenares, G. H. Karpen, **2018**, 24.

[15] aA. J. Bannister, P. Zegerman, J. F. Partridge, E. A. Miska, J. O. Thomas, R. C. Allshire, T. Kouzarides, **2001**, *410*, 5; bM. Lachner, D. O'Carroll, S. Rea, K. Mechtler, T. Jenuwein, *Nature* **2001**, *410*, 116-120.

[16] A. H. F. M. Peters, D. O'Carroll, H. Scherthan, K. Mechtler, S. Sauer, C. Schöfer, K. Weipoltshammer, M. Pagani, M. Lachner, A. Kohlmaier, S. Opravil, M. Doyle, M. Sibilia, T. Jenuwein, *Cell* **2001**, *107*, 323-337.

[17] A. H. Tencer, K. L. Cox, L. Di, J. B. Bridgers, J. Lyu, X. Wang, J. K. Sims, T. M. Weaver, H. F. Allen, Y. Zhang, J. Gatchalian, M. A. Darcy, M. D. Gibson, J. Ikebe, W. Li, P. A. Wade, J. J. Hayes, B. D. Strahl, H. Kono, M. G. Poirier, C. A. Musselman, T. G. Kutateladze, *Cell Reports* **2017**, *21*, 455-466.

[18] X. Lu, M. D. Simon, J. V. Chodaparambil, J. C. Hansen, K. M. Shokat, K. Luger, *Nature Structural & Molecular Biology* **2008**, *15*, 1122-1124.

[19] aB. Dorigo, T. Schalch, K. Bystricky, T. J. Richmond, *Journal of Molecular Biology* **2003**, *327*, 85-96; bP. Y. Kan, T. L. Caterino, J. J. Hayes, *Molecular and Cellular Biology* **2009**, *29*, 538-546.

[20] S. Rea, F. Eisenhaber, C. P. Ponting, C. D. Allis, T. Jenuwein, **2000**, *406*, 7.

[21] W. Fischle, B. S. Tseng, H. L. Dormann, B. M. Ueberheide, B. A. Garcia, J. Shabanowitz, D. F. Hunt, H. Funabiki, C. D. Allis, *Nature* **2005**, *438*, 1116-1122.

[22] P. Cheung, K. G. Tanner, W. L. Cheung, P. Sassone-Corsi, J. M. Denu, C. D. Allis, *Molecular Cell* **2000**, *5*, 905-915.

[23] A. Zippo, R. Serafini, M. Rocchigiani, S. Pennacchini, A. Krepelova, S. Oliviero, *Cell* **2009**, *138*, 1122-1136.

[24] T. C. James, S. C. Elgin, *Molecular and Cellular Biology* **1986**, *6*, 3862-3872.

[25] J. C. Eissenberg, T. C. James, D. M. Foster-Hartnett, T. Hartnett, V. Ngan, S. C. Elgin, *Proceedings of the National Academy of Sciences* **1990**, *87*, 9923-9927.

[26] aG. LeRoy, I. Chepelev, P. A. DiMaggio, M. A. Blanco, B. M. Zee, K. Zhao, B. A. Garcia, *Genome Biology* **2012**, *13*, R68; bG. Nishibuchi, S. Machida, A. Osakabe, H. Murakoshi, K. Hiragami-Hamada, R. Nakagawa, W. Fischle, Y. Nishimura, H. Kurumizaka, H. Tagami, J.-i. Nakayama, *Nucleic Acids Research* **2014**, *42*, 12498-12511.

[27] D. E. Cryderman, M. H. Cuaycong, S. C. R. Elgin, L. L. Wallrath, *Chromosoma* **1998**, *107*, 277-285.

[28] aF. L. Sun, M. H. Cuaycong, S. C. R. Elgin, *Molecular and Cellular Biology* **2001**, *21*, 2867-2879; bD. Canzio, E. Y. Chang, S. Shankar, K. M. Kuchenbecker, M. D. Simon, H. D. Madhani, G. J. Narlikar, B. Al-Sady, *Molecular Cell* **2011**, *41*, 67-81.

[29] aT. Cheutin, *Science* **2003**, *299*, 721-725; bK. Hiragami-Hamada, S. Soeroes, M. Nikolov, B. Wilkins, S. Kreuz, C. Chen, I. A. De La Rosa-Velázquez, H. M. Zenn, N. Kost, W. Pohl, A. Chernev, D. Schwarzer, T. Jenuwein, M. Lorincz, B. Zimmermann, P. J. Walla, H. Neumann, T. Baubec, H. Urlaub, W. Fischle, *Nature Communications* **2016**, *7*, 11310; cS. Kilic, A. L. Bachmann, L. C. Bryan, B. Fierz, *Nature Communications* **2015**, *6*.

[30] aK. Muller-Ott, F. Erdel, A. Matveeva, J. P. Mallm, A. Rademacher, M. Hahn, C. Bauer, Q. Zhang, S. Kaltofen, G. Schotta, T. Hofer, K. Rippe, *Molecular Systems Biology* **2014**, *10*, 746-746; bL. C. Bryan, D. R. Weilandt, A. L. Bachmann, S. Kilic, C. C. Lechner, P. D. Odermatt, G. E. Fantner, S. Georgeon, O. Hantschel, V. Hatzimanikatis, B. Fierz, *Nucleic Acids Research* **2017**, *45*, 10504-10517.

[31] aR.-S. Nozawa, K. Nagao, H.-T. Masuda, O. Iwasaki, T. Hirota, N. Nozaki, H. Kimura, C. Obuse, *Nature Cell Biology* **2010**, *12*, 719-727; bR. Eskeland, A. Eberharter, A. Imhof, *Molecular and Cellular Biology* **2007**, *27*, 453-465.

[32] aA. Thiru, D. Nietlispach, H. R. Mott, M. Okuwaki, D. Lyon, P. R. Nielsen, M. Hirshberg, A. Verreault, N. V. Murzina, E. D. Laue, *The EMBO Journal* **2004**, *23*, 489-499; bY. Liu, S. Qin, M. Lei, W. Tempel, Y. Zhang, P. Loppnau, Y. Li, J. Min, *Journal of Biological Chemistry* **2017**, *292*, 5655-5664.

[33] C. Rosnoblet, J. Vandamme, P. Völkel, P. O. Angrand, *Biochem Biophys Res Commun* **2011**, *413*, 206-211.

[34] J. M. Swenson, S. U. Colmenares, A. R. Strom, S. V. Costes, G. H. Karpen, *eLife* **2016**, *5*.

[35] S. P. Sripathy, J. Stevens, D. C. Schultz, *Molecular and cellular biology* **2006**, *26*, 8623-8638.

[36] aV. Ostapcuk, F. Mohn, S. H. Carl, A. Basters, D. Hess, V. Iesmantavicius, L. Lampersberger, M. Flemr, A. Pandey, N. H. Thomä, J. Betschinger, M. Bühler, *Nature* **2018**, *557*, 739-743; bH. Hoffmeister, A. Fuchs, F. Erdel, S. Pinz, R. Gröbner-Ferreira, A. Bruckmann, R. Deutzmann, U. Schwartz, R. Maldonado, C. Huber, A.-S. Dendorfer, K. Rippe, G. Längst, *Nucleic Acids Research* **2017**, *45*, 10534-10554.

[37] S. H. Kwon, L. Florens, S. K. Swanson, M. P. Washburn, S. M. Abmayr, J. L. Workman, *Genes & Development* **2010**, *24*, 2133-2145.

[38] aA. V. Emelyanov, A. Y. Konev, E. Vershilova, D. V. Fyodorov, *Journal of Biological Chemistry* **2010**, *285*, 15027-15037; bD. Sadic, K. Schmidt, S. Groh, I. Kondofersky, J. Ellwart, C. Fuchs, F. J. Theis, G. Schotta, *EMBO reports* **2015**, *16*, 836-850; cN. Kourmouli, Y.-m. Sun, S. van der Sar, P. B. Singh, J. P. Brown, *Biochemical and Biophysical Research Communications* **2005**, *337*, 901-907; dA. A. Alekseyenko, A. A. Gorchakov, B. M. Zee, S. M. Fuchs, P. V. Kharchenko, M. I. Kuroda, *Genes & Development* **2014**, *28*, 1445-1460.

[39] A. L. Nielsen, *The EMBO Journal* **2002**, *21*, 5797-5806.

[40] aSamuel P. Rowbotham, L. Barki, A. Neves-Costa, F. Santos, W. Dean, N. Hawkes, P. Choudhary, W. R. Will, J. Webster, D. Oxley, Catherine M. Green, P. Varga-Weisz, Jacqueline E. Mermoud, *Molecular Cell* **2011**, *42*, 285-296; bR. Santoro, J. Li, I. Grummt, *Nature Genetics* **2002**, *32*, 393-396.

[41] aH. Walter, D. E. Brooks, *FEBS Letters* **1995**, *361*, 135-139; bC. Liu, A. Lomakin, G. M. Thurston, D. Hayden, A. Pande, J. Pande, O. Ogun, N. Asherie, G. B. Benedek, *The Journal of Physical Chemistry* **1995**, *99*, 454-461; cC. P. Brangwynne, C. R. Eckmann, D. S. Courson, A. Rybarska, C. Hoege, J. Gharakhani, F. Julicher, A. A. Hyman, *Science* **2009**, *324*, 1729-1732.

[42] D. M. Mitrea, R. W. Kriwacki, *Cell Communication and Signaling* **2016**, *14*.

[43] S. F. Banani, H. O. Lee, A. A. Hyman, M. K. Rosen, *Nature Reviews Molecular Cell Biology* **2017**, *18*, 285-298.

[44] Clifford P. Brangwynne, P. Tompa, Rohit V. Pappu, *Nature Physics* **2015**, *11*, 899-904.

[45] R. D. Cohen, G. J. Pielak, *Protein Science* **2017**, *26*, 403-413.

[46] J. Wang, J.-M. Choi, A. S. Holehouse, H. O. Lee, X. Zhang, M. Jahnel, S. Maharana, R. Lemaitre, A. Pozniakovsky, D. Drechsel, I. Poser, R. V. Pappu, S. Alberti, A. A. Hyman, *Cell* **2018**, *174*, 688-699.e616.

[47] J. A. Ditlev, L. B. Case, M. K. Rosen, *Journal of Molecular Biology* **2018**.

[48] aB. A. Gibson, L. K. Doolittle, M. W. G. Schneider, L. E. Jensen, N. Gamarra, L. Henry, D. W. Gerlich, S. Redding, M. K. Rosen, *Cell* **2019**, *179*, 470-484.e421; bS. Sanulli, M. J. Trnka, V. Dharmarajan, R. W. Tibble, B. D. Pascal, A. L. Burlingame, P. R. Griffin, J. D. Gross, G. J. Narlikar, *Nature* **2019**; cH. Strickfaden, T. O. Tolsma,

A. Sharma, D. A. Underhill, J. C. Hansen, M. J. Hendzel, *Cell* **2020**, *183*, 1772-1784.e1713.

[49] A. R. Strom, C. P. Brangwynne, *Journal of Cell Science* **2019**, *132*, jcs235093.

[50] aA. R. Strom, A. V. Emelyanov, M. Mir, D. V. Fyodorov, X. Darzacq, G. H. Karpen, *Nature* **2017**, *547*, 241-245; bA. G. Larson, D. Elnatan, M. M. Keenen, M. J. Trnka, J. B. Johnston, A. L. Burlingame, D. A. Agard, S. Redding, G. J. Narlikar, *Nature* **2017**, *547*, 236-240; cL. Wang, Y. Gao, X. Zheng, C. Liu, S. Dong, R. Li, G. Zhang, Y. Wei, H. Qu, Y. Li, C. D. Allis, G. Li, H. Li, P. Li, *Molecular Cell* **2019**, *76*, 646-659.e646.

[51] aA. J. Plys, C. P. Davis, J. Kim, G. Rizki, M. M. Keenen, S. K. Marr, R. E. Kingston, *Genes & Development* **2019**, *33*, 799-813; bR. Tatavosian, S. Kent, K. Brown, T. Yao, H. N. Duc, T. N. Huynh, C. Y. Zhen, B. Ma, H. Wang, X. Ren, *Journal of Biological Chemistry* **2019**, *294*, 1451-1463.

[52] A. L. Turner, M. Watson, O. G. Wilkins, L. Cato, A. Travers, J. O. Thomas, K. Stott, *Proceedings of the National Academy of Sciences* **2018**, 201805943.

[53] aL. Wang, M. Hu, M.-Q. Zuo, J. Zhao, D. Wu, L. Huang, Y. Wen, Y. Li, P. Chen, X. Bao, M.-Q. Dong, G. Li, P. Li, *Cell Research* **2020**; bH. Strickfaden, K. Missiaen, M. J. Hendzel, D. A. Underhill, *bioRxiv* **2019**.

[54] aA. Boija, I. A. Klein, B. R. Sabari, A. Dall'Agnese, E. L. Coffey, A. V. Zamudio, C. H. Li, K. Shrinivas, J. C. Manteiga, N. M. Hannett, B. J. Abraham, L. K. Afeyan, Y. E. Guo, J. K. Rimel, C. B. Fant, J. Schuijers, T. I. Lee, D. J. Taatjes, R. A. Young, *Cell* **2018**, *175*, 1842-1855.e1816; bB. R. Sabari, A. Dall'Agnese, A. Boija, I. A. Klein, E. L. Coffey, K. Shrinivas, B. J. Abraham, N. M. Hannett, A. V. Zamudio, J. C. Manteiga, C. H. Li, Y. E. Guo, D. S. Day, J. Schuijers, E. Vasile, S. Malik, D. Hnisz, T. I. Lee, I. I. Cisse, R. G. Roeder, P. A. Sharp, A. K. Chakraborty, R. A. Young, *Science* **2018**, *361*, eaar3958; cW.-K. Cho, J.-H. Spille, M. Hecht, C. Lee, C. Li, V. Grube, I. I. Cisse, *Science* **2018**, *361*, 412-415.

[55] P. B. Talbert, S. Henikoff, *Nature Reviews Genetics* **2006**, *7*, 793-803.

[56] F. Erdel, K. Rippe, *Biophys J* **2018**, *114*, 2262-2270.

[57] B. E. Ackermann, G. T. Debelouchina, *Angew. Chem. Int. Ed.* **2019**, *58*, 6300-6305.

[58] L. Wang, Y. Gao, X. Zheng, C. Liu, S. Dong, R. Li, G. Zhang, Y. Wei, H. Qu, Y. Li, C. D. Allis, G. Li, H. Li, P. Li, *Mol Cell* **2019**, *76*, 646-659.e646.

[59] aF. Erdel, A. Rademacher, R. Vlijm, J. Tünnermann, L. Frank, R. Weinmann, E. Schweigert, K. Yserentant, J. Hummert, C. Bauer, S. Schumacher, A. Al Alwash, C. Normand, D.-P. Herten, J. Engelhardt, K. Rippe, *Molecular Cell* **2020**; bA. R. Strom, R. J. Biggs, E. J. Banigan, X. Wang, K. Chiu, C. Herman, J. Collado, F. Yue, J. C. Ritland Politz, L. J. Tait, D. Scalzo, A. Telling, M. Groudine, C. P. Brangwynne, J. F. Marko, A. D. Stephens, *eLife* **2021**, *10*, e63972; cS. Machida, Y. Takizawa, M. Ishimaru, Y. Sugita, S. Sekine, J.-i. Nakayama, M. Wolf, H. Kurumizaka, *Molecular Cell* **2018**, *69*, 385-397.e388.

[60] S. E. Farr, E. J. Woods, J. A. Joseph, A. Garaizar, R. Collepardo-Guevara, *Nature Communications* **2021**, *12*, 2883.

[61] aM. Lavigne, R. Eskeland, S. Azebi, V. Saint-André, S. M. Jang, E. Batsché, H.-Y. Fan, R. E. Kingston, A. Imhof, C. Muchardt, *PLoS Genetics* **2009**, *5*, e1000769; bS. M. Jang, S. Azebi, G. Soubigou, C. Muchardt, *EMBO reports* **2014**, *15*, 686-694.

[62] A. N. Richart, C. I. W. Brunner, K. Stott, N. V. Murzina, J. O. Thomas, *Journal of Biological Chemistry* **2012**, *287*, 18730-18737.

Chapter II

Nuclear Magnetic Resonance Introduction

2.1 Fundamentals of nuclear magnetic resonance

Nuclear magnetic resonance (NMR) is a spectroscopic technique that relies on a property of atomic nuclei called spin. Most atoms have non-zero spin, an inherent property of the nucleus which is affected by local magnetic environments. The nuclear spin has no effect on the chemistry of molecules, making it a useful target to probe while leaving the sample unperturbed. For this reason, hospitals can use magnetic resonance imaging (MRI) to probe the nuclear spins of the hydrogen atoms from water directly in humans. NMR spectroscopists, on the other hand, are often interested in more than just water. NMR spectroscopists aim to uncover the atomic structure and dynamics of molecules, opting for smaller samples, stronger magnets, and a variety of atoms.

Fortunately for the structural biologist, there are stable NMR-active (half-integer spin) isotopes for three prominent atoms found in proteins, ^{1}H, ^{13}C, and ^{15}N. While spin is an abstract concept to us in the macroscopic world, it can be classically described as an object spinning with angular momentum and a magnetic moment susceptible to magnetic fields **(Figure 2.1A)**.[1] Some population of the spins will align in the direction of an applied external magnetic field B_0, scaling with the strength of B_0 and the magnetism of the nucleus given by its gyromagnetic ratio (γ) **(Table 2.1)**. The

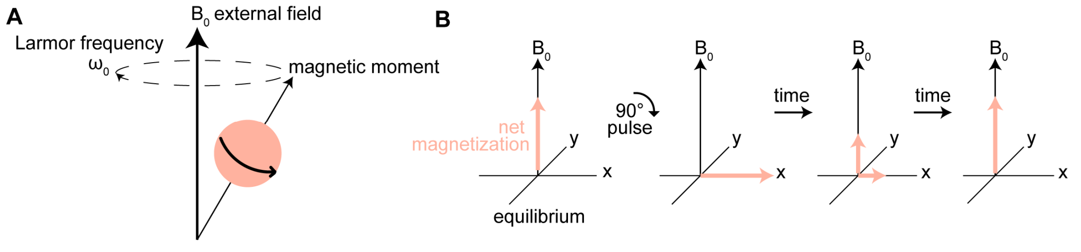
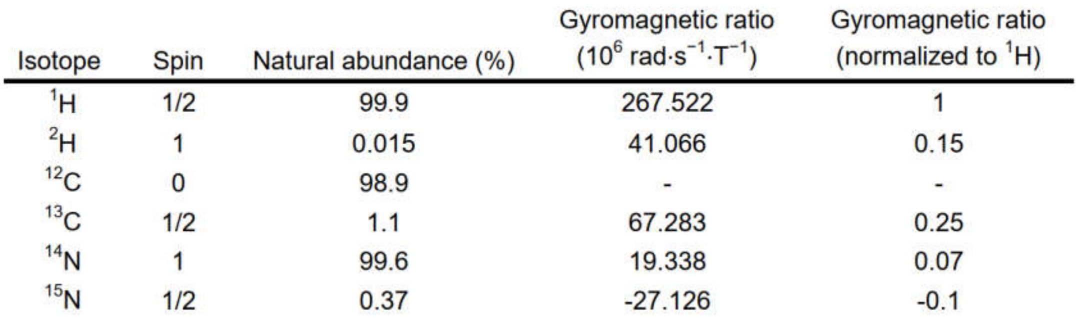


Figure 2.1. Nuclear spin precession and net magnetization. (A) The nuclear spin precesses around an applied external field at the Larmor frequency, which scales with the intrinsic gyromagnetic ratio. (B) NMR signal is generated by rotating the aligned net magnetization and recording the return to equilibrium over time.

Table 2.1. Nuclear isotope properties. NMR-active and inactive nuclei relevant to this thesis are shown. Nuclear properties can be found on the website www.webelements.com.



Isotope	Spin	Natural abundance (%)	Gyromagnetic ratio (10^6 rad·s^{-1}·T^{-1})	Gyromagnetic ratio (normalized to ^{1}H)
^{1}H	1/2	99.9	267.522	1
^{2}H	1	0.015	41.066	0.15
^{12}C	0	98.9	-	-
^{13}C	1/2	1.1	67.283	0.25
^{14}N	1	99.6	19.338	0.07
^{15}N	1/2	0.37	-27.126	-0.1

population difference of proton spins using a common 14.1 Tesla magnet would be 0.0042%, estimated by $\hbar\gamma B_0/2kT$. Here, $\hbar$ is Planck's constant, k is the Boltzmann constant, and T is temperature. The relationship between the gyromagnetic ratio and the given magnetic field can be described as the Larmor frequency ($\omega = -\gamma B_0$), which is also the precession frequency of the spin around the axis of the external field. It is common to refer to NMR magnets as the Larmor frequency of the ^{1}H nucleus, for example a 14.1 Tesla magnet is referred to as a 600 MHz magnet (the ω_o of ^{1}H at this magnetic field).

To spectroscopically probe nuclear spins, NMR spectroscopists apply radiofrequency pulses (B_1) that are in resonance with the Larmor frequency but at an angle perpendicular to B_0 **(Figure 2.1B)**. This breaks equilibrium and rotates all nuclear spins by 90 degrees. Once the pulse is removed, the nuclear spins can precess back towards the B_0 axis to return to equilibrium. This return process creates an oscillating electric current which is captured by the NMR coil and generates the NMR signal. Because every atom in a molecule is situated in a different chemical environment, each atom experiences a slightly different magnetic field and resonates at a slightly shifted Larmor frequency. This idea is ingrained in the axes of NMR spectra which are called chemical shifts. The precession of an ensemble of spins is tracked over time as a free induction decay. This time domain signal is Fourier transformed into the frequency domain which assists in visualizing the unique resonances from distinct atoms (**Figure 2.2)**. The main observables of NMR are the resonance frequency, which reports on the chemical environment; the lineshape, which reports on the environment inhomogeneity surrounding the atom; and the intensity, which reports on quantity and/or dynamics of the atom.

The assignment of each resonance to a specific atom within the molecule of interest is critical to analysis. This can be straightforward for small molecules, but becomes unfeasible for the thousands of atoms in proteins. Therefore, protein NMR spectroscopy often uses multidimensional experiments that spread the signals into two or more dimensions **(Figure 2.3A)**. Experiments that correlate two proximal atoms can open up the available spectral space, reducing spectral overlap. Conducting a two-dimensional correlation experiment involves four steps: preparation, evolution, mixing, and detection **(Figure 2.3B)**. Preparation is the initial stage of pulses rotating the nuclei

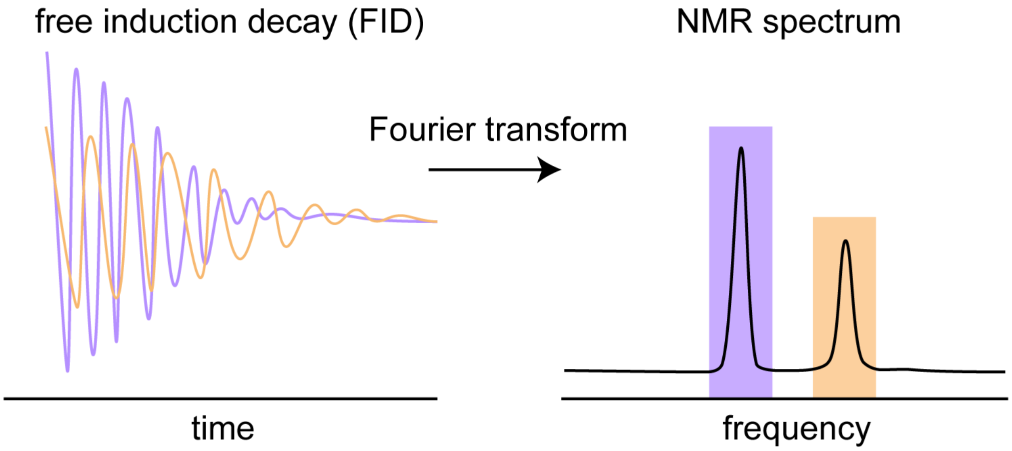


Figure 2.2. Fourier transform NMR spectroscopy. The free induction decay which detects the oscillation of nuclei over time is Fourier transformed into the frequency domain for interpretation. The intensity and length of the FID is translated into the height and width of peaks in the NMR spectrum.

of interest, evolution is an incremented time block that is used to create the indirect dimension (t_1), mixing is a period of signal transfer to nearby atoms, and detection is the recorded free induction decay of the remaining signal (t_2). Fourier transformation of each time domain gives a two-dimensional spectrum correlating nucleus 1 (from t_1) and nucleus 2 (from t_2). The axes show the chemical shift, a value in units of parts per million (ppm) that is normalized to the Larmor frequency of the chosen nucleus. This makes for facile comparison between magnets of different strengths.

Beyond two dimensions, three-dimensional experiments are commonly used for proteins to correlate a longer chain of adjacent atoms. Several overlapping 3D experiments allow "walking" through the backbone of a protein and assignment of each resonance to an amino acid atom based on the known sequence. Resonance assignments and the ability to correlate between them is fundamental to protein NMR, allowing for spatial and dynamic information to be extracted.

2.2 Solution NMR spectroscopy of proteins

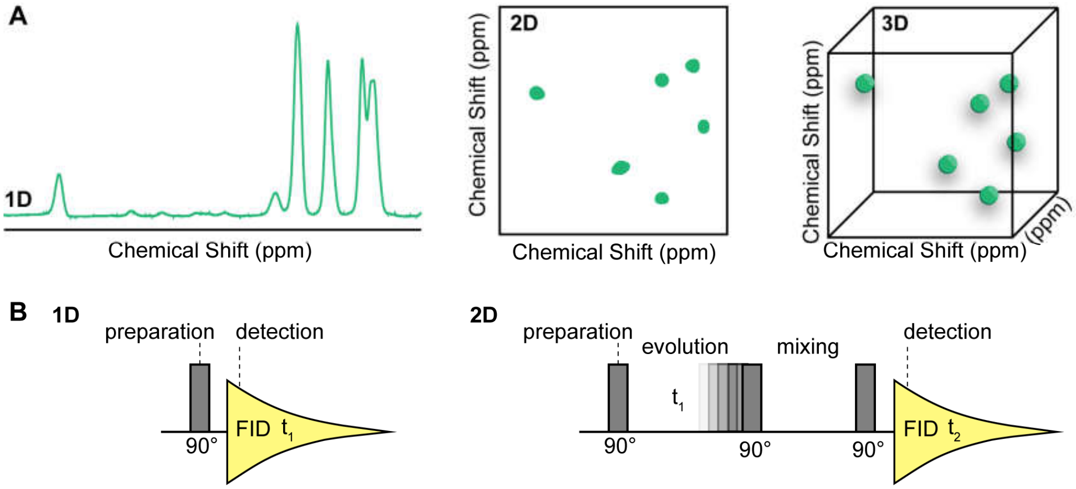


Figure 2.3. Multi-dimensional NMR spectroscopy. (A) An example of spectral space gained through expanding from 1D to 3D experiments. (B) Generic pulse experiments for generating 1D and 2D NMR spectra.

Solution state NMR spectroscopy has been the "bread and butter" of protein NMR. Since first recording a protein spectrum[2] and later solving the first structure of a protein,[3] solution NMR is now an essential tool for structural biologists to gain insight into the structure, motion and interactions of proteins at an atomic level.

Standard procedure for solution NMR starts with the recombinant production of the protein of interest from isotopically enriched growth media ($^{13}C/^{15}N$) to ensure ~99% NMR-active isotope labeling of the protein. Isotope enrichment ensures that there are more active nuclei for improved sensitivity and that the active nuclei are adjacent to each other for correlation experiments. The enriched protein is purified and concentrated in a suitable buffer to the micromolar-millimolar range. From here, the sample is simply

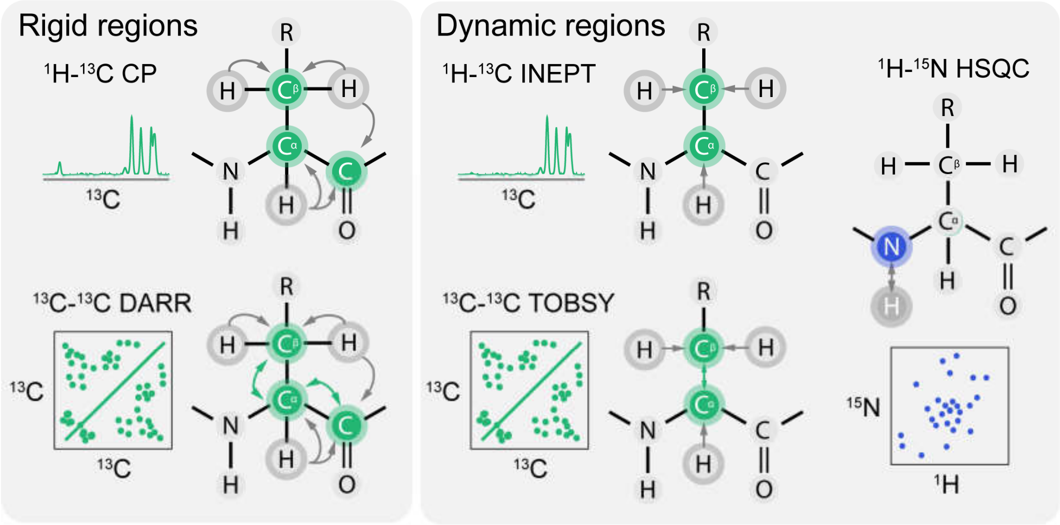


Figure 2.4. Amino acid atom correlations for solution and solid-state experiments. (Left) Solid-state experiments used to detect rigid protein segments. A 1D ^{1}H-^{13}C CP transfers polarization from protons through space to nearby carbon atoms. A 2D ^{13}C-^{13}C DARR begins with the ^{1}H-^{13}C CP transfer, then allows carbon atoms to interact through space. (Right) Solid-state and solution-state experiments to detect proteins regions with rapid motions. A 1D ^{1}H-^{13}C INEPT transfers spin polarization to bonded carbon atoms. A 2D ^{13}C-^{13}C TOBSY begins with the ^{1}H-^{13}C INEPT transfer, then allows carbon atoms to transfer information through covalent bonds. A 2D ^{1}H-^{15}N HSQC correlating amide proteins and their bonded backbone nitrogen.

inserted into the NMR magnet and regulated at a temperature that ensures protein stability for hours or days.

One of the most popular experiments for solution NMR is the Heteronuclear Single Quantum Coherence (HSQC) experiment **(Figure 2.4)**. HSQC is used to correlate ^{1}H-^{15}N or ^{1}H-^{13}C which provide highly resolved signals for the amino acid backbone amide or side chain carbons, respectively. Because each amino acid only has a single backbone nitrogen, each amino acid in the protein matches a single ^{1}H-^{15}N cross-peak. This is a valuable way to monitor each protein residue during dynamic and structural changes from biomolecular interactions, mutations, and modifications.

One caveat of solution NMR is its reliance on the rapid tumbling of molecules in solution. Proteins beyond 45 kDa or even small proteins engaging in weak interactions, begin to rotate, or "tumble", too slowly to be measured effectively. At any time point, the proteins in the sample have a distribution of orientations with respect to the magnetic field. Each of these orientations has a unique shielding effect on the atom of interest. Molecular tumbling assists by averaging out any orientation differences among protein molecules. If fast enough, tumbling will average all orientations, maintaining coherence between all proteins and providing a sharp isotropic signal. If tumbling is too slow, the coherence is not maintained, and the signal is broadened and often undetectable. This is especially problematic for 2D or 3D experiments which require coherence for long periods of time.

To alleviate the tumbling dilemma; the temperature of the sample can be increased, the protein can be partially deuterated (^{2}H) to reduce the broadening effect of the large coupled ^{1}H network, higher magnetic fields can be applied, and smaller truncations of the protein can be studied in isolation. Truncation works effectively for studying proteins like HP1α which have folded domains connected by disordered regions. Truncations, however, cannot reproduce holistic events such as LLPS. Additionally, LLPS is a phenomenon that emerges from abundant weak interactions that reduce molecular tumbling, therefore different NMR approaches altogether may be necessary for phase separated systems. One suitable option is solid-state NMR spectroscopy.

2.3 Solid-state NMR spectroscopy of proteins

Solid-state NMR spectroscopy has been developed to study the molecular details of powders, crystals and sediments. In the absence of tumbling, anisotropic interactions that carry spatial information present themselves in the NMR spectra. One anisotropic interaction is the chemical shift anisotropy (CSA), where the chemical shift of a nucleus is subject to the local electron environment which depends on orientation with respect to the external field. Dipolar couplings also impose anisotropy on a nucleus. Dipolar coupling is the through-space interaction between magnetic dipoles and is dependent on the angle between nuclear vectors with respect to the magnetic field. Altogether, these

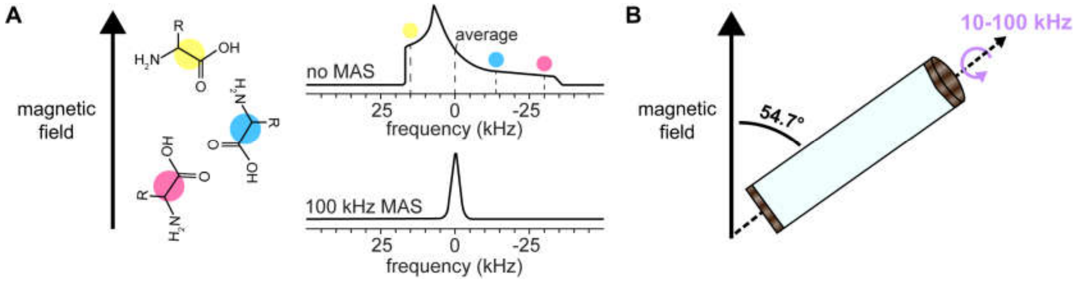


Figure 2.5. Magic angle spinning effects on anisotropy. (A) Molecules with different orientations with respect to the external magnetic field result in a wide range of chemical shifts for the same atom. MAS can average out all orientations. (B) The experimental procedure for MAS requires a sample rotor that is angled at 54.7°. The rotor spinning is typically propelled by nitrogen gas.

anisotropic interactions can spread the chemical shift range of a nucleus up to 100 kHz depending on the nucleus (**Figure 2.5A**). Without tumbling, these signal broadening interactions must be removed in a different way. Magic angle spinning (MAS) NMR resolves this by rapidly spinning the sample at the magic angle (54.74°) with respect to the external magnetic field (**Figure 2.5B**). The magic angle represents the orientation in which anisotropic interactions that depend on the term, $3\cos^2\theta - 1$, can be reduced to zero (where θ is the angle between nuclei). Spinning on this axis at a rate greater than the frequency impact of the anisotropic interactions (~10-100 kHz) will average all orientations during the NMR experiment and provide a sharp isotropic peak. This method is necessary to obtain high-resolution spectra in the solid-state comparable to those of solution NMR spectroscopy.

Importantly, these same anisotropic interactions can also be harnessed to gain structural information in solid-state NMR experiments. For example, the dipolar coupling of protons with carbon or nitrogen is utilized in the fundamental solid-state experiment cross-polarization (CP) **(Figure 2.4)**. CP transfers magnetization through space and actually requires that molecules remain rigid to retain their orientation throughout the NMR experiment. A CP experiment typically generates magnetization on the higher gyromagnetic ratio nucleus (^{1}H) and transfers this magnetization to less sensitive nuclei such as ^{13}C or ^{15}N. After generating polarization through CP, the common 2D fingerprinting experiment for MAS NMR is the Dipolar Assisted Rotational Resonance (DARR) experiment **(Figure 2.4)**.[4] This experiment correlates ^{13}C-^{13}C pairs through space, reaching atoms several bonds away. The DARR fingerprint shows entire amino acid side chains and can therefore sense more structural changes than the amide backbone alone. Similar to solution NMR, these resonances can be assigned by several overlapping experiments which "walk" sequentially through the protein.

Solid-state NMR can also borrow experiments from solution NMR, such as the Insensitive Nuclei Enhanced by Polarization Transfer (INEPT) experiment which is the building block of the HSQC experiment described above **(Figure 2.4)**. INEPT transfers spin information through-bond and therefore requires rapid dynamics. To probe the mobile dynamic regime, polarization can be generated through INEPT and correlated between ^{1}H-^{13}C or ^{1}H-^{15}N similar to the HSQC, or ^{13}C-^{13}C using the Total Through Bond Correlation Spectroscopy (TOBSY) experiment. Using these two experiments, INEPT and CP, alongside each other, allows for the protein of interest to be detected in two separate dynamic regimes. Solid-state MAS NMR is then uniquely suited to simultaneously study large complexes and rigid systems that have highly flexible domains. This pairs well with chromatin, a large polymer with rigid histone cores and dynamic histone tails,[5] as well as gel-like materials which require rigid cross-links as well as dynamic bridges.[6] Altogether, large biomolecules that are beyond the limitations of solution NMR can be studied in their entirety by solid-state NMR, in contrast to other popular structural techniques such as x-ray crystallography and cryogenic electron microscopy which typically miss dynamic regions.

2.4 Dynamic nuclear polarization signal enhancement

Solid-state NMR possesses the power to atomically characterize any molecule or material, however, suffers from low sensitivity. While isotopic enrichment certainly addresses the issue, the inability to produce or fit enough of the expensive sample into

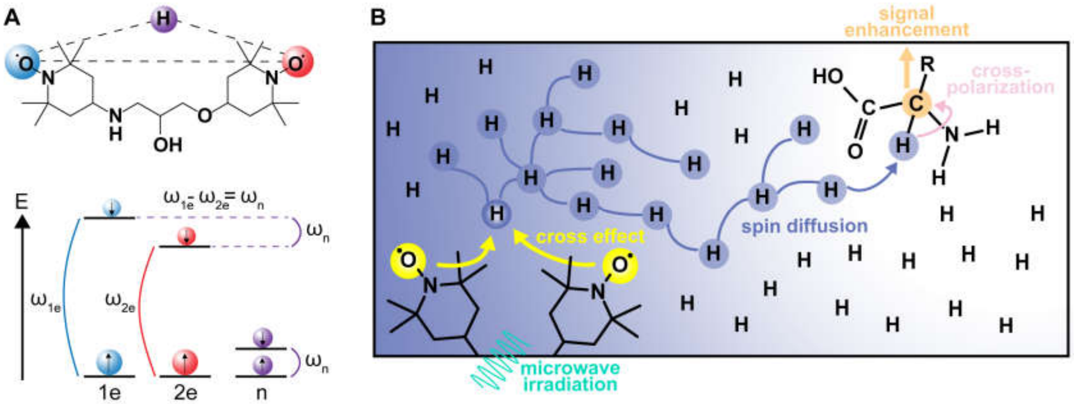


Figure 2.6. Cross-effect mechanism for DNP sensitivity enhancement. (A) The cross-effect mechanism requires coupling between two unpaired electron spins and a nuclear spin. This is commonly achieved through the addition of a biradical PA (TOTAPOL pictured here). To satisfy the cross-effect condition, the difference in nutation frequencies between the two electron spins must be equal to the nutation frequency of the nuclear spin. (B) Microwave irradiation saturates the electron energy transition to induce the cross-effect. Hyperpolarized nuclei spread polarization through the sample via spin diffusion. Polarization can finally be transferred to the nuclei of interest through cross-polarization.

the NMR rotor can extend NMR experiments to last weeks to months. If we expand beyond the realm of nuclear spin, we find that the electron spin can be harnessed to enhance NMR signal, an idea akin to the cross-polarization experiment which transfers the greater polarization of the proton to less sensitive nuclei such as ^{13}C or ^{15}N.[7] This process is called dynamic nuclear polarization (DNP). DNP uses the spin of an unpaired electron which has a gyromagnetic ratio 660 times that of the proton. If fully utilized, experimental NMR times can be reduced up to 660 x 660 (435,600) fold, since signal-to-noise scales with the square root of the number of scans ($S/N \propto \sqrt{\text{scans}}$) during signal acquisition. In this thesis we focus on a form of DNP typically used for biological samples which performs MAS NMR at high magnetic field and cryogenic temperatures.

DNP MAS NMR currently achieves optimal signal enhancement using a polarization transfer mechanism called the cross effect.[8] The cross effect requires a cooperation between three spins coupled through space (two electron and one nuclear)

(Figure 2.6A). To satisfy the conditions of polarization transfer, the difference in nutation frequency between the electron spins must be equal to the nuclear Larmor frequency (ω_{1e} - ω_{2e} = ω_n) **(Figure 2.6A)**. This condition allows for mixing between the electron and nuclear spin energies such that the nuclear spin can achieve the polarization level of the electrons when the electrons are excited by microwave irradiation. From here, the coupled proton can distribute polarization throughout the sample by ^{1}H-^{1}H spin diffusion **(Figure 2.6B)**.[9] Spin diffusion transfers magnetization to nearby spins with similar Larmor frequencies. In biological samples, H_2O molecules serve as a powerful matrix for this diffusion of polarization. For a typical aqueous matrix, 5-10 % protonation is optimal for spin diffusion and achieved through the judicial use of D_2O. Spin diffusion eventually reaches the molecule of interest and can be transfered over to ^{13}C or ^{15}N nuclei through the standard cross-polarization experiment.

DNP requires additional practical considerations that include instrumentation and sample preparation. The first consideration is the source of unpaired electron spins, these are typically synthesized molecules called polarization agents (PAs) that are added to the sample in millimolar quantities. To efficiently enable the cross effect, PAs typically link together two unpaired electrons.[8] One of the earliest examples of this is TOTAPOL **(Figure 2.8A).[10]** PAs are actively being developed because of their massive influence on signal enhancement. Some engineerable characteristics of PAs include solubility, orientation between radicals, radical stability, size, and bonus features such as fluorescence or chemically reactive groups. PA design is discussed further in section 2.6.

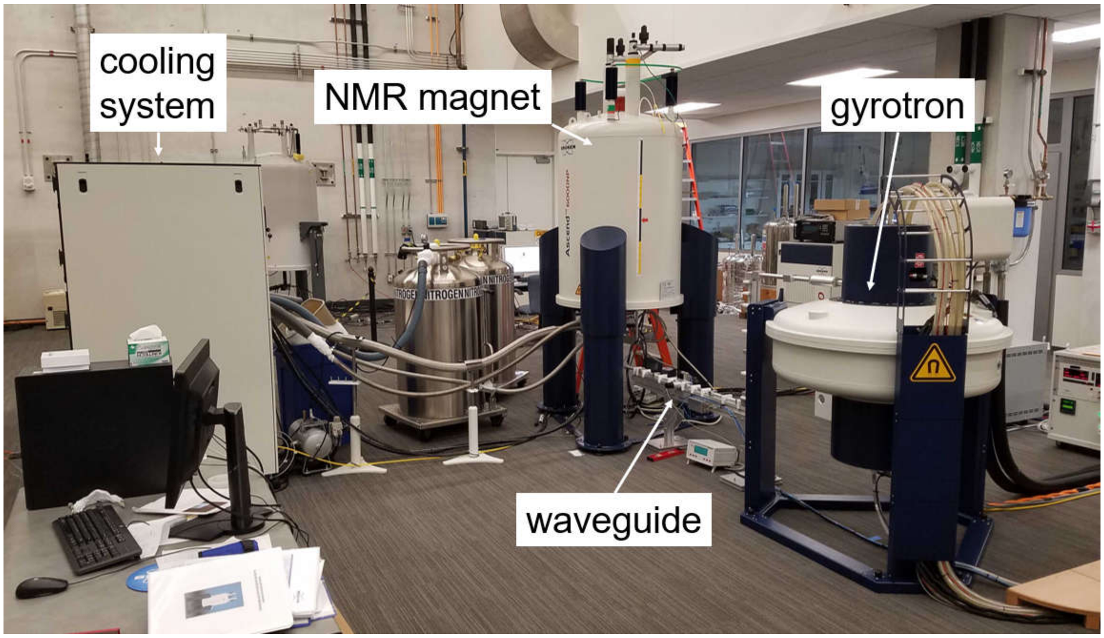


Figure 2.7. DNP instrument layout. A NMR magnet is shown in the center connected to a gyrotron by a waveguide, and a cooling system by a cryogenic transfer line.

A major instrumental need for DNP is a source of microwave irradiation to saturate the electron polarization. For high-field setups like ours, this is typically done using a gyrotron, a separate magnet alongside the NMR magnet that is used to generate high power microwaves on resonance with the electron frequency (395 GHz to scale with our 600 MHz NMR magnet) **(Figure 2.7)**.[11] The microwave travels through a waveguide to irradiate the sample inside the NMR magnet continuously throughout the experiment.

Another requirement for DNP is low temperature operation. Electron spins return to thermal equilibrium much faster than nuclei, interfering with the cross effect by preventing sufficient time for polarization transfer. Lowering the sample temperature slows down this relaxation process and is currently necessary to achieve large DNP enhancements. A routine way to cool is through a liquid N_2 cooling system that produces gas flows down to 100 K and is accompanied by a cryo-friendly DNP NMR probe under constant vacuum. A bonus feature of low temperature is the generous temperature dependence of spin polarization. For our experimental setup at 14.1 T and 100 K, the electron spin can achieve 9.5 % polarization (~3.2 % at room temperature), compared to the .014% proton polarization. Another consequence of low temperature occurs during sample freezing. Ice formation can destroy the structure of macromolecules and can cause PA aggregation that diminishes signal enhancements. Therefore, glassing agents like d8-glycerol are added (up to 60% of the volume) to prevent ice formation during slow cooling. The greatest consequence of low temperatures is the loss of spectral resolution. Inhomogeneous broadening of the NMR signal derives from freezing out many molecular conformations. At room temperature, these different conformations continually interchange to result in a sharp averaged peak. Low temperature traps all conformational variation which adds valuable structural information to the spectra but also complicates analysis by overlapping peaks.

DNP MAS NMR is a promising technique, but still requires improvements to be routinely used for structural biology. The approach to gain structural information by DNP must be designed carefully to work around the significant peak broadening. Protein structural biology applications have found the most success *in vitro* with amyloids, viral capsids, membrane proteins, chromatin, proteins undergoing disorder to order transitions, and active sites of proteins.[12] These systems succeed for several reasons. Rigid structures, like amyloids, maintain highly resolved peaks at low temperature because there is little conformational heterogeneity. DNP provides enough signal enhancement for amyloids to determine the distance between protein monomers and therefore build a model for fibril structure. Proteins undergoing major structural rearrangement also succeed with DNP because secondary structure (alpha helix, beta strand, and disorder) have a predictable effect on the chemical shift of an atom. Lastly, proteins with unique chemical shifts simplify the identification of structural change. For example, an active site that binds a small molecule cofactor or an active site that transitions through a distinctive functional group. In these cases, DNP may enhance NMR signals to detect low population states during enzyme activity. DNP offers such large sensitivity improvements that the field is excited to perform these experiments on molecules naturally residing within cells.

2.5 DNP NMR in cells

Conducting NMR on proteins within the cell has long captured the imagination of spectroscopists. Methodology to do so was originally developed for solution NMR.[13] However, it was quickly realized that NMR signal is deteriorated by the many non-specific biomolecular interactions experienced within the cell.[14] DNP MAS NMR,

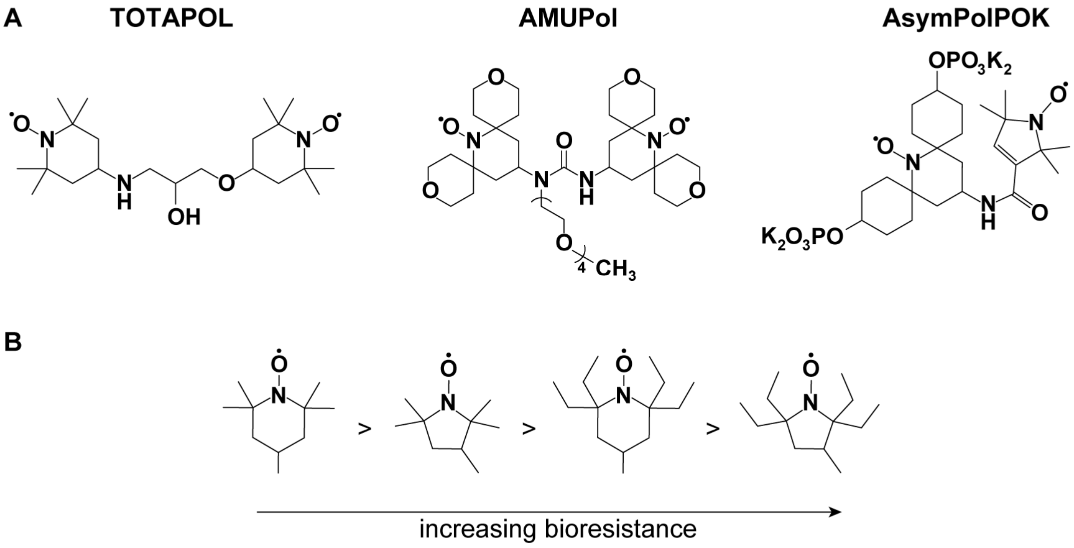


Figure 2.8. Polarization agent design. (A) Popular commercially available polarization agents. (B) Nitroxide modifications for increasing resistance to reduction.

on the other hand, both disregards tumbling rates and enhances the weak signal of intracellular proteins (ng-μg quantities per one million cells).

The majority of studies implementing DNP for native environments have reported on bacterial membrane proteins. The focus of this thesis looks to the future application of DNP NMR to soluble proteins within mammalian cells. Only a few studies have successfully performed DNP in cells, from which we glean insight for development. Most have tracked the presence of a biomolecule of interest using 1D experiments or characterized the general cellular biomass using 2D experiments.[15] Narasimhan et al. paved the way to using the full multi-dimensional NMR toolkit on the intracellular protein ubiquitin.[16] They used electroporation to deliver recombinant ^{13}C-^{15}N ubiquitin into HeLa cells. The cells were cryoprotected with glycerol and incubated with the PA AMUPol **(Figure 2.8A)**. They were able to achieve signal enhancements ranging from 35 to 130, which produced enough signal to unambiguously assign one-third of the ubiquitin residues.

Recent efforts have focused on maintaining cellular integrity during the DNP freezing process. The two main factors at play are the glassing agent used (DMSO or glycerol) and the rate of freezing (slow or rapid).[15c, 17] Based on these studies, as well as in our hands, rapid freezing drastically reduces cell viability. Further, the 60% glycerol ideal for *in vitro* DNP is likely lethal to mammalian cells. While the evidence of low glycerol and slow freezing conditions that maintain cell integrity is encouraging, the bulk of work now needed is in describing biologically relevant consequences in the cellular environment. On one hand, this requires clever experimental design and protein selection to be able to identify spectral changes and their relation to structure. On the other hand, the protein of interest must have significant signal enhancement by DNP and be detectable beyond the abundant ^{13}C and ^{15}N atoms present in the cellular milieu. We suggest that developments in PAs may assist in both of these demands.

2.6 Polarization agent design

Since the initial creation of the biradical, BTnE,[8] two nitroxides stitched together became the standard for generating the cross effect. From here, TOTAPOL became popular for its improved water solubility.[10] Soon after, AMUPol was generating signal enhancements four times greater than those of TOTAPOL.[18] Because of its commercial availability and reliable performance, AMUPol is still the most commonly used PA. However, as DNP studies continue to test new ground, the PAs must adapt as well. For example, PAs are designed for general enhancement improvements (AsymPolPOK) **(Figure 2.8A)**,[19] performance at high magnetic fields (SNAPol-1),[20] or highly protonated environments (cAsympolPOK).[21] For DNP in cells, however, PA design must consider redox resistance and localization.

The cell interior is a reducing environment, in part to protect against oxidative species. The nitroxide radicals of PAs which are stable in water, may be become rapidly reduced to hydroxylamine upon cell entry. A PA like TOTAPOL becomes ineffective in bacterial cells,[22] and AMUPol only excels in mammalian cells if added in great excess.[23] Fortunately, there are design strategies to increase the resistance to reduction. This includes switching the 6-membered ring nitroxide to a 5-membered ring, or adding ethyl groups adjacent to the nitroxide rather than methyl groups present on TOTAPOL **(Figure 2.8B)**.[24] These changes focus on shielding the radical from solvent exposure but may also interrupt cross effect properties. Future PAs must be fine-tuned to balance bioresistance and enhancement.

Bioresistance becomes particularly important for PA-targeting strategies which promise to localize to the protein of interest within the cell.[25] Some ways to do this are through cysteine chemistry,[26] a ligand linked PA,[27] or bio-orthogonal tetrazine chemistry.[28] These strategies rely on a stoichiometric ratio of PA to protein of interest, therefore each PA that gets reduced depletes the signal of its bound protein. The future of PA-targeting must develop PAs that are efficient at the low concentrations of intracellular proteins. Targeting chemistry that proceeds faster than nitroxide reduction is also needed. In this thesis we develop and characterize two novel PAs; TTz designed for versatility and highly specific targeting in biological settings (Chapter 5), and POPAPOL designed for improved bioresistance (Chapter 6).

References

[1] M. J. Duer, *Solid-State NMR Spectroscopy Principles and Applications* **2001**, 1-72.
[2] M. Saunders, A. Wishnia, J. G. Kirkwood, *Journal of the American Chemical Society* **1957**, *79*, 3289-3290.
[3] M. P. Williamson, T. F. Havel, K. Wüthrich, *Journal of Molecular Biology* **1985**, *182*, 295-315.
[4] K. Takegoshi, S. Nakamura, T. Terao, *Chemical Physics Letters* **2001**, *344*, 631-637.
[5] M. Gao, P. S. Nadaud, M. W. Bernier, J. A. North, P. C. Hammel, M. G. Poirier, C. P. Jaroniec, *Journal of the American Chemical Society* **2013**, *135*, 15278-15281.
[6] aS. B. Kennedy, E. R. deAzevedo, W. A. Petka, T. P. Russell, D. A. Tirrell, M. Hong, *Macromolecules* **2001**, *34*, 8675-8685; bC. Ader, S. Frey, W. Maas, H. B. Schmidt, D. Gorlich, M. Baldus, *Proceedings of the National Academy of Sciences* **2010**, *107*, 6281-6285.
[7] A. W. Overhauser, *Physical Review* **1953**, *92*, 411-415.
[8] K.-N. Hu, H.-h. Yu, T. M. Swager, R. G. Griffin, *Journal of the American Chemical Society* **2004**, *126*, 10844-10845.
[9] P. C. A. van der Wel, K.-N. Hu, J. Lewandowski, R. G. Griffin, *Journal of the American Chemical Society* **2006**, *128*, 10840-10846.
[10] C. Song, K.-N. Hu, C.-G. Joo, T. M. Swager, R. G. Griffin, *Journal of the American Chemical Society* **2006**, *128*, 11385-11390.

[11] M. Rosay, L. Tometich, S. Pawsey, R. Bader, R. Schauwecker, M. Blank, P. M. Borchard, S. R. Cauffman, K. L. Felch, R. T. Weber, R. J. Temkin, R. G. Griffin, W. E. Maas, **2010**, 13.

[12] aB. Uluca, T. Viennet, D. Petrović, H. Shaykhalishahi, F. Weirich, A. Gönülalan, B. Strodel, M. Etzkorn, W. Hoyer, H. Heise, *Biophysical Journal* **2018**, *114*, 1614-1623; bR. Gupta, H. Zhang, M. Lu, G. Hou, M. Caporini, M. Rosay, W. Maas, J. Struppe, J. Ahn, I.-J. L. Byeon, H. Oschkinat, K. Jaudzems, E. Barbet-Massin, L. Emsley, G. Pintacuda, A. Lesage, A. M. Gronenborn, T. Polenova, *The Journal of Physical Chemistry B* **2019**, *123*, 5048-5058; cG. T. Debelouchina, M. J. Bayro, A. W. Fitzpatrick, V. Ladizhansky, M. T. Colvin, M. A. Caporini, C. P. Jaroniec, V. S. Bajaj, M. Rosay, C. E. MacPhee, M. Vendruscolo, W. E. Maas, C. M. Dobson, R. G. Griffin, *Journal of the American Chemical Society* **2013**, *135*, 19237-19247; dM. Kaplan, S. Narasimhan, C. de Heus, D. Mance, S. van Doorn, K. Houben, D. Popov-Čeleketić, R. Damman, E. A. Katrukha, P. Jain, W. J. C. Geerts, A. J. R. Heck, G. E. Folkers, L. C. Kapitein, S. Lemeer, P. M. P. van Bergen en Henegouwen, M. Baldus, *Cell* **2016**, *167*, 1241-1251.e1211; eKendra K. Frederick, Vladimir K. Michaelis, B. Corzilius, T.-C. Ong, Angela C. Jacavone, Robert G. Griffin, S. Lindquist, *Cell* **2015**, *163*, 620-628; fJ. Maciejko, J. Kaur, J. Becker-Baldus, C. Glaubitz, *Proceedings of the National Academy of Sciences* **2019**, 201817665; gS. Lange, W. T. Franks, N. Rajagopalan, K. Döring, M. A. Geiger, A. Linden, B.-J. van Rossum, G. Kramer, B. Bukau, H. Oschkinat, *Science Advances* **2016**, *2*, e1600379; hQ. Z. Ni, T. V. Can, E. Daviso, M. Belenky, R. G. Griffin, J. Herzfeld, *Journal of the American Chemical Society* **2018**, *140*, 4085-4091; iP. Fricke, D. Mance, V. Chevelkov, K. Giller, S. Becker, M. Baldus, A. Lange, *Journal of Biomolecular NMR* **2016**, *65*, 121-126; jN. Elathram, B. E. Ackermann, G. T. Debelouchina, *Journal of Magnetic Resonance Open* **2022**, *10-11*, 100057.

[13] Z. Serber, A. T. Keatinge-Clay, R. Ledwidge, A. E. Kelly, S. M. Miller, V. Dötsch, *Journal of the American Chemical Society* **2001**, *123*, 2446-2447.

[14] P. B. Crowley, E. Chow, T. Papkovskaia, *ChemBioChem* **2011**, *12*, 1043-1048.

[15] aA. Bertarello, P. Berruyer, M. Artelsmair, C. S. Elmore, S. Heydarkhan-Hagvall, M. Schade, E. Chiarparin, S. Schantz, L. Emsley, *Journal of the American Chemical Society* **2022**, jacs.1c12442; bS. A. Overall, L. E. Price, B. J. Albert, C. Gao, N. Alaniva, P. T. Judge, E. L. Sesti, P. A. Wender, G. B. Kyei, A. B. Barnes, *International Journal of Molecular Sciences* **2020**, *21*, 4649; cJ. Schlagnitweit, S. Friebe Sandoz, A. Jaworski, I. Guzzetti, F. Aussenac, R. J. Carbajo, E. Chiarparin, A. J. Pell, K. Petzold, *ChemBioChem* **2019**, *20*, 2474-2478; dR. Ghosh, Y. Xiao, J. Kragelj, K. K. Frederick, *Journal of the American Chemical Society* **2021**, jacs.1c06680.

[16] S. Narasimhan, S. Scherpe, A. Lucini Paioni, J. van der Zwan, G. E. Folkers, H. Ovaa, M. Baldus, *Angew Chem Int Ed Engl* **2019**, *58*, 12969-12973.

[17] aS. A. Overall, A. B. Barnes, *Front. Mol. Biosci.* **2021**, *8*, 743829; bY. Xiao, R. Ghosh, K. K. Frederick, *Front. Mol. Biosci.* **2022**, *8*, 789478.

[18] C. Sauvée, M. Rosay, G. Casano, F. Aussenac, R. T. Weber, O. Ouari, P. Tordo, *Angewandte Chemie International Edition* **2013**, *52*, 10858-10861.

[19] F. Mentink-Vigier, I. Marin-Montesinos, A. P. Jagtap, T. Halbritter, J. van Tol, S. Hediger, D. Lee, S. T. Sigurdsson, G. De Paëpe, *Journal of the American Chemical Society* **2018**, *140*, 11013-11019.

[20] X. Cai, A. L. Paioni, A. Adler, R. Yao, W. Zhang, D. Beriashvili, A. Safeer, A. Gurinov, A. Rockenbauer, Y. Song, M. Baldus, Y. Liu, *Chemistry – A European Journal* **2021**, chem.202102253.

[21] R. Harrabi, T. Halbritter, F. Aussenac, O. Dakhlaoui, J. van Tol, K. Damodaran, D. Lee, S. Paul, S. Hediger, F. Mentink-Vigier, S. Sigurdsson, G. De Paepe, *Angew. Chem. Int. Ed.* **2022**, anie.202114103.

[22] K. M. McCoy, R. Rogawski, O. Stovicek, A. E. McDermott, *J Magn Reson* **2019**, *303*, 115-120.

[23] R. Ghosh, R. Dumarieh, Y. Xiao, K. K. Frederick, *Journal of Magnetic Resonance* **2022**, 107150.

[24] A. P. Jagtap, I. Krstic, N. C. Kunjir, R. Hänsel, T. F. Prisner, S. T. Sigurdsson, *Free Radical Research* **2015**, *49*, 78-85.

[25] D. Gauto, O. Dakhlaoui, I. Marin-Montesinos, S. Hediger, G. De Paepe, *Chem Sci* **2021**, *12*, 6223-6237.

[26] T. Viennet, A. Viegas, A. Kuepper, S. Arens, V. Gelev, O. Petrov, T. N. Grossmann, H. Heise, M. Etzkorn, *Angew. Chem. Int. Ed.* **2016**, *55*, 10746-10750.

[27] R. Rogawski, I. V. Sergeyev, Y. Li, M. F. Ottaviani, V. Cornish, A. E. McDermott, *The Journal of Physical Chemistry B* **2017**, *121*, 1169-1175.

[28] B. J. Lim, B. E. Ackermann, G. T. Debelouchina, *ChemBioChem* **2020**, *21*, 1315-1319.

Chapter III

Heterochromatin Protein HP1α Gelation Dynamics Revealed by Solid-State NMR Spectroscopy

NMR Spectroscopy

International Edition: DOI: 10.1002/anie.201901141
German Edition: DOI: 10.1002/ange.201901141

Heterochromatin Protein HP1α Gelation Dynamics Revealed by Solid-State NMR Spectroscopy

*Bryce E. Ackermann and Galia T. Debelouchina**

Abstract: *Heterochromatin protein 1α (HP1α) undergoes liquid–liquid phase separation (LLPS) and forms liquid droplets and gels in vitro, properties that also appear to be central to its biological function in heterochromatin compaction and regulation. Here we use solid-state NMR spectroscopy to track the conformational dynamics of phosphorylated HP1α during its transformation from the liquid to the gel state. Using experiments designed to probe distinct dynamic modes, we identify regions with varying mobilities within HP1α molecules and show that specific serine residues uniquely contribute to gel formation. The addition of chromatin disturbs the gelation process while preserving the conformational dynamics within individual bulk HP1α molecules. Our study provides a glimpse into the dynamic architecture of dense HP1α phases and showcases the potential of solid-state NMR to detect an elusive biophysical regime of phase separating biomolecules.*

Proteins and nucleic acids have evolved collective properties that enable the spatial and temporal regulation of cellular functions without the need of lipid membranes.[1] Such phenomena are often mediated by a process called liquid–liquid phase separation (LLPS) and have been observed in a diverse set of biological events, including the stress response, gene regulation, signaling, and transport.[2–5] LLPS results in the formation of droplets where molecules form a dense liquid phase but can still diffuse freely. Sometimes, the droplets mature into less mobile gel or solid-like states where a stronger interaction network limits molecular diffusion and reorientation.[4,6–8] For example, immobile condensates have been identified under physiological conditions in heterochromatin, nucleoli, and stress granules.[3,4,7,9–11] While reduced condensate dynamics can have profound implications for physiological function, aging, and disease, the molecular basis of these transformations has largely remained elusive.

Here, we focus our attention on the gel transition of HP1α, a highly conserved protein that mediates gene silencing, genome stability, and chromatin compaction.[12,13] HP1α is composed of two folded domains, the chromodomain (CD) that binds methylated histone H3K9, and the chromoshadow domain (CSD) that enables the formation of the stable 44 kDa HP1α dimer (Figure 1a).[14,15] These domains interrupt three disordered regions enriched in charged amino acids: the N-terminal segment (NTE), the hinge region, and the C-terminal tail (CTE). For many years, it had been known that HP1α in cells exhibits differential dynamics ranging from fast to immobile, and recent studies propose that this is due to LLPS behavior.[3,7,16,17] For example, HP1α can form liquid droplets and gels in vitro upon N-terminal phosphorylation or the addition of DNA.[7] HP1α foci with properties consistent with LLPS behavior have also been observed at heterochromatin domains in *Drosophila* and mammalian cells.[3] These studies suggest that HP1α forms distinct dilute, dense liquid and gel-like phases within the nucleus to partition heterochromatin from the transcriptional machinery by either phase incompatibility or a physical sieving mechanism.[18]

Biophysical treatments of LLPS proteins, such as HP1α, suggest that the compaction state, interaction strength, and multivalency of the biomolecule can shape the collective material properties.[19–21] Within this framework, however, specific molecular interactions and dynamics have been challenging to describe experimentally due to the dense and amorphous nature of biological condensates.[22] Some details have emerged from solution nuclear magnetic resonance (NMR) spectroscopy experiments of the dense liquid phase that point to the importance of intrinsically disordered protein regions and their transient interactions with other LLPS components.[22–24] Solution NMR spectroscopy, however, cannot capture the molecular picture of biomolecules in the gel state where the increased viscosity and cross-linking interactions reduce molecular tumbling and cause severe line-broadening in the NMR spectra. To circumvent this problem, here we monitor the gelation of HP1α by magic angle spinning (MAS) solid-state NMR spectroscopy. With this methodology, line-broadening interactions, such as the chemical shift anisotropy, are averaged out by mechanical rotation at the magic angle, allowing the acquisition of resolved spectroscopic data independent of molecular size. Information rich interactions, such as the dipolar and scalar nuclear couplings, can then be reintroduced by selective pulse sequences to allow the characterization of secondary and tertiary protein structure, protein–protein interactions, and the description of dynamic and rigid regions of the molecule. Already successfully applied to biological hydrogels and fibrils,[25–29] MAS NMR spectroscopy can be a powerful technique for studying the molecular driving forces of biomolecular LLPS and gelation. In the context of HP1α condensates, this approach allowed us to observe the mobile and rigid regions of HP1α molecules at residue specific resolution and to identify specific sites that participate in the transformation from a dense liquid to a gel state.

Previous studies have indicated that the LLPS behavior of HP1α depends on the phosphorylation of several serine

---

[*] B. E. Ackermann, Prof. G. T. Debelouchina
Department of Chemistry and Biochemistry, University of California, San Diego, 9500 Gilman Dr., La Jolla, CA 92093 (USA)
E-mail: gdebelouchina@ucsd.edu

Supporting information and the ORCID identification number(s) for the author(s) of this article can be found under: https://doi.org/10.1002/anie.201901141.

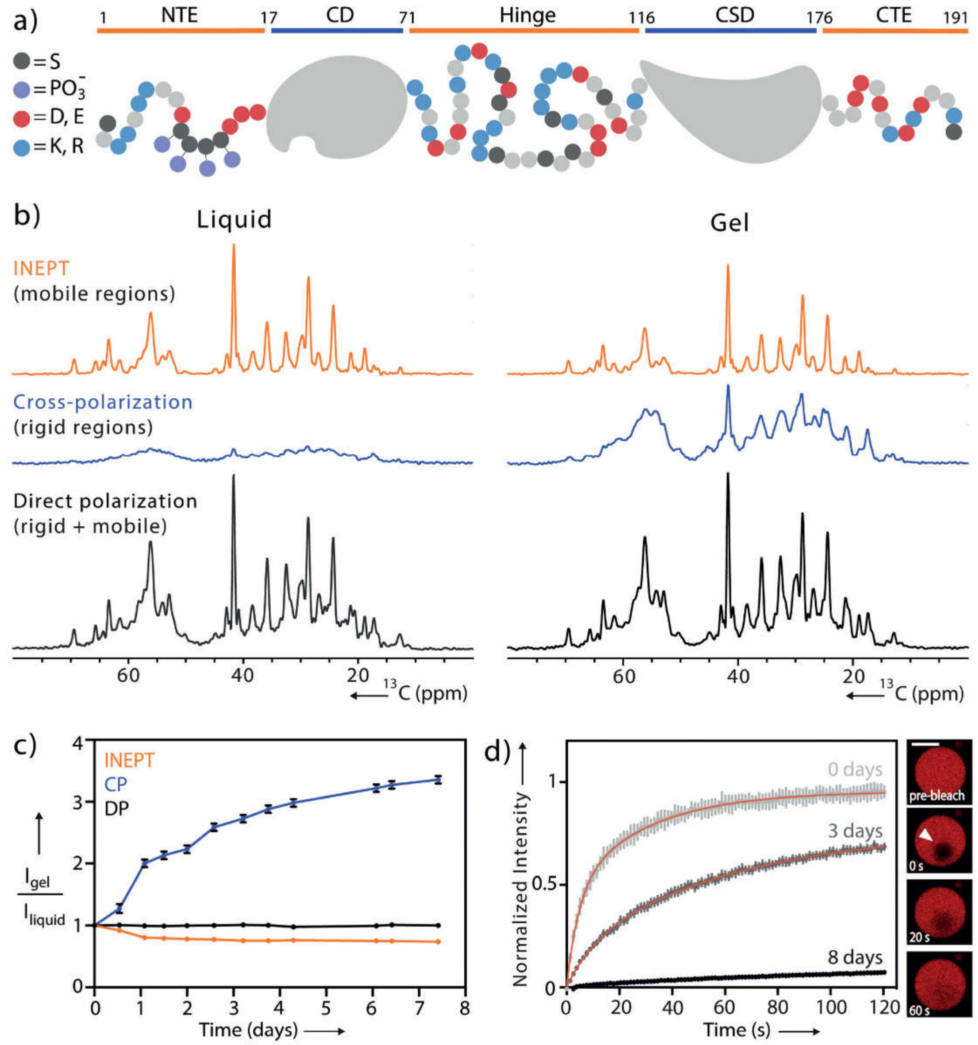


Figure 1. pHP1α dynamics monitored during gelation. a) Domain map of monomeric pHP1α. Negatively charged residues are colored in red, positively charged residues are in blue, serine residues are in dark gray, and phosphorylation sites are in purple. b) 1D ^{13}C INEPT, CP, and DP spectra acquired during the initial dense liquid state and the final gel state. c) Integrated signal of the aliphatic regions of the spectra collected every 12 hours. Data are normalized to the starting point of each respective experiment. For error analysis, see the Supporting Information. d) FRAP recovery curves for pHP1α collected immediately upon droplet formation, 3 days, and 8 days later. Red lines represent exponential fits. Inset: FRAP of a single pHP1α droplet at day 3, scale bar 20 μm. White arrow indicates bleach site.

residues in the N-terminal region of the protein (Figure 1 a).[7] Following published protocols,[7,30] we obtained high yields (50–80 mg) of pure, phosphorylated HP1α (pHP1α) per liter of ^{13}C, ^{15}N enriched *Escherichia coli* cell culture (Supporting Information, Figure S1). Upon phase separation of pHP1α, we coalesced the dense pHP1α condensates into a single phase to be used in NMR spectroscopy experiments (Supporting Information, Figure S2). As expected, the solution NMR spectra of pHP1α condensates lost all resolved information leaving only a congested region in the middle of the spectrum presumably corresponding to the disordered segments of the protein (Supporting Information, Figure S3).

Proceeding with MAS NMR spectroscopy, we performed three simple one-dimensional (1D) experiments that nevertheless provide rich information about the overall dynamics in the condensate (Figure 1 b).[31] First, the scalar-based 1D INEPT transfer experiment selects for highly mobile sites that have long transverse relaxation times similar to those detected under solution NMR conditions. Second, rigid regions that exhibit strong dipole–dipole couplings can be selected with 1D cross-polarization (CP) experiments. And finally, the ^{13}C direct polarization (DP) experiment can, in principle, be used to detect all the ^{13}C nuclei in the sample independent of dynamics. At the start of our experiments, the ^{1}H–^{13}C CP signal was remarkably lower than the ^{1}H–^{13}C INEPT signal (Figure 1 b). Comparing the intensities of these spectra with the DP experiment, it also became clear that CP and INEPT together do not account for all the ^{13}C signals in the sample. Thus, in the dense liquid droplet state there are pHP1α molecules or regions of each individual protein that exhibit motion in an intermediate dynamic range rendering them invisible in both scalar-based and dipolar-based MAS NMR spectroscopy experiments, similar to the behavior of other complex biological assemblies.[31]

Expecting that pHP1α condensates will equilibrate into gels given sufficient time,[7] we continuously performed MAS and monitored the intensities of the CP and INEPT spectra at 12 hour intervals over the course of a week (Figure 1 c and Supporting Information, Figure S4). Reaching a gel state after 7 days (Supporting Information, Figure S2), the total CP signal increased 3.3 fold, the INEPT signal decreased 1.35 fold, and the ^{13}C DP signal remained unchanged. We interpret the dramatic change in total CP intensity to arise from an expanding cross-linking network during gelation that restricts the motion of individual pHP1α dimers and shifts their overall dynamics to a slower timescale where dipolar-based polarization transfer mechanisms are more efficient. As an independent measure of this process, we also performed concurrent fluorescence recovery after photobleaching (FRAP) experiments using Cy3-labeled HP1α (Figure 1 d). In the initial liquid droplet state, the recovery of fluorescence intensity occurs on a timescale of several seconds, consistent with observations in other dense liquid systems.[16,22] Over the course of a week, the fluorescence recovery slowed down, eventually reaching a flat line. Reminiscent of the observations made at heterochromatin foci, these results align well with our NMR spectroscopy time course and provide further evidence that the motion of individual molecules within the gel state is severely limited.

Next, we turned our attention to the INEPT signals, which remained strong even after gelation was complete. The superior resolution in these spectra allowed a more careful analysis of the intensity changes for individual peaks. For example, in the Ser/Thr Cα/Cβ region certain peaks (for example, at 65.8 ppm) decreased steadily, while other peaks (for example, at 64.4 ppm) lost intensity much more slowly (Figure 2 a,b and Supporting Information, Figure S5). To obtain further insight, we proceeded to record two-dimensional (2D) ^{13}C–^{13}C correlation experiments (Figure 2 c and Supporting Information, Figure S6). In particular, INEPT can be coupled with TOBSY[32] (total through-bond correlation spectroscopy) mixing that relies on the isotropic scalar couplings to transfer polarization between adjacent nuclei and to report on highly mobile residues. A comparison of the INEPT-TOBSY spectra of the dense liquid and gel state revealed several Ser Cα/Cβ cross-peaks that remained strong and one Ser cross-peak that lost intensity over time. The downfield shift of the Cβ atom of the latter Ser compared to random coil values is consistent with phosphorylation. Thus, it appears likely that phosphorylated residues in the N-terminal tail experience changes in dynamics during gelation distinct from all other residues (Supporting Information, Figure S5).

To complement the INEPT-TOBSY 2D correlation experiment, we also recorded a 2D CP-DARR[33] (dipolar-assisted rotational resonance) correlation spectrum, which enables transfer between dipolar-coupled, and hence more rigid, ^{13}C spin systems (Figures 2 c and 3 a). A visual inspection of the complete DARR and INEPT-TOBSY spectra illustrates the dramatic difference in content and linewidths for the two sets of targeted residues (Figure 3 a). The INEPT-TOBSY spectrum shows narrow cross-peaks located at typical random coil chemical shifts that could easily be assigned to specific amino acid types. We leveraged the distinct sequence composition of the structured and unstructured regions of HP1α to match the assigned amino types with those enriched in the disordered regions (Figure 3 b). We note the cross-peak for an Ile residue, potentially originating from the hinge, the Ala and Thr correlations consistent with the N- and C-terminal tails, and the multiple Ser cross-peaks consistent with all three disordered regions. On the other hand, the DARR spectrum shows a greater diversity of chemical shifts, indicative of the sequence variety in the CD and CSD. The relatively broad Cα–Cβ region may also indicate the presence of intermediate motions that interfere with decoupling during acquisition and that contribute to the linewidth.[34] Nevertheless, many individually resolved peaks are present in the spectrum suggesting that the addition of other spectral dimensions (for example, ^{15}N) will enable sequence assignments and the interrogation of HP1α gel interactions in more detail. Note that these regions are completely intractable in solution NMR spectroscopy experiments of the dense liquid or gel state.

Finally, we wondered how the addition of relevant components, such as chromatin, will affect the dynamics of the gelation process. We therefore prepared chromatin arrays containing 12 nucleosomes with a lysine methylation mimic at position 9 of histone H3 (Supporting Information, Figure S7). These arrays reproduce the physiologically relevant state of the nuclear heterochromatin environment where HP1α binds and undergoes LLPS.[3,7,16,17] We added the equivalent of 200 μM nucleosomes to freshly prepared dense liquid droplets and repeated the experiments described above. To our surprise, the addition of chromatin significantly slowed down the rigidification process as evidenced by the smaller increase in overall CP signal (1.5 fold) and the retention of mobility in FRAP experiments (Supporting Information, Figure S8). On the other hand, the content of INEPT experiments was preserved, albeit with a smaller change in the dynamics of the phosphorylated Ser peak.

Previous models of HP1α condensation suggest that the pHP1α dimer adopts a more open state that facilitates interactions with neighboring dimers, resulting in a dense

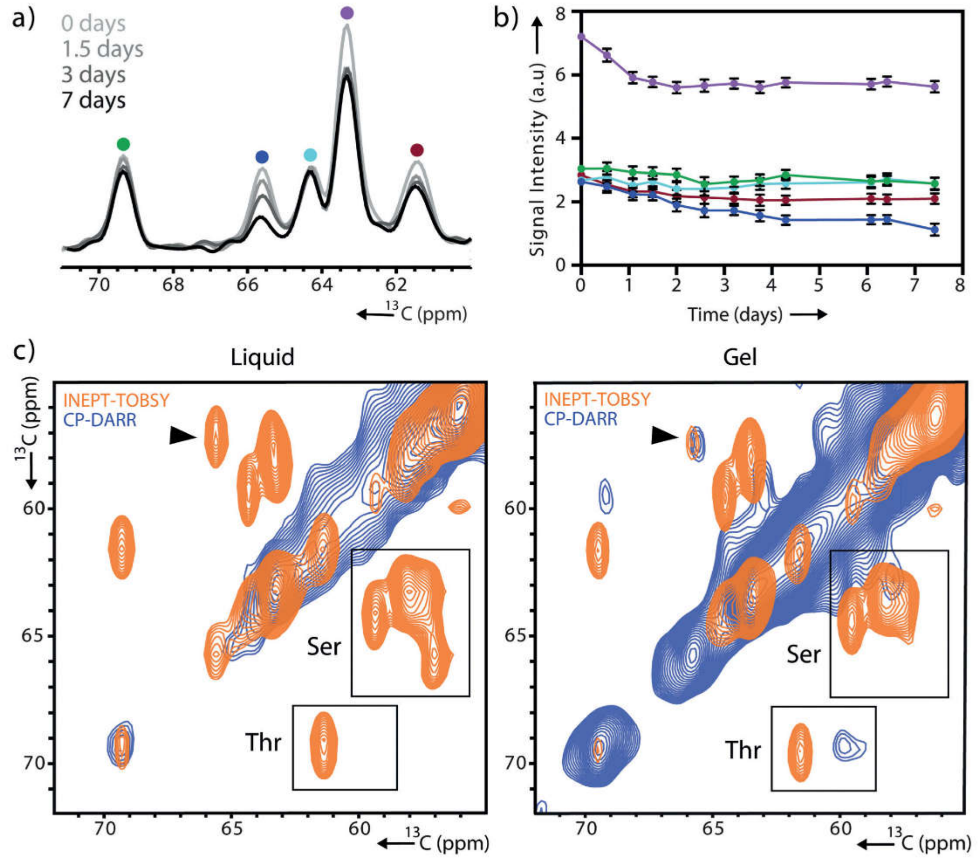


Figure 2. Site-specific variation in the gelation dynamics of pHP1α. a) 1D INEPT spectra collected 0, 1.5, 3, and 7 days after initiation of phase separation. A unique color is assigned to each peak for tracking in (b). b) Signal intensity of the five peaks shown in (a) tracked over time with corresponding colors. For error analysis, see the Supporting Information. c) Serine/threonine region of 2D ^{13}C–^{13}C CP-DARR and INEPT-TOBSY spectra recorded 3 hours after droplet formation and 7 days later. A black arrow denotes the serine cross-peak with greatest signal loss.

liquid phase.[7,18] These models, however, do not identify specific interactions and stop short of describing the maturation process from the dense liquid to the gel state. Our study provides the first molecular glimpse of this transformation. In our hands, in the absence of other components, pHP1α dense liquid droplets mature into gels over the course of a week. As the interaction network within the condensate strengthens, the molecular orientation of individual pHP1α molecules becomes restricted, a change reflected in the increase of dipolar-based signals in our MAS NMR experiments and in the incomplete recovery of our FRAP data. Yet, different regions of the protein exhibit distinct dynamic states. While the disordered tails and hinge region remain dynamic, other sites, such as the folded CD and CSD domains, experience much more restricted motions, potentially reflecting increased participation in interactions within the gel network. The loss of signal in the INEPT spectra suggests that the cross-linking interactions within the gel state may also involve the phosphorylated Ser residues, a significant observation considering their important role in promoting pHP1α phase separation.[7] The addition of physiologically relevant components, such as chromatin, disrupts the cross-linking interaction network of pHP1α condensates and prevents their complete maturation into a more rigid gel state. In addition to chromatin, "solubilization" of pHP1α condensates in the nucleus may also be achieved by the presence of non-phosphorylated HP1α and other HP1 isoforms that do not appear to phase separate. We note that due to the overwhelming concentration of pHP1α in our samples, our current study reports primarily on the behavior of individual molecules in the bulk condensed phase and not on the chromatin binding dynamics of pHP1α dimers where complex kinetics

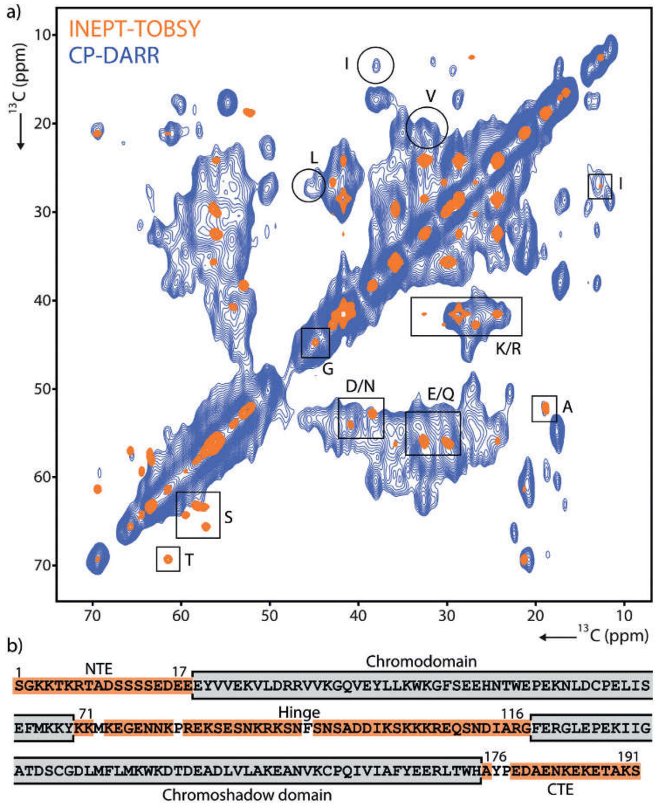


Figure 3. 2D ^{13}C–^{13}C correlation spectra of pHP1α condensates. a) Overlay of CP-DARR and INEPT-TOBSY spectra, probing the rigid and highly mobile sites of pHP1α 4 days after droplet formation. Amino acid types with characteristic and identifiable chemical shifts are marked in boxes (INEPT-TOBSY) and circles (CP-DARR). b) Complete sequence of HP1α showing structured regions (gray), and disordered regions with identified amino acid types (orange).

and states have been reported.[35] Nevertheless, the preservation of mobility in chromatin-containing pHP1α dense phases has important implications for heterochromatin compaction and regulation in the cell where access to genetic information is preserved and can be granted on demand. We expect that implementing the full suite of solid-state NMR spectroscopy experiments will help illuminate the specific intermolecular interactions and structural ensembles of biomolecules in condensed environments.

Acknowledgements

We would like to thank Ratan Rai and Anna de Angelis for assistance with the NMR spectrometers. B.A. was supported by NIH Molecular Biophysics Training Grant T32 GM008326. This work utilized the Biotechnology Research Center for NMR Molecular Imaging of Proteins at UCSD, supported by NIH grant P41 EB002031, and the UCSD Microscopy Core supported by NINDS NS047101.

Conflict of interest

The authors declare no conflict of interest.

Keywords: biophysics · heterochromatin protein 1α (HP1α) · liquid–liquid phase separation (LLPS) · solid-state NMR spectroscopy · structural biology

How to cite: *Angew. Chem. Int. Ed.* **2019**, *58*, 6300–6305
Angew. Chem. **2019**, *131*, 6366–6371

---

[1] D. M. Mitrea, R. W. Kriwacki, *Cell Commun. Signaling* **2016**, https://doi.org/10.1186/s12964-015-0125-0127.

[2] S. Kroschwald, S. Maharana, D. Mateju, L. Malinovska, E. Nüske, I. Poser, D. Richter, S. Alberti, *eLife* **2015**, https://doi.org/10.7554/eLife.06807.

[3] A. R. Strom, A. V. Emelyanov, M. Mir, D. V. Fyodorov, X. Darzacq, G. H. Karpen, *Nature* **2017**, *547*, 241–245.

[4] A. Patel, H. O. Lee, L. Jawerth, S. Maharana, M. Jahnel, M. Y. Hein, S. Stoynov, J. Mahamid, S. Saha, T. M. Franzmann, et al., *Cell* **2015**, *162*, 1066–1077.

[5] X. Su, J. A. Ditlev, E. Hui, W. Xing, S. Banjade, J. Okrut, D. S. King, J. Taunton, M. K. Rosen, R. D. Vale, *Science* **2016**, *352*, 595–599.

[6] S. Alberti, A. A. Hyman, *BioEssays* **2016**, *38*, 959–968.

[7] A. G. Larson, D. Elnatan, M. M. Keenen, M. J. Trnka, J. B. Johnston, A. L. Burlingame, D. A. Agard, S. Redding, G. J. Narlikar, *Nature* **2017**, *547*, 236–240.

[8] S. Elbaum-Garfinkle, Y. Kim, K. Szczepaniak, C. C.-H. Chen, C. R. Eckmann, S. Myong, C. P. Brangwynne, *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 7189–7194.

[9] M. Feric, N. Vaidya, T. S. Harmon, D. M. Mitrea, L. Zhu, T. M. Richardson, R. W. Kriwacki, R. V. Pappu, C. P. Brangwynne, *Cell* **2016**, *165*, 1686–1697.

[10] J. A. Riback, C. D. Katanski, J. L. Kear-Scott, E. V. Pilipenko, A. E. Rojek, T. R. Sosnick, D. A. Drummond, *Cell* **2017**, *168*, 1028–1040.e19.

[11] S. Kroschwald, M. C. Munder, S. Maharana, T. M. Franzmann, D. Richter, M. Ruer, A. A. Hyman, S. Alberti, *Cell Rep.* **2018**, *23*, 3327–3339.

[12] J. C. Eissenberg, T. C. James, D. M. Foster-Hartnett, T. Hartnett, V. Ngan, S. C. Elgin, *Proc. Natl. Acad. Sci. USA* **1990**, *87*, 9923–9927.

[13] R. C. Allshire, H. D. Madhani, *Nat. Rev. Mol. Cell Biol.* **2017**, *19*, 229–244.

[14] S. H. Kwon, J. L. Workman, *Mol. Cell* **2008**, *26*, 217–227.

[15] N. P. Cowieson, J. F. Partridge, R. C. Allshire, P. J. McLaughlin, *Curr. Biol.* **2000**, *10*, 517–525.

[16] T. Cheutin, *Science* **2003**, *299*, 721–725.

[17] G. K. Dialynas, S. Terjung, J. P. Brown, R. L. Aucott, B. Baron-Luhr, P. B. Singh, S. D. Georgatos, *J. Cell Sci.* **2007**, *120*, 3415–3424.

[18] A. G. Larson, G. J. Narlikar, *Biochemistry* **2018**, *57*, 2540–2548.

[19] T. S. Harmon, A. S. Holehouse, M. K. Rosen, R. V. Pappu, *eLife* **2017**, https://doi.org/10.7554/elife.30294.001.

[20] Y.-H. Lin, J. D. Forman-Kay, H. S. Chan, *Biochemistry* **2018**, *57*, 2499–2508.

[21] L. Leibler, M. Rubinstein, R. H. Colby, *Macromolecules* **1991**, *24*, 4701–4707.

[22] D. M. Mitrea, B. Chandra, M. C. Ferrolino, E. B. Gibbs, M. Tolbert, M. R. White, R. W. Kriwacki, *J. Mol. Biol.* **2018**, https://doi.org/10.1016/j.jmb.2018.07.006.

[23] K. A. Burke, A. M. Janke, C. L. Rhine, N. L. Fawzi, *Mol. Cell* **2015**, *60*, 231–241.

[24] J. P. Brady, P. J. Farber, A. Sekhar, Y.-H. Lin, R. Huang, A. Bah, T. J. Nott, H. S. Chan, A. J. Baldwin, J. D. Forman-Kay, et al., *Proc. Natl. Acad. Sci. USA* **2017**, *114*, E8194–E8203.

[25] S. B. Kennedy, E. R. deAzevedo, W. A. Petka, T. P. Russell, D. A. Tirrell, M. Hong, *Macromolecules* **2001**, *34*, 8675–8685.

[26] C. Ader, S. Frey, W. Maas, H. B. Schmidt, D. Gorlich, M. Baldus, *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 6281–6285.

[27] H.-K. Lin, J. C. Boatz, I. E. Krabbendam, R. Kodali, Z. Hou, R. Wetzel, A. M. Dolga, M. A. Poirier, P. C. A. van der Wel, *Nat. Commun.* **2017**, *8*, 15462.

[28] R. Tycko, *Annu. Rev. Phys. Chem.* **2011**, *62*, 279–299.

[29] D. T. Murray, M. Kato, Y. Lin, K. R. Thurber, I. Hung, S. L. McKnight, R. Tycko, *Cell* **2017**, *171*, 615–627.e16.

[30] F. Munari, M. J. Gajda, K. Hiragami-Hamada, W. Fischle, M. Zweckstetter, *FEBS Lett.* **2014**, *588*, 1094–1099.

[31] I. Matlahov, P. C. A. van der Wel, *Methods* **2018**, *148*, 123–135.

[32] O. C. Andronesi, S. Becker, K. Seidel, H. Heise, H. S. Young, M. Baldus, *J. Am. Chem. Soc.* **2005**, *127*, 12965–12974.

[33] K. Takegoshi, S. Nakamura, T. Terao, *Chem. Phys. Lett.* **2001**, *344*, 631–637.

[34] V. S. Bajaj, P. C. A. van der Wel, R. G. Griffin, *J. Am. Chem. Soc.* **2009**, *131*, 118–128.

[35] L. C. Bryan, D. R. Weilandt, A. L. Bachmann, S. Kilic, C. C. Lechner, P. D. Odermatt, G. E. Fantner, S. Georgeon, O. Hantschel, V. Hatzimanikatis, et al., *Nucleic Acids Res.* **2017**, *45*, 10504–10517.

Manuscript received: January 28, 2019
Revised manuscript received: March 7, 2019
Accepted manuscript online: March 7, 2019
Version of record online: April 3, 2019

Supplemental Information

Lab Techniques

All commonly used reagents were purchased from Sigma-Aldrich or Fisher Scientific. Isotopically enriched reagents were purchased from Cambridge Isotope Laboratories. Primer synthesis and gene synthesis were completed by Integrated DNA Technologies and Genewiz, respectively. PCR was performed using a Bio-Rad T100 thermocycler, while purification of the DNA relied on kits from Biomiga and New England BioLabs. Dialysis kits were purchased from Thermo Fisher Scientific and protein concentrators were obtained from Sartorius. Reverse-phase (RP) HPLC was performed on a 2545 Binary Gradient Module Waters system equipped with a 2484 UV/vis detector. For prep-scale RP-HPLC purification, we relied on a Waters XBridge BEH C18 19 mm x 250 mm, 10 µm particle size column, while for analytical measurements we used a Symmetry300 C18 4.6 mm x 150 mm, 5 µm particle size column. HPLC solvent A contained 100% H_2O + 0.1% trifluoroacetic acid (TFA), while solvent B contained 100% acetonitrile + 0.1 % TFA. For size-exclusion chromatography, a Superdex 200 10/300 column (GE Healthcare Life Sciences) was used with an ÄKTA pure protein purification system (GE Healthcare Life Sciences). LC-ESI-TOF MS analysis was conducted on an Agilent 6230 Accurate-Mass TOFMS. Gel images were acquired using a camera and light box from Fotodyne Incorporated. Protein and fluorophore absorbances were measured using a Nanodrop One Spectrophotometer by Thermo Scientific.

Constructs

The HP1α gene was excised from a GST HP1α plasmid provided by Naoko Tanese[1](Addgene plasmid # 24074 ; http://n2t.net/addgene:24074 ; RRID:Addgene_24074) and cloned into the pET vector backbone of a 2BT MacroLab plasmid (generously provided by Dr. Kevin Corbett) using NEBuilder® HiFi DNA Assembly Master Mix. The resulting plasmid contained a His_6-TEV-HP1α construct.

CK2 was created by excising CK2beta from the pAB46 plasmid provided by David Litchfield[2](Addgene plasmid # 27085 ; http://n2t.net/addgene:27085 ; RRID:Addgene_27085) and cloned into a pCDF-Duet-CK2alpha plasmid generously provided by Dr. Neel Shah. The resulting plasmid contained CK2alpha and CK2beta constructs that could be co-expressed.

Histone H3 K9C was created from a pET30-histone H3.1 plasmid with three mutations, C96A, C110A and K9C. Mutations were made using the NEBuilder® HiFi DNA Assembly Master Mix.

The following are the sequences of purified proteins used in this study:

HP1 alpha:

SGKKTKRTADSSSSEDEEEYVVEKVLDRRVVKGQVEYLLKWKGFSEEHNTWEPEKNLDCPELISEFMKKYKKMKEGENN
KPREKSESNKRKSNFSNSADDIKSKKKREQSNDIARGFERGLEPEKIIGATDSCGDLMFLMKWKDTDEADLVLAKEANV
KCPQIVIAFYEERLTWHAYPEDAENKEKETAKS*

HP1 alpha – GSKCK:

SGKKTKRTADSSSSEDEEEYVVEKVLDRRVVKGQVEYLLKWKGFSEEHNTWEPEKNLDCPELISEFMKKYKKMKEGENN
KPREKSESNKRKSNFSNSADDIKSKKKREQSNDIARGFERGLEPEKIIGATDSSGDLMFLMKWKDTDEADLVLAKEANV
KCPQIVIAFYEERLTWHAYPEDAENKEKETAKSGSKCK*

Histone H3 K9C:
ARTKQTARCSTGGKAPRKQLATKAARKSAPATGGVKKPHRYRPGTVALREIRRYQKSTELLIRKLPFQRLVREIAQDFKT DLRFQSSAVMALQEAAEAYLVGLFEDTNLAAIHAKRVTIMPKDIQLARRIRGERA*

HP1 alpha expression and purification

BL21(DE3)-Rosetta cells were co-transformed with the HP1 alpha and CK2 plasmids. Pre-cultures were grown at 37 °C overnight under ampicillin and streptomycin antibiotic selection in 2X LB. Pre-cultures were pelleted in the morning, LB was removed and the pellet was resuspended in ^{13}C, ^{15}N enriched M9 media and added to 1L ^{13}C, ^{15}N enriched M9 expression cultures. Many of the following steps have been previously described for HP1α.[3] Cells were grown at 37 °C until OD600 = 0.6 (~12 hours), transferred to 18 °C and induced with 0.3 mM IPTG. Cells were collected 20 hours later by centrifuging at 5,000 xg. The cell pellet was resuspended in lysis buffer (1x PBS, pH 7.4, 300 mM NaCl, 10% glycerol, 7.5 mM imidazole, Roche protease inhibitor tablet) and sonicated at 4 °C. Cell debris was cleared by centrifuging for 30 minutes at 30,000 xg. Roughly 5mL of Ni-NTA resin (per 1L culture) was added to the lysate and rotated at 4 °C for one hour. The lysate-resin mixture was loaded into a Bio-Rad Econo-Column and washed with wash buffer (1x PBS, pH 7.4, 300 mM NaCl, 10% glycerol, 7.5 mM imidazole). Elution buffer (20 mM HEPES, pH 7.2, 300 mM KCl, 400 mM imidazole, 1 mM DTT) was then ran over the column. The resulting elution was incubated with TEV-protease during dialysis against imidazole-free buffer overnight at 4 °C. The sample was analyzed by SDS-PAGE to ensure the complete removal of the His tag. At this point better purity was achieved utilizing reverse-phase HPLC than using size-exclusion or ion-exchange chromatography. The phosphorylated HP1α solution was adjusted to 6 M Guanidinium HCl concentration and filtered. This solution was loaded onto a Waters XBridge BEH C18 prep-size reverse-phase HPLC column and fractions were collected using a gradient of 10%-60% solvent B. HP1α eluted around 40% solvent B and due to protein retention on the column, multiple runs were required to collect the protein. The elution product was lyophilized and stored at -80 °C. Purity was assessed by SDS-PAGE and analytical RP-HPLC.

Cy3-HP1α conjugation

Cy3 labeling was completed as described previously (3). Briefly, the GSKCK motif was added to the C-term of HP1α and cysteine 133 was mutated to serine. The HP1α GSKCK construct was co-expressed with the CK2 kinase and purified. Excess Cy3-maleimide was added to 100 µM HP1 alpha and reacted for ~10 seconds, then quenched with excess β-mercaptoethanol. The reaction product was purified on a prep-size HPLC column and only the peak corresponding to a single Cy3 modification was collected and confirmed by analytical HPLC and mass spectrometry.

HP1α preparation

HP1α was refolded by dissolving the lyophilized stock in resuspension buffer (20 mM HEPES pH 7.2, 6 M G-HCl, 20 mM KCl, 2 mM DTT), dialyzing against 1 M GuanidiniumHCl for 5 hours, then dialyzing against guanidinium-free buffer with 300 mM KCl overnight at 4 °C. We routinely achieved 70% yield during the refolding process with proper refolding confirmed by solution NMR. At 300 mM KCl, HP1α does not undergo LLPS. The dialysis product was first centrifuged at 4,000 x *g* to remove any precipitate, then concentrated using Sartorius centrifugal concentrators. The sample was concentrated to completion and then diluted 1:4 in 0 M KCl buffer to achieve final concentration of 75 mM KCl. The final buffer used for all NMR and FRAP experiments was 20 mM HEPES pH 7.2, 75 mM KCl, 2 mM DTT, 0.01% NaN_3. Since

HP1α immediately phase separates upon dilution in lower salt conditions, protein concentration was determined prior to dilution using an A280 extinction coefficient of 29,160 $M^{-1} \cdot cm^{-1}$. The resulting phase separated sample was incubated at the desired temperature, then gently centrifuged at 100 x *g* to coalesce droplets. For solution NMR, this was done while in the NMR tube and resulted in a single droplet phase within the NMR coil range of the tube. For solid-state NMR, this was done in a 1.7 mL tube and the droplets alone were then transferred to the rotor.

Histone H3 methyl-lysine installation

To install the methyl-lysine analog we followed a previously established protocol.[4] Briefly, histone H3 K9C, C110A was expressed, purified and lyophilized as previously described[5]. 5 mM of H3 K9C was completely reduced at 50 °C for an hour in 1M HEPES pH 7.8, 4M GuanidiniumHCl, 10 mM D/L-methionine and 20 mM DTT. Afterwards, 400 mM of (2-bromoethyl)-trimethylammonium bromide (Sigma) was added to the solution. The sample was kept in the dark at 50 °C and occasionally inverted to mix. After 2.5 hours, 10 mM of fresh DTT was added, and the reaction was allowed to proceed for another 2.5 hours. The reaction was finally quenched with 700 mM 2-mercaptoethanol. The product was purified by reverse-phase HPLC and confirmed by ESI-TOFMS.

Octamer and 12-mer nucleosome array assembly

DNA and histones were prepared using established protocols.[6,7] 12x601 DNA (12 repeats of the 601 DNA sequence separated by 30 bp linkers) for the 12-mer array and MMTV for buffer DNA were expressed in DH5α cells. Plasmids were digested and the DNA fragments were purified by phenol-chloroform extraction and ethanol precipitation. The DNA pellets were redissolved in 10 mM Tris pH 8.0, 0.1 mM EDTA buffer and stored at –20 °C.

The 601-site sequence is:
CTGGAGAATCCCGGTGCCGAGGCCGCTCAATTGGTCGTAGACAGCTCTAGCACCGCTTAAACGCACGTACGCGCT
GTCCCCCGCGTTTTAACCGCCAAGGGGATTACTCCCTAGTCTCCAGGCACGTGTCAGATATATACATCCTGTGCATG
TAAGATCCAGTACTACGCGGCCGCC

The MMTV sequence is:
ACTTGCAACAGTCCTAACATTCACCTCTTGTGTGTTTGTGTCTGTTCGCCATCCCGTCTCCGCTCGTCACTTATCCTTC
ACTTTCCAGAGGGTCCCCCCGCAGACCCCGGCGACCCTGGTCGGCCGACTGCGGCACAGTTTTTTG

Histones were expressed in BL21(DE3)-Rosetta cells, purified from inclusion bodies by reverse-phase HPLC and lyophilized. The histones were co-dissolved in 6M GuanidiniumHCl, 20mM Tris pH 7.5, then dialyzed against 0 M GuanidiniumHCl to refold and form octamers that were further purified with a size-exclusion Superdex 200 10/300 column (GE Healthcare Life Sciences) as in published protocols.[8] Octamers were stored in buffer containing 1M NaCl and 50% glycerol at –20 °C.

12-mer nucleosome arrays were prepared using previously described protocols.[5] Briefly, 1.6:1 12-mer DNA to octamer and 0.3:1 MMTV DNA to octamer were added to 2M TEK buffer (2 M KCl, 10 mM Tris pH 7.5, 0.1 mM EDTA) and 2 µM octamers in 10 mL of solution. Samples were dialyzed into 10 mM TEK buffer (10 mM KCl), precipitated by the addition of 4 mM $MgCl_2$ and resuspended in assay buffer (20 mM HEPES pH 7.2, 75 mM KCl, 2 mM DTT, 0.01% NaN_3) to be used immediately. Chromatin arrays were analyzed by native 1% agarose/2% polyacrylamide (APAGE) gels stained with SyBr Gold (Life Technologies). Chromatin concentrations were the equivalent of 150 µM and 200 µM 601-sites in the phase separated condensates for solid-state NMR and FRAP, respectively. Concentrations were determined using molar extinction coefficient ε_{A260} = 2822957.

Solution NMR spectroscopy

Experiments were performed using a Bruker Avance Neo 800 MHz (^{1}H Larmor frequency) NMR spectrometer equipped with a triple resonance TCO cryoprobe. Chemical shifts were referenced to the TMS frequency. The ^{1}H-^{15}N TROSY-HSQC spectra were acquired using the standard Bruker trosyf3gpph19 pulse sequence. The dilute state HSQC spectrum was acquired with 16 scans while the phase separated spectrum required 96 scans. Experiments used the following parameters: 298 K, 2048 points and 90 ms acquisition in the direct dimension, 256 points and 44 ms acquisition in the indirect dimension, interscan delay of 1.25 s, ^{1}H center frequency of 4.7 ppm, ^{15}N center frequency of 120 ppm, and J_{NH} = 90 Hz.

Solid-state NMR spectroscopy

Experiments were performed on a 750 MHz (^{1}H Larmor frequency) magnet equipped with an AVANCE III Bruker console. All spectra were recorded using a triple resonance (^{1}H, ^{13}C, ^{15}N) 3.2 mm MAS Bruker probe (PH MAS DVT 750SB EFREE BL3.2). The magic angle was set with KBr and the ^{13}C chemical shift was referenced to the downfield peak of adamantane (40.49 ppm).[9] Sample temperature was maintained at 293 K, while spinning at 11.1 kHz using a 3.2 mm thin-wall Bruker rotor.

1D CP experiments had the following parameters: 293 K, 64 scans, 4 s interscan delay, 9 ms acquisition, 1024 points, ^{13}C center frequency set at 100 ppm. For CP, a 500 µs contact time was used with the Hartmann-Hahn matching condition satisfied by setting the ^{13}C power to 50 kHz and optimizing on the ^{1}H channel. 100 kHz spinal64 decoupling was performed during acquisition.

1D INEPT experiments had the following parameters: 293 K, 64 scans, 4 s interscan delay, 9 ms acquisition, 1024 points, ^{13}C center frequency set at 100 ppm, and J_{CH} = 140 Hz.

2D DARR experiments used the cpSPINDIFF pulse sequence, with the following parameters: 293 K, 16 or 48 scans, 4 s interscan delay, 1024 points and 9 ms acquisition in the direct dimension, 256 points and 2.7 ms acquisition in the indirect dimension, ^{13}C center frequency 100 ppm, and DARR condition set to 11 kHz. For CP, a 500 µs contact time was used, together with 20 ms of DARR mixing. 100 kHz spinal64 decoupling was performed during acquisition.

2D INEPT-TOBSY used the pulse sequence described in ref. [7], with the following parameters: 293 K, 16 scans, 4 s interscan delay, 1536 points and 13.5 ms acquisition in the direct dimension, 720 points (9 ms) or 144 points (1.8 ms) indirect, in the indirect dimension with ^{13}C center frequency set at 100 ppm and J_{CH} = 155 Hz. 5 TOBSY cycles were used for a total of 7 milliseconds mixing.

Spectral analysis

Spectra were processed using Bruker TopSpin 4.0.5 and analyzed using SPARKY.[10] Peak and integrated signal intensities were quantified using TopSpin 4.0.5.

All quantified peaks were over 6 times the average noise level of each spectrum. Error bars for graphs were calculated by:

$$\sqrt{\left(\frac{\sigma_{gel}}{I_{gel}}\right)^2 + \left(\frac{\sigma_{liquid}}{\mathrm{I}_{liquid}}\right)^2}$$

where σ_{gel} and σ_{liquid} is the noise level of the gelation time point and the initial dense liquid, respectively. I_{gel} and I_{liquid} corresponds to the peak or integrated signal intensities of the gel and liquid, respectively.

Fluorescence microscopy

Cy3-labeled HP1α from a 90 µM stock was added to phase separated HP1α samples to a final concentration of 1 µM. The solution was loaded into a channel between a glass slide and a coverslip raised by double-sided tape. Fresh slides were prepared for each time point. Droplets were imaged with an Olympus FV1000 confocal microscope at 40x using 543 nm wavelength for detection. 10 µm diameter circles within the droplets were bleached over 5 seconds, then monitored for fluorescence recovery over 2 minutes. To quantify fluorescence recovery, signals were normalized between the initial droplet fluorescence intensity and the signal achieved directly after photobleaching. Six droplets were bleached per time point. Using GraphPad Prism version 8.0, data were fit to a two phase association model, $I = I_0 + a\left(1 - e^{-\frac{t}{\tau_{fast}}}\right) + b\left(1 - e^{-\frac{t}{\tau_{slow}}}\right)$ where a and b are fit parameters that account for the fraction of fast and slow components and τ_{fast} and τ_{slow} are the recovery times of those components, while I_0 = 0.

Methods references

[1] M. F. Vassallo, N. Tanese, *Proceedings of the National Academy of Sciences* **2002**, *99*, 5919–5924.

[2] J. P. Turowec, J. S. Duncan, A. C. French, L. Gyenis, N. A. St. Denis, G. Vilk, D. W. Litchfield, in *Methods in Enzymology*, Elsevier, **2010**, pp. 471–493.

[3] A. G. Larson, D. Elnatan, M. M. Keenen, M. J. Trnka, J. B. Johnston, A. L. Burlingame, D. A. Agard, S. Redding, G. J. Narlikar, *Nature* **2017**, *547*, 236–240.

[4] M. D. Simon, F. Chu, L. R. Racki, C. C. de la Cruz, A. L. Burlingame, B. Panning, G. J. Narlikar, K. M. Shokat, *Cell* **2007**, *128*, 1003–1012.

[5] G. T. Debelouchina, K. Gerecht, T. W. Muir, *Nature Chemical Biology* **2017**, *13*, 105–110.

[6] A. Flaus, T. J. Richmond, *Journal of Molecular Biology* **1998**, *275*, 427–441.

[7] P. N. Dyer, R. S. Edayathumangalam, C. L. White, Y. Bao, S. Chakravarthy, U. M. Muthurajan, K. Luger, in *Methods in Enzymology*, Elsevier, **2003**, pp. 23–44.

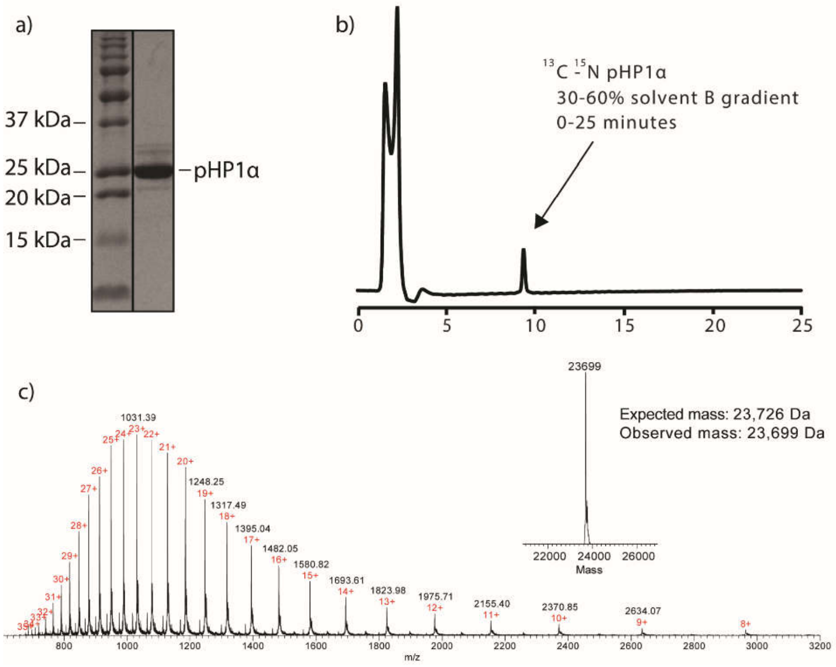
[8] B. Fierz, C. Chatterjee, R. K. McGinty, M. Bar-Dagan, D. P. Raleigh, T. W. Muir, *Nature Chemical Biology* **2011**, *7*, 113–119.

[9] C. R. Morcombe, K. W. Zilm, *Journal of Magnetic Resonance* **2003**, *162*, 479–486.

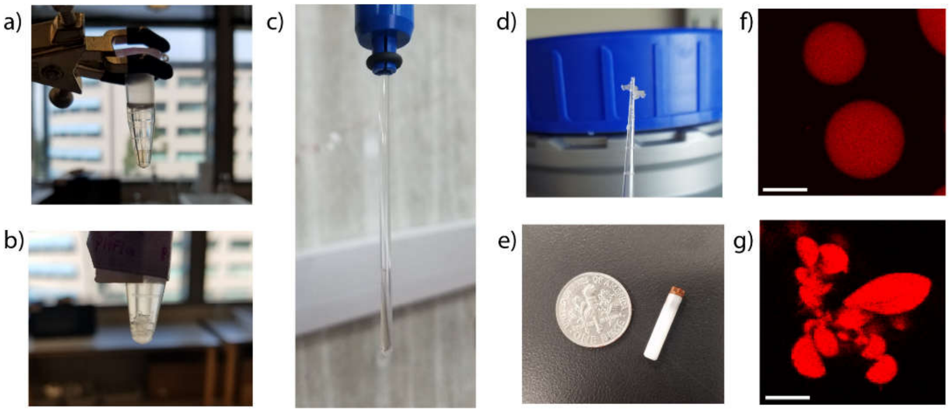
[10] W. Lee, M. Tonelli, J. L. Markley, *Bioinformatics* **2015**, *31*, 1325–1327.

[11] J. Leppert, O. Ohlenschläger, M. Görlach, R. Ramachandran, *Journal of Biomolecular NMR* **2004**, *29*, 167–173.

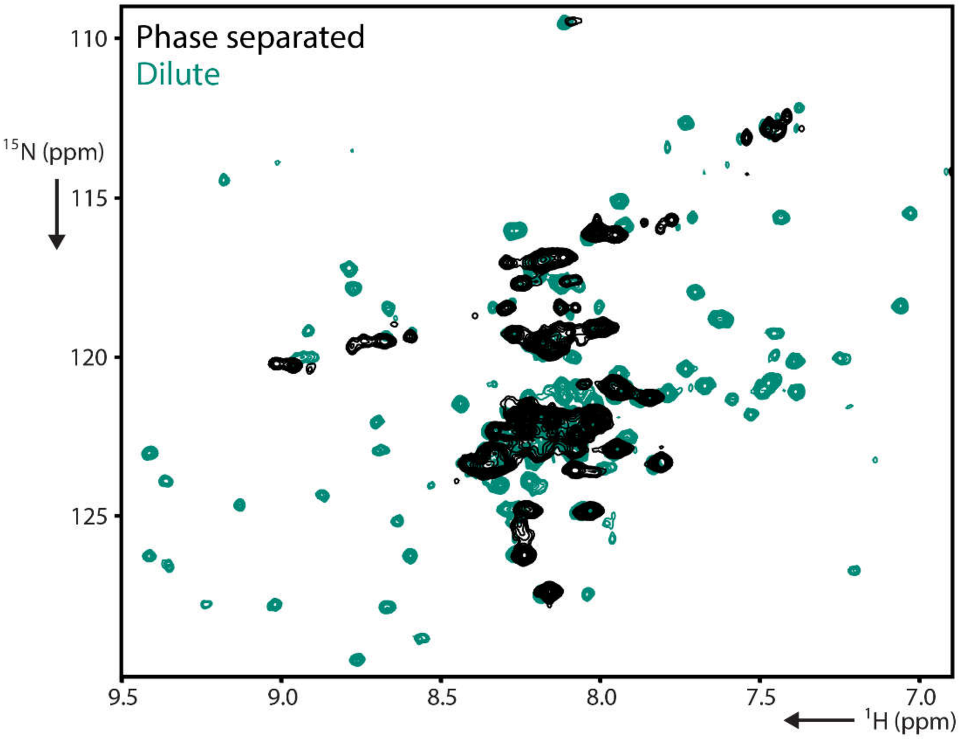
Supplemental Figures



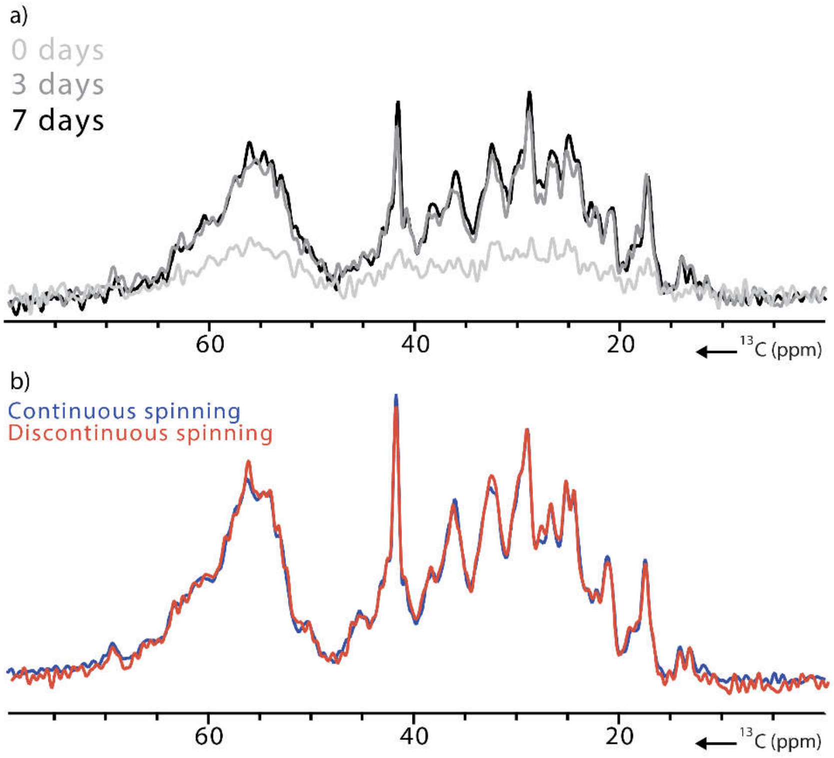
Supplemental Figure 1. Analysis of recombinantly prepared and phosphorylated HP1α. **(a)** SDS-PAGE gel analysis of the purified protein. **(b)** Analytical RP-HPLC A214 trace of pHP1α taken directly from a pHP1α droplet sample. **(c)** ESI-TOF MS results. Calculated mass for WT ^{13}C,^{15}N- labeled HP1α = 23,406 Da, calculated mass for tetraphosphorylated ^{13}C,^{15}N- labeled protein = 23,726 Da, mass determined by ESI-TOF MS = 23 699 Da, corresponding to an average of 3.66 phosphorylation sites per protein molecule.



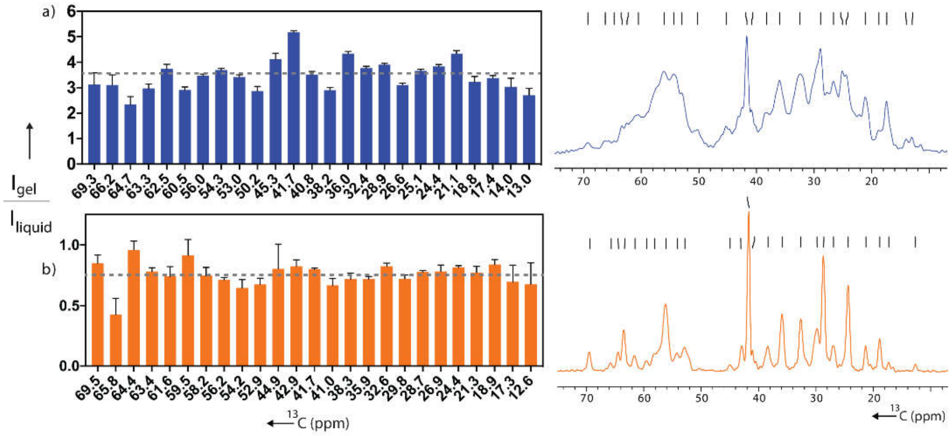
Supplemental Figure 2. Images of pHP1a liquid and gel condensates. **(a)** Liquid droplets coalesced at the bottom of a tube. **(b)** The same droplets left for a week at room temperature form a gel. **(c)** An NMR tube with coalesced droplets at the bottom. **(d)** A pipette tip stab of the sample after 8 days of MAS NMR. **(e)** An image of the 3.2 mm rotor used for these experiments, a U.S. dime displayed for scale. **(f)** Cy3-labeled liquid droplets of pHP1α, scale bar 10 microns. **(g)** pHP1α gel formed after 8 days, scale bar 10 microns.



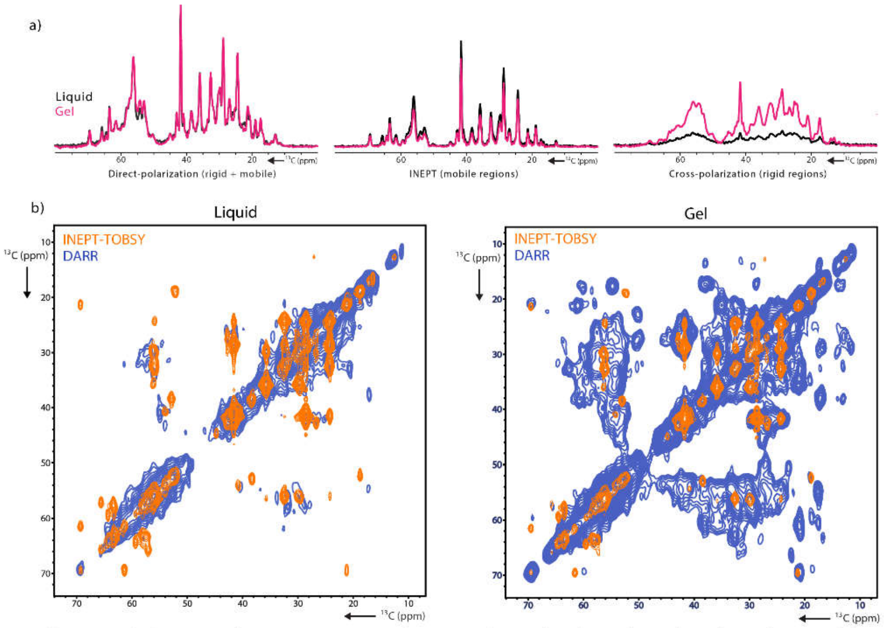
Supplemental Figure 3. Solution NMR spectra of pHP1α. Overlay of ^{1}H-^{15}N TROSY-HSQC spectra for dilute (100 μM) pHP1α (green) and phase separated (~5 mM) pHP1α in the dense liquid state (black). Both samples are dissolved in 20 mM HEPES pH 7.2, 75 mM KCl and 2 mM DTT at 298 K.



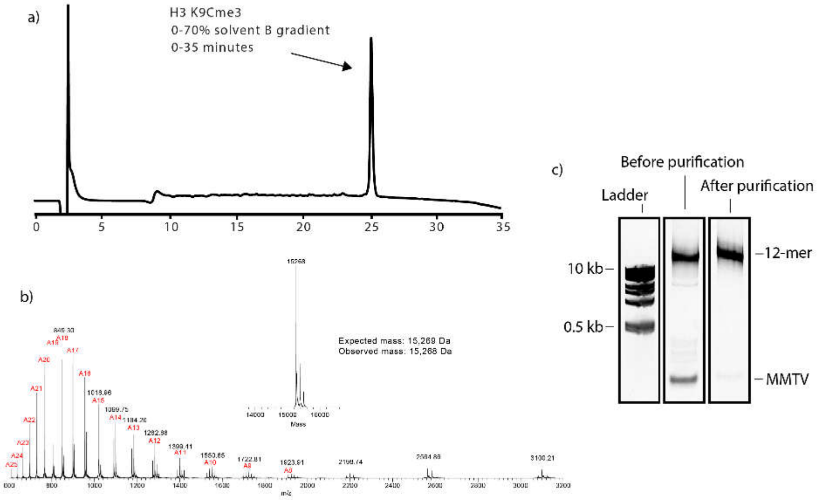
Supplemental Figure 4. Gelation control without continuous MAS spinning. **(a)** A separate ^{13}C-^{15}N pHP1α sample was prepared and the same 1D experiments were recorded at 0, 3, and 7 days. Between 1D experiments, the sample was removed from the magnet and left to sit at room temperature. The increase in CP intensity was 2.6 and 2.8 after 3 and 7 days respectively. **(b)** Overlaid 1D CP experiments in the gel state of the MAS (7 days) and control sample (14 days) confirm consistent chemical shifts and relative peak intensities. Taken together, these data indicate that the CP changes observed in the MAS sample are due to gelation and not due to dehydration or sedimentation.



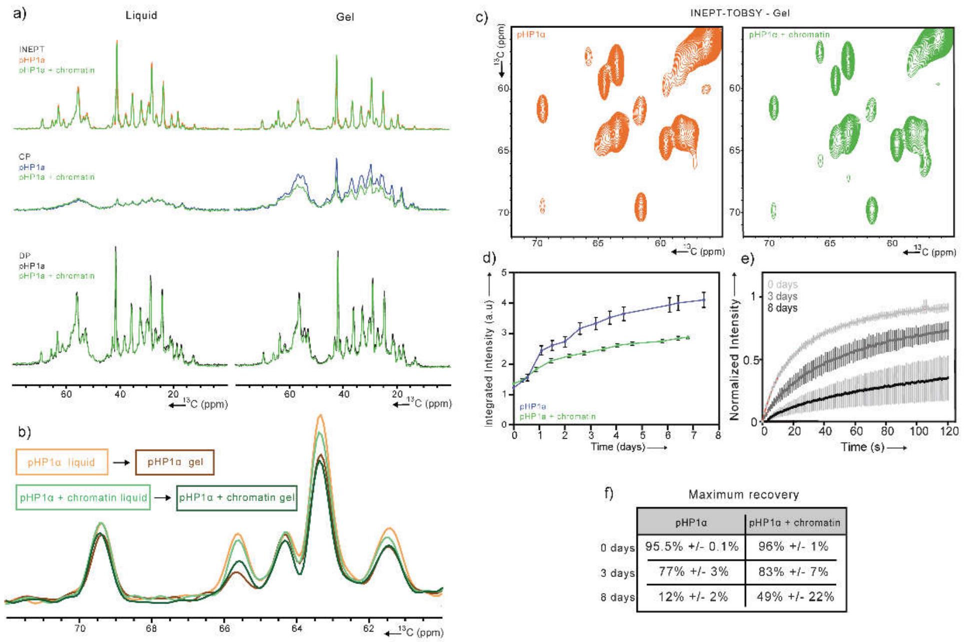
Supplemental Figure 5. Peak intensity changes in the 1D spectra of pHP1α as a result of gelation. **(a)** Intensity change comparison of the 25 strongest CP peaks before and after gelation. Selected peaks are shown on the 1D CP spectrum of the final gel state (right). **(b)** Intensity change comparison of the 25 strongest INEPT peaks, before and after gelation. Selected peaks are shown on the 1D INEPT spectrum of the final gel state (right). Dashed lines represent the average change in intensity. Only peaks with S/N > 6 were selected.



Supplemental Figure 6. Changes in pHP1α dynamics from the dense liquid to the gel state. **(a)** 1D overlays of the spectra shown in Figure 1B. **(b)** 2D ^{13}C-^{13}C DARR and INEPT-TOBSY overlays from the dense liquid (left) and gel state (right). These are the complete spectra from which Figure 2c was derived.



Supplemental Figure 7. Preparation of H3K9Cme3 chromatin arrays. **(a)** Analytical RP-HPLC trace of the H3K9Cme3 protein. (b) ESI-TOF MS results for the H3K9Cme3 sample. Expected mass = 15,269 Da, observed mass = 15,268 Da. **(c)** Native APAGE gel analysis of the chromatin assembly. Before Mg^{2+} purification the sample contains leftover “buffer” MMTV DNA and MMTV nucleosomes, while after purification, only the 12-mer nucleosome arrays remain.



	pHP1α	pHP1α + chromatin
0 days	95.5% +/- 0.1%	96% +/- 1%
3 days	77% +/- 3%	83% +/- 7%
8 days	12% +/- 2%	49% +/- 22%

Supplemental Figure 8. Comparative analysis of pHP1α gelation in the presence and absence of H3K9Cme3 chromatin. **(a)** Comparison of ^{13}C 1D INEPT, CP and DP spectra of pHP1α dense liquid and gel condensates in the presence and absence of chromatin. The data without chromatin are the same as in Fig. 1a. **(b)** Overlay of the Ser/Thr region of 1D INEPT spectra of pHP1α with and without chromatin before and after gelation. **(c)** 2D ^{13}C-^{13}C INEPT-TOBSY of pHP1α gels with and without chromatin. Only the Ser/Thr region of the spectra is shown. **(d)** Integrated intensity of the aliphatic regions of 1D ^{13}C CP spectra recorded over time for pHP1α only and pHP1α with chromatin. **(e)** FRAP recovery curves for pHP1α condensates containing 200 μM chromatin collected immediately upon droplet formation, 3 days and 8 days later. Red lines represent exponential fits. **(f)** Maximum recovery of FRAP signal of pHP1α condensates with and without chromatin.

Acknowledgements

Chapter 3, in full, is a reprint of the material as it appears in Heterochromatin protein HP1α gelation dynamics revealed by solid-state NMR spectroscopy. 2019. Ackermann, Bryce E. and Debelouchina, Galia T. Angewandte Chemie International Edition. The dissertation author was the primary investigator and author of this material.

Chapter IV

HP1 CSD interactions with chromatin and chromatin effectors

Abstract

Heterochromatin protein 1α (HP1α) is a conserved protein prominent in heterochromatin formation and maintenance, associated with the repression of gene expression. While HP1α is a relatively small protein, it has several domains with distinct functions. The chromodomain binds to the epigenetic histone marks H3 K9me2/3, the hinge binds to DNA, and the chromoshadow domain (CSD) binds to a long list of partner proteins. Here, we investigate whether the αN helix of histone H3 is a binding target of the CSD to induce gene repression. We use NMR spectroscopy and gel-based assays to assess CSD binding to an αN helix peptide and mononucleosomes. Based on recent structural models of transcription factors and chromatin remodelers unwrapping DNA from contact with the αN helix, we introduce these proteins to prompt HP1α binding to nucleosomes. However, we are unable to detect CSD docking to the nucleosome. Moving beyond binding assays, we design a chromatin remodeling assay to detect any effect of the HP1α CSD on nucleosome accessibility. We discover that it is not binding to the αN helix but rather the compaction of nucleosome arrays by HP1α that exerts repressive effects on the gene regulatory machinery.

4.1 Introduction

Heterochromatin is well-known for its role as a gene silencing feature of the eukaryotic genome. The constitutive heterochromatin territory can be recognized by its DNA content (telomeres, tandem repeats, pericentromeres, late-stage replication), post-translational modifications (H3 K9me3 and H4 K20me3), and effector proteins such as heterochromatin protein 1α (HP1α).[1] However, viewing heterochromatin only as a gene silencing domain oversimplifies its active roles in establishing inheritance,[2] driving nuclear topology,[3] fine-tuning (rather than preventing) transcription,[4] guiding cell differentiation,[5] and segregating chromosomes during cell division.[6] And while these diverse functions are enacted by a wide range of biomolecules, central to these actions is the protein HP1α.

HP1α is a heterochromatic protein intertwined in seemingly all heterochromatin events. It is exactly this widespread distribution, existence of partially redundant paralogs (β, and γ) and vast list of binding partners that makes it difficult to describe what it is that HP1α does. HP1 is haunting in its contradiction, simultaneously localizing to transcriptionally silent and active regions,[7] both rapidly diffusing and immobile,[8] paralogs unique but also redundant.[9] We are interested in characterizing the interaction and structural capabilities of HP1α to construct a base for its cellular complexity. In particular, we aim to probe the interactions of HP1α with the nucleosome because such interactions may be foundational to gene regulation.

The interaction between the chromodomain (CD) and the di/trimethylated lysine 9 on histone H3 is well documented and critical to the localization, maintenance and spreading of heterochromatin **(Figure 1.3).**[8, 10] The hinge domain is able to interact with nucleosomal DNA and can even distinguish DNA tertiary structures like G-quadruplexes.[11] HP1α is also suggested to bind directly the folded regions of histones H3 and H2A.Z.[12] While the mechanism of binding to H2A.Z is unknown, the binding to H3 is believed to be through the CSD domain on HP1α and the H3 αN helix.[13] The H3 αN helix has been co-crystalized as a small peptide with the CSD of HP1γ, and can interact with all paralogs of HP1.[14] Further, the phosphorylation of the H3 αN helix is associated with cell-cycle dependent removal of HP1α.[15] The CSD can interact with the PXVXL motif (where X is any amino acid), but has known sequence deviation.[14] For example, the CSD can bind the non-canonical motif on H3 (PGTVAL), in addition to a vast array of chromatin effectors that enact their own function of methylation, nucleosome remodeling, DNA sequence recognition, and many more.[16] A direct interaction to the H3 core predicts HP1α to have a direct and disruptive interaction that would have significant impact on many nucleosome regulating processes. With such implications, we sought to better probe the interaction of the CSD with nucleosomes in the presence of chromatin effectors that require access to the histone cores and the wrapped DNA.

Assessing the interaction between the CSD and the αN helix (or any binding partner) in cells is challenging because any manipulation of the PXVXL-binding region abrogates interaction with hundreds of binding partners and therefore complicates any causal relationships. For this reason, such interactions are typically studied with purified components. This chapter focuses on the interplay of HP1 (transcription repressor), BRG1 (remodeler), and OCT4/SOX2 (transcription factors) in the context of the nucleosome **(Figure 4.1A)**. HP1 has been shown to interact with every family of chromatin remodelers through its CSD,[12b, 17] making it clear that one function of HP1 is to affect the proper sliding or removal of nucleosomes. BRG1 in particular is a common hit for proteomic studies looking for HP1α interactions in cells. *In vitro* studies have even identified the PXVXL region of BRG1 and shown a competitive relationship between HP1α and BRG1 over access to the nucleosome.[12b, 18] Interestingly, both HP1α and BRG1 were been shown to interact with the H3 αN helix.[18] The location of the H3 αN helix is important because it establishes the entry/exit site of the nucleosome, the last DNA-histone interface to complete DNA wrapping. This DNA site is called the super helical location (SHL) +/- 6 and SHL +/- 7. Access to this region would be required to enact major structural changes to the nucleosome. Unsurprisingly, this helix is a hotspot for post-translational modifications that impact DNA wrapping and binding partners. For example, the H3 K56 acetylation mark forces the DNA to unwrap partially from the nucleosome, exposing new sites of DNA for transcription factors.[19] This helix can also be phosphorylated at H3 T41 and T45, which disperses HP1α in cells, and prevents binding by the CSD in *in vitro* peptide binding assays.[15b] While preliminary studies have laid the groundwork for the BRG1-HP1α relationship, more work is required to elucidate the source of competition between them and the biological implication.

Beyond remodelers, HP1 may have similar relationships with transcription factors. Recent structures have revealed that OCT4 and SOX2 bind directly to nucleosomal DNA at SHL -6 and induce slight DNA unwrapping. OCT4 and SOX2 are pluripotency factors, able to orchestrate the conversion of differentiated cells into a pluripotent state, essentially erasing the identity of a cell.[20] Being able to bind directly to the nucleosome and thus bypass the steric exclusion of the histone octamer, means that these special transcription factors must be strategically regulated. HP1 may be in part responsible for this regulation, especially with the understanding that constitutive heterochromatin (likely involving all three HP1 paralogs) is an impediment to OCT4-SOX2-induced pluripotency.[21] Both H3 K9me3 and HP1γ must be removed for OCT4 to initiate cell reprogramming to pluripotency.[5c, 22] Upon initiation by OCT4, BRG1 is

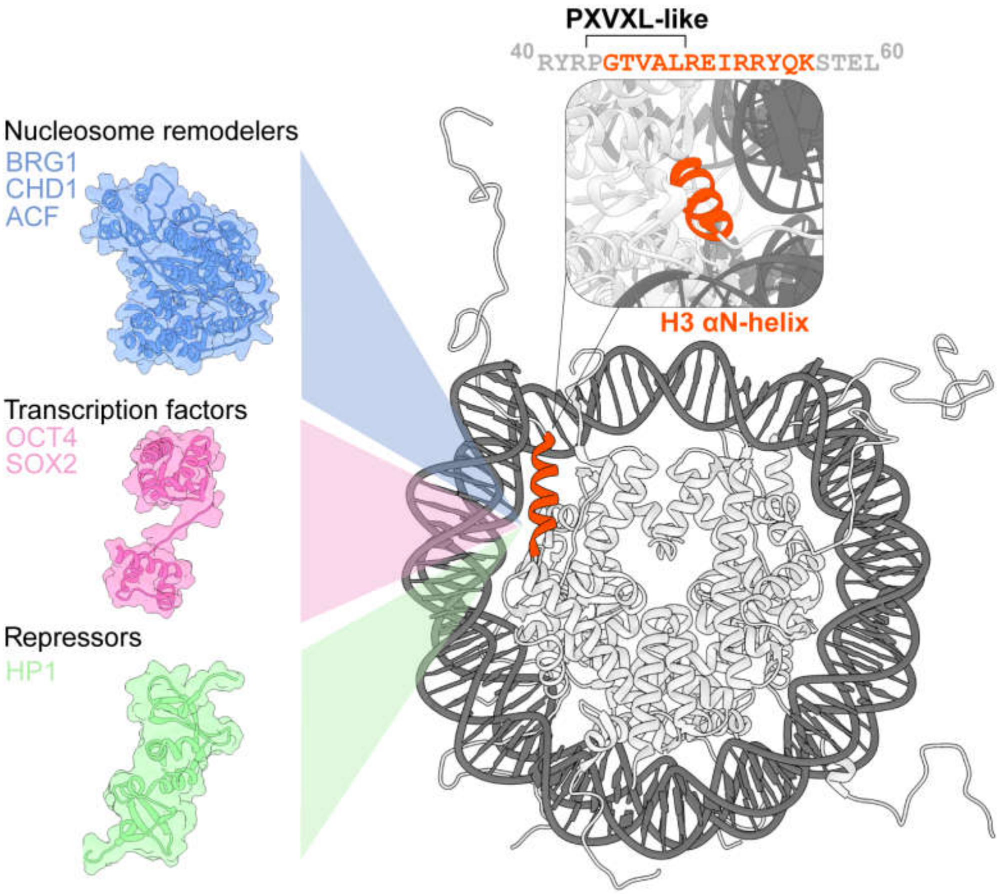


Figure 4.1. Landscape of nucleosome interactions. The H3 αN helix, tucked into the nucleosome, is shown in red. The PXVXL-like sequence (PGTVAL) is the last folded sequence of H3 before the disordered N-terminal tail begins. Three types of chromatin effectors are shown; nucleosome remodelers, transcription factors, and repressors. Each type benefits from access to the nucleosome surface and likely the H3 αN helix. PDB structures used (1KX5, 3Q6S, 6T93, 5X0X).

recruited to open up chromatin and establish a landscape for new gene expression.[23] Interestingly, while OCT4 can bind to closed heterochromatic regions, BRG1-dependent remodeling is necessary to make chromatin accessible to other factors.[23b] Additionally, HP1γ was able to interact with OCT4 and other pioneer factors, but interaction interfaces were not fully determined.[5c] Altogether, the regulation of pluripotency may in part depend on a regulatory triad of HP1, OCT4, and BRG1 where all components have proven interactions with each other, and their interactions are either required for function (OCT4 and BRG1) or prevention of function (HP1 and BRG1/OCT4).

This chapter was built from the desire to describe the observed cellular relationships between these three protein families on a molecular level. The model described above, while captivating, is complicated by the reliance on αN helix binding experiments that only use peptides or free histones (without DNA) when probing interactions with the CSD.[12b, 14] The αN helix is largely folded and buried within the nucleosome and remains that way in most currently determined structures, even with bound proteins.[24] However, there has been a recent appreciation of histone core flexibility accounting for significant structural movement within the nucleosome,[25] a phenomenon exacerbated by remodeler and HP1 binding.[26] The first aim of this chapter is to determine if the CSD of HP1α is sufficient to bind to the nucleosome, as it is for free histones. The following aim is to test how HP1α or HP1γ competes or cooperates with transcription factors (OCT4 and SOX2) and remodelers (BRG1, CHD1, ACF) for access to the nucleosome. Finally, we aim to identify the mechanism of HP1α regulation of active chromatin remodeling and if it relies on interaction with the αN helix.

4.2 Results

HP1 interactions with the nucleosome

It has been suggested through a range of studies that the CSD of HP1 can bind to the H3 αN helix.[12b, 14, 15b] To test this in our own hands, we sought to repeat the known interaction with the H3 αN helix peptide and progressively increase the complexity of the system toward a fully wrapped nucleosome. We started by testing for interactions between the HP1α CSD and the H3-H4 histone tetramer. Co-elution of the

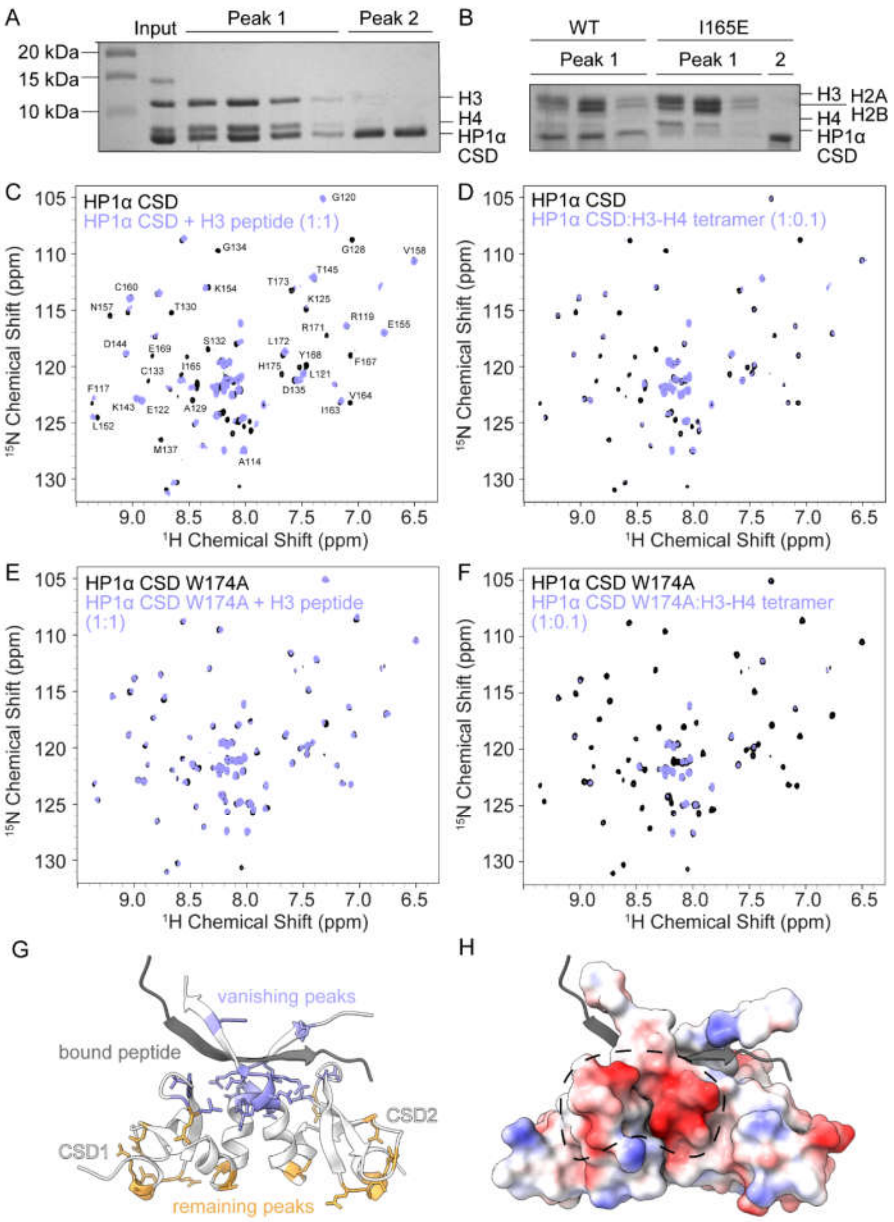


Figure 4.2. Chromoshadow domain binding to histone H3 in the absence of DNA. A) SDS-PAGE of two peaks eluted from size-exclusion chromatography. Peak 1 co-elutes the histone tetramer and CSD, peak 2 contains unbound CSD. B) SDS-PAGE of wildtype (WT) CSD co-eluting with the histone octamer, while the I165E construct shows no co-elution. C) ^{1}H-^{15}N HSQC of HP1α CSD (black) and HP1α CSD with equimolar H3 αN helix peptide (purple). D) 1H-15N HSQC of HP1α CSD and HP1α CSD 1:0.1 with the H3-H4 tetramer. E) 1H-15N HSQC of HP1α CSD W174A (black) and HP1α CSD W174A with equimolar H3 αN helix peptide (purple). D) 1H-15N HSQC of HP1α CSD W174A and HP1α CSD W174A 1:0.1 with the H3-H4 tetramer. G) Structural model of the HP1 CSD dimer bound to the H3 αN helix peptide (PDB: 3Q6S). Purple residues are the first peaks to disappear upon tetramer addition, while orange peaks are those still remaining. H) CSD dimer colored by residue electrostatics, acidic surface adjacent to the PXVXL pocket is circled.

between the two components **(Figure 4.2A)**. We repeated this test with fully formed octamers and included a mutant I165E that is deficient in dimerization and thus cannot make the PXVXL binding surface **(Figure 4.2B)**. In this case, the dimer mutant was unable to co-elute with octamers, initially suggesting that binding the PGTVAL motif of H3 was necessary for complex formation.

To gain more insight into the interaction, we employed solution nuclear magnetic resonance (NMR) spectroscopy. We started with a peptide of the H3 αN helix previously used to determine a bound CSD crystal structure, where the otherwise disordered peptide adopts a beta strand inside the binding pocket **(Figure 4.2G)**. Mixing a stoichiometric amount of peptide and HP1α CSD dimer, we observed several significant chemical shift perturbations and many vanishing peaks, consistent with previously studies of the CSD with PXVXL peptides **(Figure 4.2C)**.[13] Next, we tested if this pattern is the same for folded histones, which would require a helix to beta strand structural transition. The binding signature appears to match the peptide, albeit there is even greater peak loss at the lower stoichiometric ratio used (1.0 CSD: 0.1 H3-H4 tetramer) **(Figure 4.2D)**. The loss of peak intensity commonly comes from an intermediate exchange rate between bound and unbound states that broadens the signal between each conformation. Alternatively, a complex may become too large to be detected by solution NMR. The peak loss with the small peptide suggests that this is a case of intermediate exchange and that the H3-H4 tetramer interacts with a much larger surface area of the CSD.

To ensure that the CSD binding is specific for the PXVXL-like sequence, we tested the mutant, W174A, which removes the affinity for PXVXL peptides. Repeating the same NMR experiment with the W174A mutant and the H3 αN helix peptide confirmed that this interaction is completely abrogated in the mutant **(Figure 4.2E)**. Surprisingly, the addition of H3-H4 tetramer returned the same binding signature with similar affinity as wildtype CSD **(Figure 4.2F).** This could be explained by the strong electrostatic interactions with the rest of the histone core dominating the interaction. Notably, the overall signal remaining is not generally broadened, which would occur if a stable complex formed with the core of the ~52 kDa H3-H4 tetramer. Looking at the structure of the CSD, the area which has the greatest signal loss from the spectra aligns with the PXVXL pocket and the acidic surface, while the peaks remaining in the spectra are opposite to the pocket or apart of the disordered c-terminal extension **(Figure 4.2G, H)**. We interpret these results to mean that the acidic surface of the CSD drives interaction with histone H3. This is supported by the relatively weak binding of the H3 αN helix compared to other peptides, as well as the loss of binding upon phosphorylation (reduction of positive charge) of the peptide. Extensive interactions throughout the CSD have been seen before for other PXVXL-bearing proteins,[17g] however, it is still unclear how the presence of DNA would influence these interactions.

We therefore transitioned our focus to mononucleosomes (MNs) which allowed us to detect binding using an electrophoretic mobility shift assay (EMSA). Here, the negative charge of DNA draws the nucleosomes down the gel and any change to the shape, size or charge of the molecule would shift the band. First, we compared wildtype HP1α and the W174A mutant for binding affinity to MNs, finding that the W174A mutation has no bearing on affinity for MNs **(Figure 4.3A)**. We sought binding with the CSD domain alone, however, even in conditions of ~1000 fold CSD excess, no interaction was detected. We repeated the EMSAs with H3 K9me3 MNs and a phosphorylated HP1α construct, as well as a chromodomain truncation **(Figure 4.S1)**. The presence of H3 K9me3 appears to slightly increase affinity for full-length HP1α but is insufficient to detect interaction from just the chromodomain. Notably, phosphorylated HP1α has significantly reduced affinity to MNs, confirming that much of the affinity is derived from the basic hinge domain and that the H3 tail interaction does not drastically expose the H3 αN helix for stable docking by the CSD.

Because the EMSAs may be too insensitive to weak interactions, we again turned to NMR to see if MNs could induce the same binding signature as the H3 αN helix and tetramer. We found that at a ratio 1.0 CSD to 0.8 MN was insufficient to detect an interaction **(Figure 4.3B, C)**. This sample has more H3 than the tetramer sample, but the wrapped DNA most likely precludes access to the H3 αN helices. To increase the likelihood of DNA unwrapping and H3 αN helix exposure, we prepared tetrasomes (TSs), which only contain the H3-H4 tetramer and are not canonically wrapped. Subjecting TSs to EMSAs and NMR with the HP1α CSD failed to produce any interaction **(Figure 4.S2)**. Therefore, the presence of DNA may be sufficient to prevent the necessary histone rearrangement for CSD docking. To test this again from the MN point of view, we used a pulldown assay that utilizes the high affinity of streptavidin for a biotin tag placed on the DNA of mononucleosomes. In this way, we could use high salt buffers that favor unwrapped MNs and may give a more physiological landscape of MN dynamics. However, once again the interaction by full-length HP1α or CSD was not improved with the high salt unwrapping condition **(Figure 4.3D)**. Overall, we were unable to detect a CSD-MN interaction and were unable to see a cooperative effect of

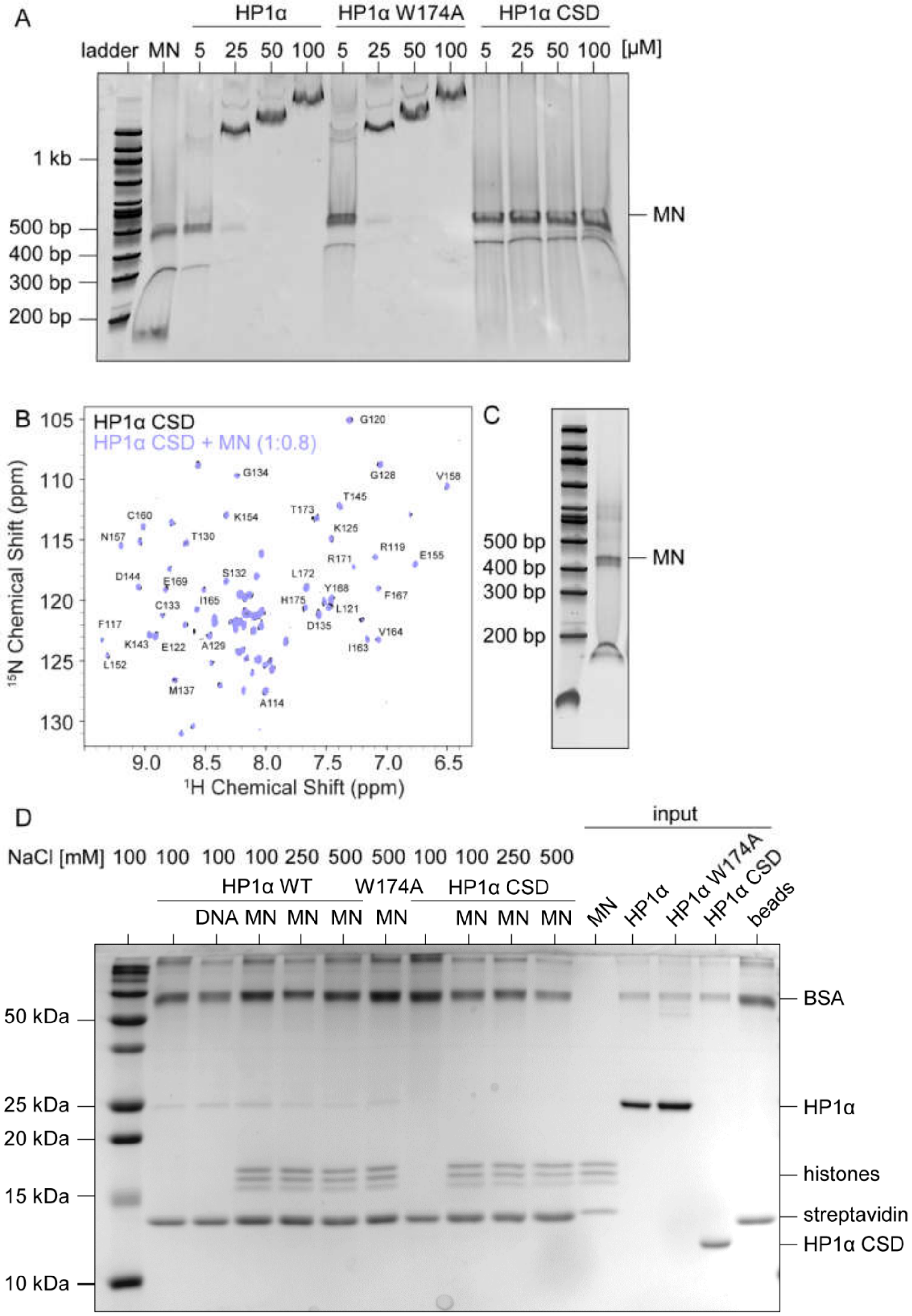


Figure 4.3. Chromoshadow domain binding to mononucleosomes. A) EMSA of 200 nM MN with HP1α, HP1α W174A, and HP1α CSD with protein concentration ranging from 5 to 100 μM.B) ^{1}H-^{15}N HSQC of HP1α CSD (black) and HP1α CSD 1:0.8 with MN (purple). C) Native polyacrylamide gel of MN used for NMR. D) SDS-PAGE of biotinylated MN pull-down. HP1α, HP1α W174A, and HP1α CSD were tested for binding in a range of NaCl concentrations.

having the full-length HP1 besides the DNA-hinge interaction. A limitation of our experiments may be the lack of other chromatin effectors that may act in concert with HP1α to make the H3 αN helix more accessible.

HP1 competition with chromatin effectors

We then aimed to harness the recently discovered MN binding mechanism of OCT4-SOX2 to activate CSD binding,[27] potentially offering functional mechanistic insight in an elusive HP1 function. To generate a docking site for OCT4-SOX2, we replaced the 601 sequence at SHL -6 with the recognition sequence for the OCT4 POU domain and SOX2 HMG domain **(Figure 4.4A)**. A more challenging undertaking is generating recombinant OCT4 and SOX2. Much of these proteins are disordered regions involved in activating downstream components, however, because we were primarily interested in the nucleosome interface, we constructed the DNA-binding domains only. While the recombinant SOX2 HMG domain behaved well, the OCT4 POU domain required several precautions **(Figure 4.S3A, C, D)**. The POU domain has high affinity for DNA regardless of sequence and the protein fold is unstable without its DNA interaction. For this reason, a GFP tag was added to the POU domain to assist with solubility and DNA was removed as a last step before experiments (more details can be found in the methods section and supplemental figure **4.S3**). Using these strategies, we were able to produce some active OCT4 and SOX2 which were able to cooperatively bind DNA with a preference for SHL -6 containing DNA **(Figure 4.4B)**. The OCT4/SOX2 complex was able to shift the SHL -6 MN, but we were unable to reproducibly achieve significant band shifts, likely due to a large population of misfolded OCT4/SOX2. Nonetheless, we were able to test binding competition via

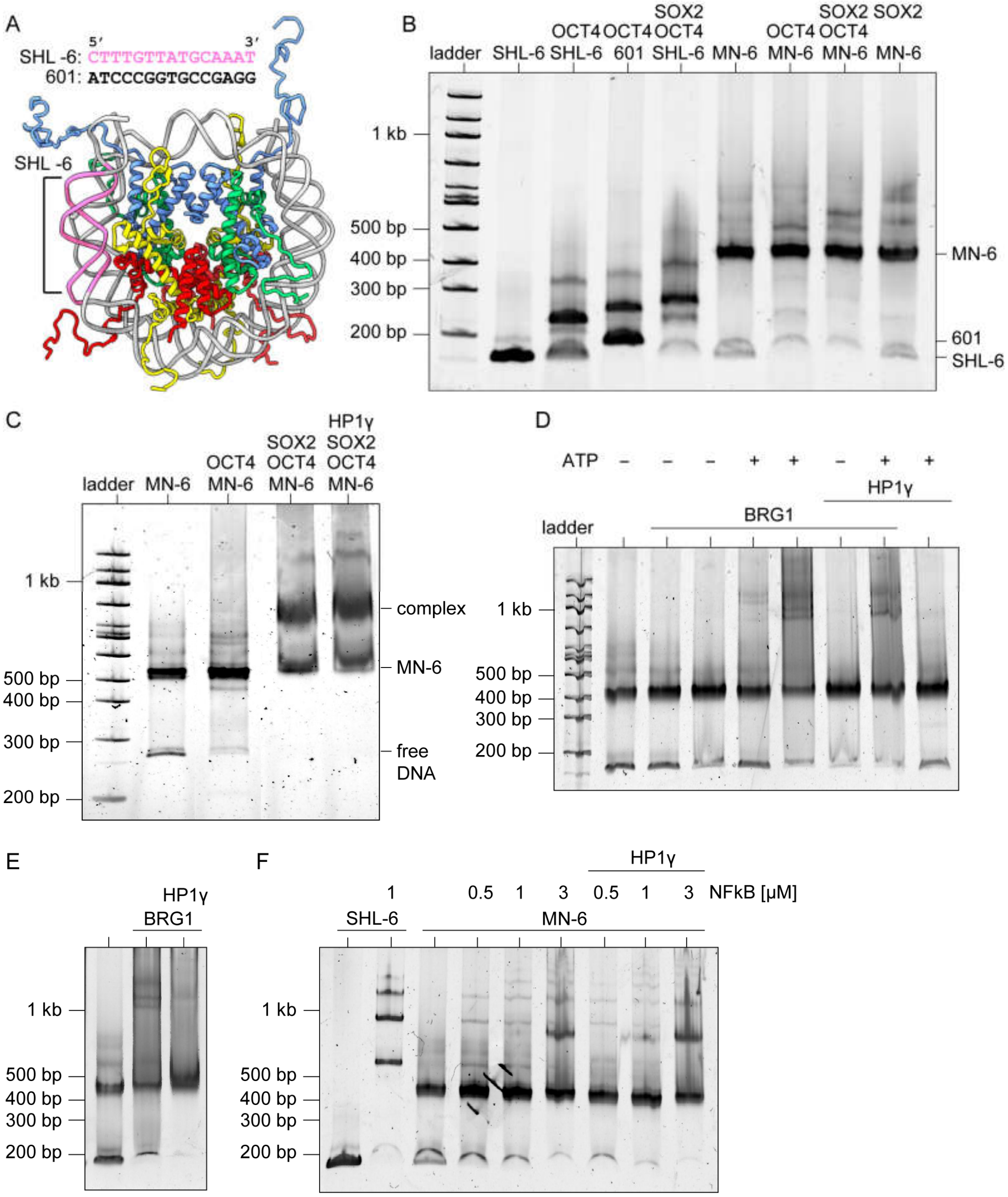


Figure 4.4. Binding relationships on the nucleosome surface. A) MN with the SHL-6 DNA region modified to the OCT4-SOX2 recognition sequence (PDB:1KX5). B) EMSA of 3 μM OCT4 and SOX2 with 300 nM 601 DNA, SHL-6 DNA, or MN SHL-6. C) EMSA of 1.05 μM MN SHL-6 with 10 μM OCT4 and SOX2, and 36 μM HP1γ. D) EMSA with 250 nM MN SHL-6 with 846 nM BRG1 ATPase and 50 μM HP1γ. E) EMSA with 100 nM MN SHL-6 with 1.12 μM BRG1 ATPase and 100 μM HP1γ. F) EMSA with 200 nM SHL-6 DNA or 500 nM MN SHL-6 mixed with a range of NFkB and 17 μM HP1γ.

EMSA **(Figure 4.4C)**. Here, we used HP1γ because of its cellular relationship to OCT4 and because it lacks DNA binding that would interfere with the EMSA. We did not detect any effect from HP1γ, where a band shift down would suggest the removal of OCT4-SOX2 and a shift up would be a complex between all 4 components.

After HP1γ and OCT4-SOX2 showed lack of binding competition, we tried another transcription factor known to bind the nucleosome, NFkB.[28] Again, HP1γ in excess was unable to assist or disrupt the complex of MN and NFkB **(Figure 4.4F)**.

Unable to show any effect of HP1γ on the purely structural effect of transcription factors, we sought to determine if an active process is required for HP1γ interaction. We used the ATPase domain (BRG1-C) of BRG1, which lacks the PXVXL motif, to test this hypothesis. Upon initial testing, BRG-1 was only able to form a complex with MN in the presence of ATP, and the addition of 50 μM HP1γ had no effect **(Figure 4.4D)**. Increasing HP1γ up to 100 μM appeared to disassemble the BRG1-MN complex, however, adding this much protein tends to smear lanes and is likely just an artifact **(Figure 4.4E)**. Altogether, we were unable to show a reproducible effect of HP1 on the binding of chromatin effectors to the nucleosome.

HP1 regulation of chromatin remodeling

We next wanted to probe the relationship of HP1 with chromatin effectors while bypassing the limitations of EMSA experiments which necessitates a strong interaction. We chose to focus on a chromatin remodeling assay that can test the impact of HP1 through a functional readout of remodeling. We started with designing a 205 bp H3 K9me3 nucleosome with a PstI cleavage site at base pairs 23-28. PstI has been regularly employed to cut exposed DNA in remodel assays.[29] Successful remodeling will shift the DNA up to 60 base pairs and expose the PstI site for cleavage **(Figure 4.5A)**. We were able to conduct this assay with three different families of remodelers; BRG1

(SWI/SNF), CHD1 (CHD), and ACF (ISWI). This set of remodelers exposes HP1α to many possible modes of action. HP1α has known interaction with both BRG1 and ACF,[12b, 17b] while CHD1 dramatically removes the DNA from the H3 αN helix.[30] First, we were able to determine conditions to achieve an intermediate level of remodeling by tuning time and remodeler concentration **(Figure 4.S4A)**, this dynamic range would best allow us to detect any remodeling inhibition. With chosen remodeler concentrations, we allowed remodeling to proceed for 30 minutes in the presence of PstI

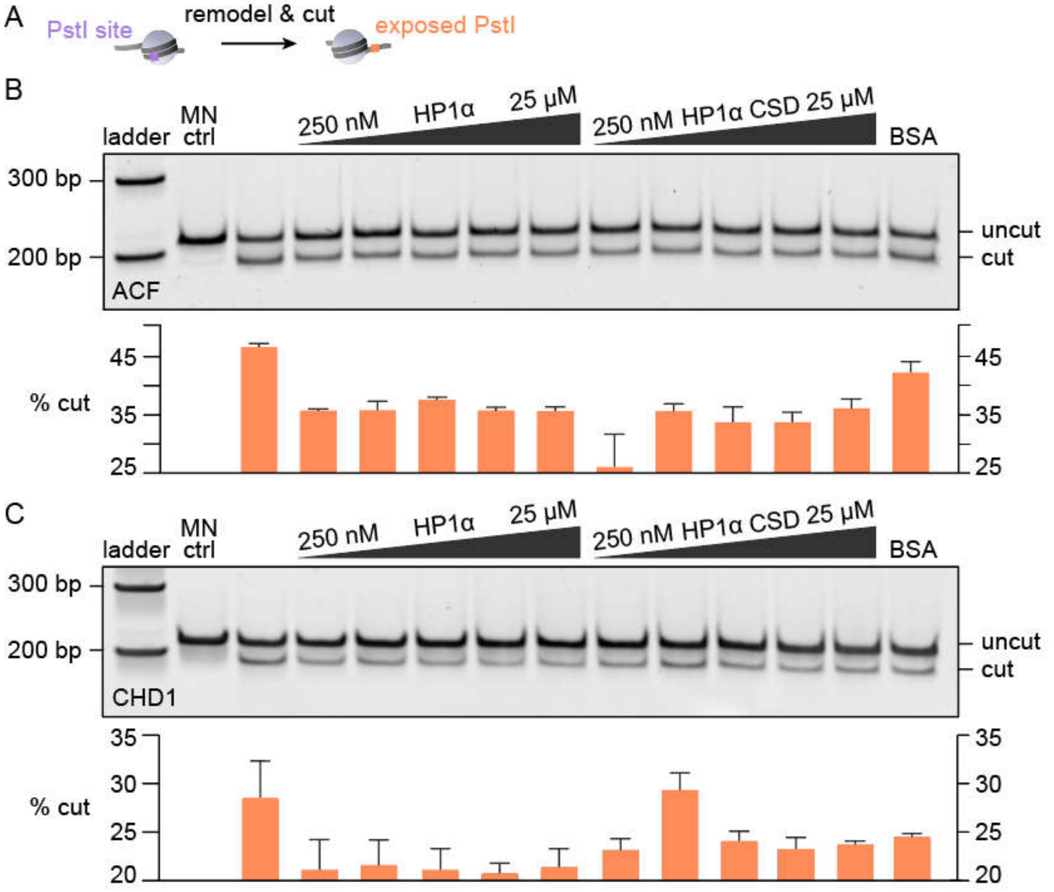


Figure 4.5. Mononucleosome remodeling in the presence of HP1. A) Scheme of the remodeling assay. PstI is available for restriction enzyme digest only upon remodeling. B) Above: Deproteinized DNA products of ACF remodeling assay with 50 nM ACF and 250 nM, 500 nM, 1 μM, 5 μM, and 25 μM HP1α or HP1α CSD. Below: Bar graph of average cutting percentages with standard error for each condition (n = 3). C) Same as (B) but remodeled with 250 nM CHD1 (n = 3).

and titrated in either HP1α or HP1α CSD **(Figure 4.5B, C)**. With the presence of H3 K9me3 and the known hinge interaction, we expect HP1α to have extensive contact with the nucleosomes during remodeling. Following the titration, we don't see a trend for repression or activation of either CHD1 or ACF. Due to insufficient enzyme quantity, we were unable to quantify remodeling with BRG1 in triplicate, but the result appears to be the same **(data not shown)**. The general repression of all samples with HP1α/CSD is a result of the protein buffer, we were able to show this same repression with buffer alone **(data not shown)**. These assays are highly sensitive to buffer conditions, yet are insensitive to the presence of excess levels of HP1α.

Perplexed with our nucleosome remodeling results in conflict with previous studies on HP1,[12b, 17b] we decided to probe for repression on 12-mer nucleosome arrays. This context may align better with the bridging model for HP1 function.[17b] For nucleosome arrays, we used the HpaII restriction enzyme site already within the 601 sequence, as this approach has been used previously to study ACF and HP1.[17b] In this case, there are 12 HpaII sites per DNA strand that create a ladder of products as each individual nucleosome is consecutively remodeled and cut **(Figure 4.6A)**. For this reason, we opted to conduct this assay as a time course to follow the band laddering. The 601 sequences aim to position each nucleosome 30 bp apart, however, equilibrium has slight variation that allows for some HpaII cutting even without remodeling. Nonetheless, a noticeable reduction of the complete 12-mer band is observed upon addition of CHD1 or BRG1, **(Figure 4.6B, C)**. ACF was not used here because ACF's function is to space nucleosomes evenly, which has little effect on these already regular

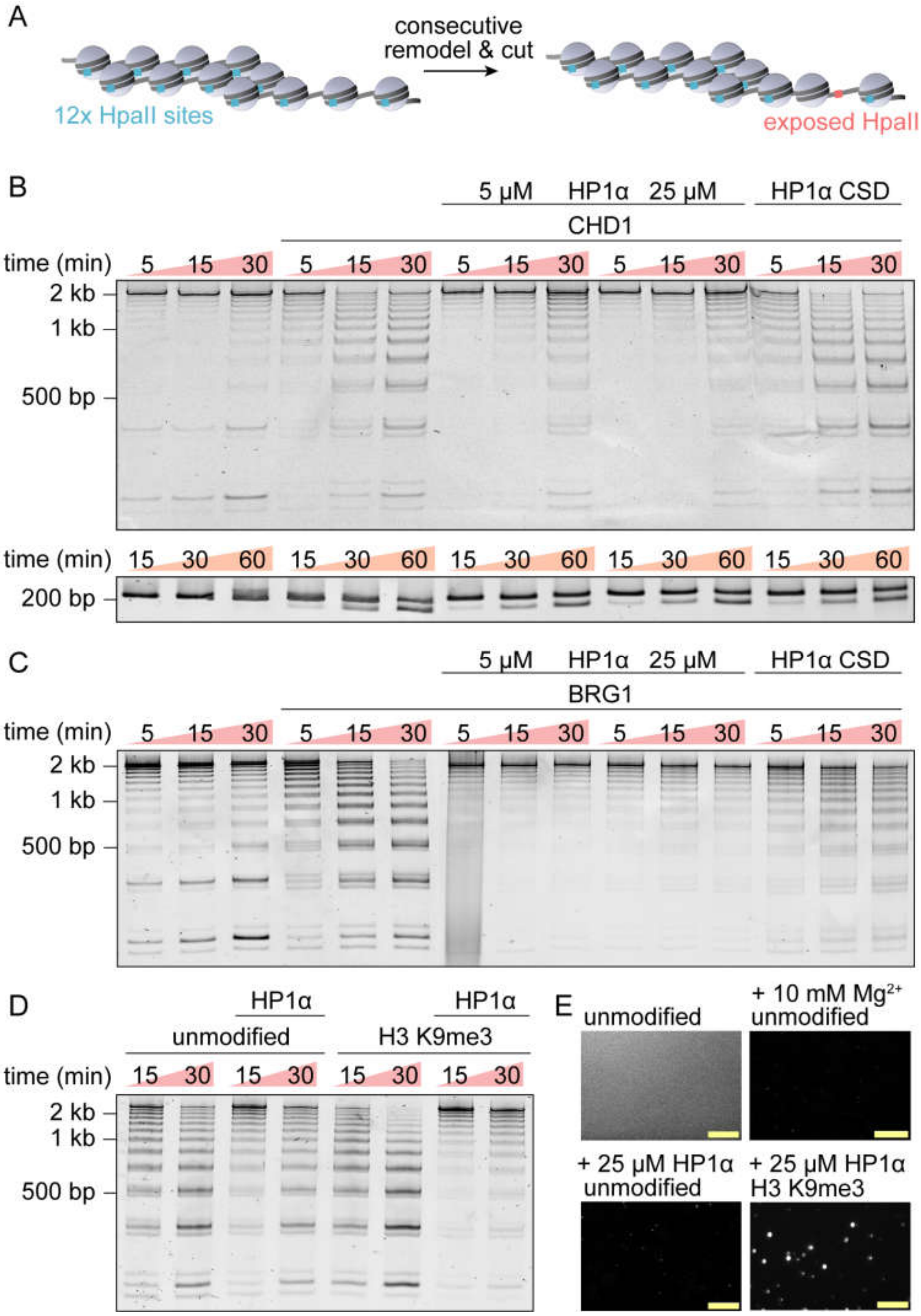


Figure 4.6. Nucleosome array remodeling in the presence of HP1. A) Nucleosome array remodeling assay scheme. Consecutive remodeling steps continually expose new HpaII sites to create a ladder of products. B) Above: Deproteinized nucleosome array remodeling products. Time points (5, 15 and 30 minutes) were taken for remodeling assays performed with 25 nM (nucleosome site concentration) chromatin, 500 nM CHD1 and 5 µM or 25 µM HP1α, or 25 µM HP1α CSD. Below: Mononucleosome remodeling assay matching the conditions above but with time points taken at 15, 30 and 60 minutes. C) Same as in (B) but remodeled with 66 nM BRG1. D) Deproteinized nucleosome array remodeling products for unmodified or H3 K9me3 arrays with or without 25 µM HP1α. E) Microscopy images of nucleosome arrays intercalated by YOYO-1 dye. Scale bars 20 µm.

12-mer arrays.[31] The addition of HP1α appears to significantly reduce the remodeling rate of both CHD1 and BRG1, while the CSD has either no effect on CHD1 or a minor effect on BRG1. A CHD1 MN time course remodeling assay confirms that this phenomenon only occurs with nucleosome arrays **(Figure 4.6B).** Because this was specific for the multi-nucleosome setting, we wondered if the bridging of nucleosomes through H3 K9me3 binding was required. Comparison of unmodified and H3 K9me3 arrays show drastically more repression with the lysine trimethylation **(Figure 4.6D)**. Looking at these samples under the microscope with a DNA intercalating dye, it is clear that the H3K9me3 arrays make large domains of collapsed chromatin consistent with previous studies **(Figure 4.6E)**,[11b, 32] however this has never been linked to remodeling efficacy. We used a two-step assay variation to show that HP1-mediated compaction is affecting remodeling and not the restriction enzyme (**Figure 4.S4B)**. Finally, we showed that the HP1α W174A mutant represses no differently than HP1α WT, suggesting that tethering via the H3 K9me3 mark is the dominating mode of repression **(Figure 4.S4C)**.

4.3 Discussion

In this chapter we were driven to discover how HP1 interacts with the nucleosome and represses other chromatin effectors as a result. We started with the interaction between the HP1α CSD and the H3 αN helix, showing that the CSD can bind to H3-H4 tetramers in the absence of DNA. While the structural switch between helix and beta strand seems possible in a tetramer, the binding appeared not to rely on the PXVXL binding pocket. Therefore, the main contacts between the CSD and H3, either on the tail or core, may be regions typically shielded by DNA or histone chaperones.

We were unable to show any binding between the CSD and nucleosomes. It is possible that because the 601 DNA is an artificially strong positioning sequence, it prohibits DNA dynamics that would allow for the CSD to bind. Our attempts to address this possibility with high salt buffers or tetrasomes showed no interaction. While the EMSAs we conducted require a high affinity interaction, the SEC co-elution between the CSD and tetramer show that these two components can maintain a stable complex. We therefore conclude that the CSD alone and full-length HP1α and γ cannot bind to the H3 αN helix in a chromatin setting. These results suggest that the recently discovered nucleosome loosening effect of Swi6 (HP1 homolog) may not involve the H3 αN helix, but may be dependent on liquid-liquid phase separation, or may be specific to the yeast homolog Swi6.[26a]

In a cellular chromatin environment, many chromatin effectors are simultaneously interfacing with the nucleosome. We intended to utilize the DNA unwrapping feature of OCT4 and SOX2 to induce HP1γ binding, but were unable to show a relationship between these chromatin effectors. This was the case as well for NFkB and the ATPase domain of BRG1. At this point, we have been unable to detect binding in the presence of DNA.

We employed functional chromatin remodeling assays to provide more insight. With these assays, an active ATP-driven process can open the nucleosome, and a functional readout is used rather than relying on binding detection. Again, HP1α initially displayed no effect on the remodeling of ACF, BRG1, or CHD1. This is in contrast to previous studies that suggest the act of remodeling may open up the nucleosome to release the H3 αN helix.[12b, 26b]

We transitioned the remodeling assay to nucleosome arrays to test if repression derives from HP1α working in concert with a crowd of nucleosomes and other HP1α molecules. The addition of HP1α to arrays significantly repressed chromatin remodeling by BRG1 and CHD1. This repression was dependent on H3 K9me3 which suggests HP1α mediates a compact state by transiently tethering arrays. Additionally, the CSD-PXVXL interaction was not required to repress remodeling. This mechanism fits well into the model of heterochromatin formation by liquid-liquid phase separation and polymer collapse.[33] We propose that the interaction with the H3 αN helix is not necessary for direct repression by HP1α, especially considering the PXVXL pocket is likely bound by other chromatin effectors which on their own can regulate phase separation.[33b, 34] We add onto the current model by positing that in addition to being able to organize the nucleus and dictate its material properties via the H3K9me3 mark, HP1α-mediated chromatin compaction may also be sufficient for chromatin remodeler inhibition.[33d, 35]

4.4 Future outlook

In this chapter we continued unpacking the direct biophysical effects of HP1 on active chromatin processes. The results presented here offer many springboards into new understanding. Fluorescently labeled remodelers can be used to test if chromatin remodelers are able to enter the chromatin foci created by HP1α. Nucleosome arrays with different patterns of H3 K9me3 marks can be used to learn if mark erasure at a nucleosome is sufficient to allow remodeling. Other factors such as histone modfiications, Mg^{2+} or histone H1 can also compact nucleosome arrays. Is there a difference in the way these compact arrays are interpreted in chromatin function and regulation processes?

It will be important to develop a protocol to routinely produce active OCT4 and BRG1. Both proteins are unstable through the purification process and limit assay reproducibility. Further, more complete constructs and complexes may be desired. For example, full length OCT4 potentially has a binding site for HP1γ, how does that affect their cooperation? Full length BRG1 has an HP1α binding site, but this site may be concealed in the physiological BAF complex.[36] BAF and other full remodeler complexes may have different relationships with HP1α.

NMR of the CSD should be continued with excess quantities of nucleosomes and tetrasomes to favor the formation of low population binding states if they exist. It will also be extremely powerful to determine the extent to which the H3 αN helix can unfold, since it unveils the possibilities of nucleosome dynamics and if CSD binding is ever feasible. There is evidence the H3 αN helix can become dynamic during unwrapping,[25a] but the extent and implications are unclear. While the histone cores are not detectable by solution NMR in the nucleosome, they may be in the tetramer alone. It would be informative to isotopically label H3 and H4 to determine which regions are contacting HP1 and if the helix does unfold. Overall, we hope to have inspired thought and provided groundwork for many future discoveries of the mechanisms underlying HP1 function.

4.5 Methods

Lab Techniques

All commonly used reagents were purchased from Sigma-Aldrich or Fisher Scientific. Isotopically enriched reagents were purchased from Cambridge Isotope Laboratories. Primer synthesis and gene synthesis were completed by Integrated DNA Technologies and Genewiz, respectively. PCR was performed using a Bio-Rad T100 thermocycler, while purification of the DNA relied on kits from New England BioLabs. Phusion polymerase and TEV protease were gifts from the Corbett Lab at UCSD. Dialysis kits were purchased from Thermo Fisher Scientific and protein concentrators were obtained from Sartorius. Reverse-phase (RP) HPLC was performed on a 2545 Binary Gradient Module Waters system equipped with a 2484 UV/vis detector. For prep-scale RP-HPLC purification, we relied on a Waters XBridge BEH C18 19 mm x 250 mm, 10 µm particle size column, while for analytical measurements we used a Symmetry300 C18 4.6 mm x 150 mm, 5 µm particle size column. HPLC solvent A contained 100% H_2O + 0.1% trifluoroacetic acid (TFA), while solvent B contained 100% acetonitrile + 0.1 % TFA. For size-exclusion chromatography, a Superdex 200 10/300 column (GE Healthcare Life Sciences) was used with an ÄKTA pure protein purification system (GE Healthcare Life Sciences). LC-ESI-TOF MS analysis was conducted on an Agilent 6230 Accurate-Mass TOFMS. Gel images were acquired using a camera and light box from Fotodyne Incorporated. Protein and fluorophore absorbances were measured using a Nanodrop One Spectrophotometer by Thermo Scientific.

Constructs

The HP1α gene was excised from a GST HP1α plasmid provided by Naoko Tanese[37] (Addgene plasmid # 24074 ; http://n2t.net/addgene:24074 ; RRID:Addgene_24074) and cloned into the vector backbone of a 2BT MacroLab plasmid (generously provided by Dr. Kevin Corbett) using NEBuilder® HiFi DNA Assembly Master Mix. The resulting plasmid contained a His_6-TEV-HP1α construct.

CK2 was created by excising CK2beta from the pAB46 plasmid, a gift from David Litchfield[38] (Addgene plasmid # 27085; http://n2t.net/addgene:27085; RRID:Addgene_27085) and cloned into a pCDF-Duet-CK2alpha plasmid generously provided by Dr. Neel Shah. The resulting plasmid contained CK2alpha and CK2beta constructs that could be co-expressed.

For *E.coli* expression, SOX2 and OCT4 genes were excised from pGem-Sox2 and pGem-OCT4, a gift from James Thomson[39] (Addgene plasmid # 16353; http://n2t.net/addgene:16353; RRID:Addgene_16353). These genes were placed into the 2BT plasmid using NEBuilder® HiFi DNA Assembly Master Mix.

For Sf9 baculoviral expression, a codon-optimized OCT4 gene was ordered from Gene Strings by GeneArt by ThermoFisher. This gene was placed into the pFastBac_HT_A plasmid using NEBuilder® HiFi DNA Assembly Master Mix.

Histone H3 K9C was created from a pET30-histone H3.1 plasmid with three mutations, C96A, C110A and K9C. All mutations were made using the NEBuilder® HiFi DNA Assembly Master Mix.

The following are the sequences of purified proteins used in this study:

HP1 alpha:

MKSSHHHHHHENLYFQ/SGKKTKRTADSSSSEDEEEYVVEKVLDRRVVKGQV
EYLLKWKGFSEEHNTWEPEKNLDCPELISEFMKKYKKMKEGENNKPREKSES
NKRKSNFSNSADDIKSKKKREQSNDIARGFERGLEPEKIIGATDSCGDLMFLM
KWKDTDEADLVLAKEANVKCPQIVIAFYEERLTWHAYPEDAENKEKETAKS*

HP1 alpha W174A:

MKSSHHHHHHENLYFQ/SGKKTKRTADSSSSEDEEEYVVEKVLDRRVVKGQV
EYLLKWKGFSEEHNTWEPEKNLDCPELISEFMKKYKKMKEGENNKPREKSES
NKRKSNFSNSADDIKSKKKREQSNDIARGFERGLEPEKIIGATDSCGDLMFLM
KWKDTDEADLVLAKEANVKCPQIVIAFYEERLTAHAYPEDAENKEKETAKS*

HP1 alpha CSD-CTE:

MKSSHHHHHHENLYFQ/SNDIARGFERGLEPEKIIGATDSCGDLMFLMKWKDT
DEADLVLAKEANVKCPQIVIAFYEERLTWHAYPEDAENKEKETAKS*

HP1 alpha CSD-CTE I165E:

MKSSHHHHHHENLYFQ/SNDIARGFERGLEPEKIIGATDSCGDLMFLMKWKDT
DEADLVLAKEANVKCPQIVEAFYEERLTWHAYPEDAENKEKETAKS*

HP1 alpha CSD-CTE W174A:

MKSSHHHHHHENLYFQ/SNDIARGFERGLEPEKIIGATDSCGDLMFLMKWKDT
DEADLVLAKEANVKCPQIVIAFYEERLTAHAYPEDAENKEKETAKS*

HP1 alpha CD:

MKSSHHHHHHENLYFQ/SEYVVEKVLDRRVVKGQVEYLLKWKGFSEEHNTW
EPEKNLDCPELISEFMKKYKKMKEG*

HP1 gamma:

MKSSHHHHHHENLYFQ/SGKKQNGKSKKVEEAEPEEFVVEKVLDRRVVNGK
VEYFLKWKGFTDADNTWEPEENLDCPELIEAFLNSQKAGKEKDGTKRKSLSD
SESDDSKSKKKRDAADKPRGFARGLDPERIIGATDSSGELMFLMKWKDSDEA
DLVLAKEANMKCPQIVIAFYEERLTWHSCPEDEAQ*

OCT4 POU:
MKSSHHHHHHENLYFQ/SDIKALQKELEQFAKLLKQKRITLGYTQADVGLTLG
VLFGKVFSQTTICRFEALQLSFKNMCKLRPLLQKWVEEADNNENLQEICKAET
LVQARKRKRTSIENRVRGNLENLFLQCPKPTLQQISHIAQQLGLEKDVVRVWF
CNRRQKGKRSSS*

OCT4 FL 5CS (all POU domain cysteines mutated to serine):
MKSSHHHHHHENLYFQ/SMAGHLASDFAFSPPPGGGGDGPGGPEPGWVDPRT
WLSFQGPPGGPGIGPGVGPGSEVWGIPPCPPPYEFCGGMAYCGPQVGVGLVP
QGGLETSQPEGEAGVGVESNSDGASPEPCTVTPGAVKLEKEKLEQNPEESQDI
KALQKELEQFAKLLKQKRITLGYTQADVGLTLGVLFGKVFSQTTISRFEALQL
SFKNMSKLRPLLQKWVEEADNNENLQEISKAETLVQARKRKRTSIENRVRGN
LENLFLQSPKPTLQQISHIAQQLGLEKDVVRVWFSNRRQKGKRSSSDYAQRED
FEAAGSPFSGGPVSFPLAPGPHFGTPGYGSPHFTALYSSVPFPEGEAFPPVSVTT
LGSPMHSN*

eGFP-TEV-OCT4 FL:
MKSSHHHHHHGMVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATY
GKLTLKFICTTGKLPVPWPTLVTTLTYGVQCFSRYPDHMKQHDFFKSAMPEG
YVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYN
YNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLP

DNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITLGMDELYKENLYFQ/SDIK
ALQKELEQFAKLLKQKRITLGYTQADVGLTLGVLFGKVFSQTTICRFEALQLS
FKNMCKLRPLLQKWVEEADNNENLQEICKAETLVQARKRKRTSIENRVRGNL
ENLFLQCPKPTLQQISHIAQQLGLEKDVVRVWFCNRRQKGKRSSS*

TEV eGFP-OCT4 FL (from Sf9):
MSYYHHHHHHDYDIPTTENLYFQ/GAMDPVSKGEELFTGVVPILVELDGDVN
GHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTLVTTLTYGVQCFSRYPD
HMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGI
DFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLAD
HYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITLG
MDELYKELMAGHLASDFAFSPPPGGGGDGPGGPEPGWVDPRTWLSFQGPPG
GPGIGPGVGPGSEVWGIPPCPPPYEFCGGMAYCGPQVGVGLVPQGGLETSQPE
GEAGVGVESNSDGASPEPCTVTPGAVKLEKEKLEQNPEESQDIKALQKELEQF
AKLLKQKRITLGYTQADVGLTLGVLFGKVFSQTTICRFEALQLSFKNMCKLRP
LLQKWVEEADNNENLQEICKAETLVQARKRKRTSIENRVRGNLENLFLQCPK
PTLQQISHIAQQLGLEKDVVRVWFCNRRQKGKRSSSDYAQREDFEAAGSPFSG
GPVSFPLAPGPHFGTPGYGSPHFTALYSSVPFPEGEAFPPVSVTTLGSPMHSN*

*SOX2 HMG***:**
MKSSHHHHHHENLYFQ/SSPDRVKRPMNAFMVWSRGQRRKMAQENPKMHN
SEISKRLGAEWKLLSETEKRPFIDEAKRLRALHMKEHPDYKYRPRRKTKT*

Histone H3 K9C:
ARTKQTAR**C**STGGKAPRKQLATKAARKSAPATGGVKKPHRYRPGTVALREIR

RYQKSTELLIRKLPFQRLVREIAQDFKTDLRFQSSAVMALQEAAEAYLVGLFE

DTNLAAIHAKRVTIMPKDIQLARRIRGERA*

Histone H3 αN helix peptide:

KKPHRPGTVALREIRRYQKST

HP1 expression and purification

The following procedure could be used for all constructs and truncations of HP1 used in this study. BL21(DE3)-Rosetta cells were transformed with the HP1 (and CK2 for pHP1α) plasmids. Pre-cultures were grown at 37 °C overnight under ampicillin antibiotic selection in LB. Pre-cultures were pelleted in the morning, LB was removed and the pellet was resuspended in ^{13}C, ^{15}N enriched M9 media and added to 1L ^{13}C, ^{15}N enriched M9 expression cultures for NMR samples, or LB for rest of samples. Many of the following steps have been previously described for HP1α.[33b, 40] Cells were grown at 37 °C until OD600 = 0.6 (~12 hours), transferred to 18 °C and induced with 0.3 mM IPTG. Cells were collected 20 hours later by centrifuging at 5,000 xg. The cell pellet was resuspended in lysis buffer (1x PBS, pH 7.4, 300 mM NaCl, 10% glycerol, 7.5 mM imidazole, Roche protease inhibitor tablet) and sonicated at 4 °C. Cell debris was cleared by centrifuging for 30 minutes at 30,000 xg. Roughly 5mL of Ni-NTA resin (per 1L culture) was added to the lysate and rotated at 4 °C for one hour. The lysate-resin mixture was loaded into a Bio-Rad Econo-Column and washed with wash buffer (1x PBS, pH 7.4, 300 mM NaCl, 10% glycerol, 7.5 mM imidazole). Elution buffer (20 mM HEPES, pH 7.2, 300 mM KCl, 400 mM imidazole, 1 mM DTT) was then ran over the column. The resulting elution was incubated with TEV-protease during dialysis against imidazole-free buffer overnight at 4 °C. The sample was analyzed by SDS-PAGE to ensure the complete removal of the His tag. The HP1α solution was adjusted to 6 M Guanidinium HCl concentration and filtered. This solution was loaded onto a Waters XBridge BEH C18 prep-size reverse-phase HPLC column and fractions were collected using a gradient of 10%-60% solvent B. HP1α eluted around 40% solvent B and due to protein retention on the column, multiple runs were required to collect the protein. The elution product was lyophilized and stored at -80 °C. Purity was assessed by SDS-PAGE and analytical RP-HPLC.

OCT4 and SOX2 expression and purification from E.coli

eGFP-OCT4, and SOX2 were expressed and purified similarly, following guidance of previously published *E.coli* strategies.[41] Pre-cultures were grown at 37 °C overnight under ampicillin antibiotic selection in LB. Pre-cultures were used as seeds for 1 L flasks which were induced with 0.05 mM IPTG at OD = 0.6-1.0 and expressed overnight at 20 °C. OCT4 without eGFP formed inclusion bodies, however, eGFP-OCT4 and SOX2 remain soluble. Cell pellets were resuspended at 4 °C in 20 mM HEPES pH 7.2, 500 mM NaCl, 100 mM KCl, 10 mM imidazole, 10% glycerol, and 1 Pierce protease inhibitor tablet. Resuspension was sonicated, centrifuged at 30,000g and the supernatant was bound to Ni-NTA beads for 1 hour. High salt wash buffer was used to remove non-specific DNA, 20 mM HEPES pH 7.2, 1.5 M NaCl, 250 mM KCl, and 40 mM imidazole. Finally, protein was eluted with 20 mM HEPES pH 7.2, 1 M NaCl, and 400 mM imidazole. The elution A260/A280 shows that DNA is co-eluted in OCT4 samples, but not SOX2. The elution was then ran on Superdex 75 10/300 GL in 20 mM HEPES pH 7.2, 1 M NaCl, 1 mM DTT. SOX2 was separated from DNA, whereas OCT4 co-eluted with DNA. SOX2 was then concentrated and 10% glycerol was added before snap-freezing and storage at -80 °C. Caution that the frozen aliquots had less binding efficiency than freshly purified SOX2. The eGFP-OCT4 sample required DNA digestion by MNase to remove DNA from OCT4. eGFP-OCT4 post-digestion can be separated using a Ni-NTA column and must be used immediately, as the protein aggregates without the presence of DNA.

OCT4 (POU domain) purification and refold

This procedure was based loosely on previous work.[42] Pre-cultures were grown at 37 °C overnight under ampicillin antibiotic selection in LB. Pre-cultures were used as seeds for 1 L flasks which were induced with 0.05 mM IPTG at OD = 0.6-1.0 and expressed overnight at 20 °C. An inclusion body purification was used, cells were pelleted at 4,000 g and resuspended in 20 mM HEPES pH 7.4, 250 mM NaCl, 10% glycerol, 1 mM BME and 1 Pierce protease inhibitor tablet. Resuspension was sonicated and centrifuged at 30,000 xg. The pellet was resuspended in 6M G-HCl, 20 mM HEPES pH 7.4, 250 mM NaCl, 10 mM imidazole, and rotated at 4 °C until dissolved. Sample was spun at 30,000 g to remove undissolved material. The supernatant was bound to Ni-NTA beads for 1 hour, beads were washed with 6M G-HCl, 20 mM HEPES pH 7.4, 40 mM Imidazole, 250 mM NaCl, and eluted with 6M G-HCl, 20 mM HEPES pH 7.4, 400 mM Imidazole, 250 mM NaCl, 1 mM TCEP. OCT4 was then refolded against 20 mM HEPES pH 7.2, 150 mM NaCl, 1 mM DTT. Several refold strategies were tested, rapid dilution of sample into a large volume of 0 M G-HCl, slow gradient dialysis, and an extremely dilute refold. Results from dilute refolding (< 500 nM OCT4) showed best OCT4 recovery. (**Supplementary Figure 3)**. Because of OCT4 aggregation, the sample reduces in purity and can be concentrated, filtered, and loaded onto Superdex 75 GL

10/300 column in dialysis buffer for purification. OCT4 should be used fresh after purification.

eGFP-OCT4 expression and purification from Sf9

Expression from *Sf9* cells was utilized to assist the folding of OCT4.[27] 1 µg of His-eGFP-OCT4 bacmid (generated through DH10Bac cells) was transfected into *Sf9* cells with a Cellfectin/DNA mix for 6 hours. Cellfectin was removed, then cells incubated for 5 days at 27 °C. The media was collected as P1, and used at a 1:20 volume ratio to generate P2 which also incubated for 5 days. P2 was used at a 1:2 volume ratio to generate P3, which showed green fluorescence 24 hours after infection. P3 media was collected at 108 hours and a 1:20 volume ratio was used to generate 20 mL of P4. P4 was collected at 48 hours, 0.1% BSA was added and the viral titer was filtered. For protein production, a suspension of 1.5 million/mL cells was infected with 5 mL of P4 and expressed for 48 hours shaking at 27 °C.

The cells were collected and spun down at 4,000g to a visibly green pellet. The pellet was resuspended in lysis buffer (50 mM Tris pH 7.5, 1 M NaCl, 100 mM PMSF, 0.25 mM DTT, 1 Pierce protease inhibitor tablet). The suspension was sonicated at 4 °C, centrifuged at 30,000g, and the supernatant was incubated with Ni-NTA beads for 15 minutes at 4 °C. The slurry was washed with wash buffer (20 mM Tris pH 7.5, 1 M NaCl, 20 mM imidazole) and eluted with 20 mM Tris pH 7.5, 500 mM NaCl, 300 mM imidazole. Resulting protein ran small on SDS-PAGE and was unable to bind DNA, possible truncations or unfolding likely occur during the expression and purification.

Commercial and gifted nucleosome remodelers and peptides

BRG1 used for binding competition assays on the mononucleosome was a BRG1-C construct gifted by Zelin Shan and Sriram Aiyer from the Lyumkis Lab at Salk Institute. This construct was residues 558-1647 which contains the ATPase domain but lacks the purported HP1 binding domain.

NFkB used for binding competition assays was gifted by Hannah Baughman from the Komives Lab at UC San Diego.

BRG1 used for remodeling assays was produced by EpiCypher (Catalog No 15-1014), recombinant full length human BRG1 (SMARCA4) was expressed in *Sf9* cells.

ACF used for remodeling assays was produced by Active Motif (Catalog No: 31509), this was a recombinant complex of *Drosophila Melanogaster* Acf1 and Iswi expressed in *Sf9* cells.

CHD1 used for remodeling assays was produced by Active Motif (Catalog No: 81607), recombinant full length human CHD1 was expressed in *Sf9* cells.

The H3 αN helix peptide used for solution NMR was synthesized by ABclonal Science as crude product, then purified by RP-HPLC on a C18 semi-prep column.

HP1 preparation

HP1 was prepared the same way as done before,[40] for all constructs and truncations used here. Preparation starts by refolding the lyophilized stock in resuspension buffer (20 mM HEPES pH 7.2, 6 M G-HCl, 20 mM KCl, 2 mM DTT), dialyzing against 1 M Guanidinium-HCl for 5 hours, then dialyzing against guanidinium-free buffer with 300 mM KCl (pHP1α) or 150 mM KCl (all other constructs) overnight at 4 °C. We routinely achieved 70% yield during the refolding process with proper refolding confirmed by solution NMR. The dialysis product was first centrifuged at 4,000 x *g* to remove any precipitate, then concentrated using Sartorius centrifugal concentrators. The sample was concentrated to completion and either stored at -80 °C with 10% glycerol or used immediately. Assay buffers using HP1 were typically 20 mM HEPES pH 7.2, 150 mM KCl, 1 mM DTT, unless specified otherwise.

Histone H3 methyl-lysine installation

To install the methyl-lysine analog we followed a previously established protocol.[43] Briefly, histone H3 K9C, C110A was expressed, purified and lyophilized as previously described.[44] 5 mM of H3 K9C was completely reduced at 50 ˚C for an hour in 1M HEPES pH 7.8, 4M GuanidiniumHCl, 10 mM D/L-methionine and 20 mM DTT. Afterwards, 400 mM of (2-bromoethyl)-trimethylammonium bromide (Sigma) was added to the solution. The sample was kept in the dark at 50 ˚C and occasionally inverted to mix. After 2.5 hours, 10 mM of fresh DTT was added, and the reaction was allowed to proceed for another 2.5 hours. The reaction was finally quenched with 700 mM 2-mercaptoethanol. The product was purified by reverse-phase HPLC and confirmed by ESI-TOFMS.

Octamer and 12-mer nucleosome array design and assembly

DNA and histones were prepared using established protocols.[45] 12x601 DNA (12 repeats of the 601 DNA sequence separated by 30 bp linkers) for the 12-mer array and MMTV for buffer DNA were expressed in DH5α cells. Plasmids were digested and the DNA fragments were purified by phenol-chloroform extraction and ethanol precipitation. The DNA pellets were redissolved in 10 mM Tris pH 8.0, 0.1 mM EDTA buffer and stored at –20 ˚C.

The 601-site sequence is: 5’CTGGAGAATCCCGGTGCCGAGGCCGCTCAATTGGTCGTAGACAGCTCTAGCACCGCTTAAACGCACGTACGCGCTGTCCCCCGCGTTTTAACCGCCAAGGGGATTACTCCCTAGTCTCCAGGCACGTGTCAGATATATACATCCTGTGCATGTAAGATCCAGTACTACGCGGCCGCC3’

Extinction coefficient A260 = 2822957 for a single nucleosome site within the 12-mer DNA.

The MMTV sequence is: 5’ACTTGCAACAGTCCTAACATTCACCTCTTGTGTGTTTGTGTCTGTTCGCCATCCCGTCTCCGCTCGTCACTTATCCTTCACTTTCCAGAGGGTCCCCCCGCAGACCCCGGCGACCCTGGTCGGCCGACTGCGGCACAGTTTTTTG3’

Histones were expressed in BL21(DE3)-Rosetta cells, purified from inclusion bodies by reverse-phase HPLC and lyophilized. The histones were co-dissolved in 6M Guanidinium-HCl, 20mM Tris pH 7.5, then dialyzed against 0 M Guanidinium-HCl to refold and form octamers that were further purified with a size-exclusion Superdex 200 10/300 column (GE Healthcare Life Sciences) as in published protocols.[46] Octamers were stored in buffer containing 1M NaCl and 50% glycerol at –20 ˚C.

12-mer nucleosome arrays were prepared using previously described protocols.[44] Briefly, 1.6:1 12-mer DNA to octamer and 0.3:1 MMTV DNA to octamer were added to 2M TEK buffer (2 M KCl, 10 mM Tris pH 7.5, 0.1 mM EDTA) and 2 µM octamers in 10 mL of solution. Samples were dialyzed into 10 mM TEK buffer (10 mM KCl), precipitated by the addition of 4 mM MgCl2 and resuspended in assay buffer (20 mM HEPES pH 7.2, 75 mM KCl, 2 mM DTT, 0.01% NaN3) to be used immediately. Chromatin arrays were analyzed by native 1% agarose/2% polyacrylamide (APAGE) gels stained with SyBr Gold (Life Technologies).

Mononucleosome and tetrasome DNA design and production

The DNA used for mononucleosome and tetrasome assembly is 147 bp with an OCT4-SOX2 recognition site (bold):

5’CTGGAGA**CTTTGTTATGCAAAT**CCGCTCAATTGGTCGTAGACAGCTCTA GCACCGCTTAAACGCACGTACGCGCTGTCCCCCGCGTTTTAACCGCCAAGG GGATTACTCCCTAGTCTCCAGGCACGTGTCAGATATATACATCCTGT3’

Extinction coefficient A260 = 2411308. The forward primer of 5CTGGAGACTTTGTTATGCAA3 and reverse primer of 5ACAGGATGTATATATCTGACAC3 were used. For biotinylated nucleosomes, biotin-5CTGGAGACTTTGTTATGCAA3 was used.

The DNA used for mononucleosome remodelling assays is 205 bp, the linker extension (bold) is the continued 601 sequence[47] and PstI cut site (underlined text):

CTGGAGAATCCCGGTGCCGAGG<u>CTGCAG</u>AATTGGTCGTAGACAGCTCTAG CACCGCTTAAACGCACGTACGCGCTGTCCCCCGCGTTTTAACCGCCAAGGG GATTACTCCCTAGTCTCCAGGCACGTGTCAGATATATACATCCTGT**GCATG TATTGAACAGCGACCTTGCCGGTGCCAGTCGGATAGTGTTCCGAGCTC CCTGT**

Extinction coefficient A260 = 3379062. The forward primer of 5CTGGAGAATCCCGGTGC3 and reverse primer of 5ACAGGGAGCTCGGAACAC3 were used.

These DNA strands were produced by PCR, simultaneously reacting 80 tubes of 100 μl to generate ~0.5 mg of DNA. Standard Phusion PCR conditions were used with the addition of 3 times the concentration of dNTPs (final concentration of 0.6 mM of each nucleotide), a final primer concentration of 200 nM, and a template concentration of .02 ng/μl. The annealing temperature was 58 ˚C and extension time was 20 seconds. PCR product was purified using Zymogen Midiprep DNA purification columns. Pure DNA was eluted in TE 10/0.1 buffer and stored at -20 ˚C.

Mononucleosome and tetrasome assembly

The mononucleosome/tetrasome samples were prepared similarly to arrays with some modifications. Briefly, 1.8:1 DNA to octamer/tetramer were added to 2M TEK buffer (2 M KCl, 10 mM Tris pH 7.5, 0.1 mM EDTA) with a target of 0.3 μM octamer/tetramer in up to 20 mL of solution. Samples were dialyzed into 1.4 mM TEK which was then slowly brought to 500 mM TEK buffer (500 mM KCl) by pump. The dialysis bag was then switched to 10 mM TEK overnight. Samples were centrifuged at 21,000g to remove precipitate and were analyzed by native 5% polyacrylamide gels in TBE buffer stained with SyBr Gold (Life Technologies). For NMR samples, nucleosome/tetrasomes were ran over FPLC in 10 mM TEK on a Superdex 200 GL 10/300 column. The octamer, the complex, and the free DNA overlap in elution, so this was primarily used to remove any free histones that could complicate NMR analysis. Otherwise, no further purification was done after centrifugation. Tetrasomes were confirmed by reference to gel retention behavior described by Morrison et al.,[48] and kept at low salt (10 mM KCl) during all experiments.

Solution NMR spectroscopy

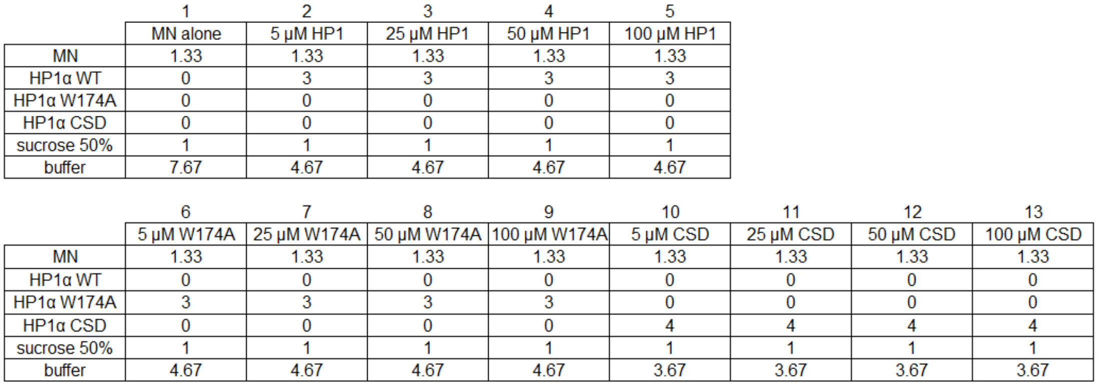
Sample	Temp (K)	Scans	Points/Acq. Direct	Points/Acq. Indirect	Figure
20 μM ^{15}N HP1α CSD (+ 8 μM H3 MN)	298	184	2048/90 ms	96/16.4 ms	3B
100 μM ^{15}N HP1α CSD (+ 50 μM H3 peptide, 5 μM tetramer)	298	24	2048/90 ms	96/16.4 ms	2C, 2D
100 μM ^{15}N HP1α CSD W174A (+ 50 μM H3 peptide, 5 μM tetramer)	298	24	2048/90 ms	96/16.4 ms	2E, 2F
50 μM ^{15}N HP1α CSD (+ 5 μM tetrasome)	293	32	2048/90 ms	128/21.9 ms	S2

Experiments were performed using a Bruker Avance Neo 800 MHz (1H Larmor frequency) NMR spectrometer equipped with a triple resonance TCO cryoprobe. Chemical shifts were referenced to the TMS frequency. The ^{1}H-^{15}N TROSY-HSQC spectra were acquired using the standard Bruker trosyf3gpph19 pulse sequence. Experiments used the following parameters: interscan delay of 1.25 s, ^{1}H center frequency of 4.7 ppm, ^{15}N center frequency of 120 ppm, and JNH = 90 Hz. All samples were prepared in 20 mM HEPES pH 7.2, 75mM KCl (10 mM KCl for tetrasome sample), 1 mM DTT with 10% D_2O.

Electrophoretic mobility shift assays (EMSAs) **Figure 2 A/B**

Samples were mixed to a final volume of 10 µl at room temperature and incubated for 10 minutes before running EMSA. All samples were set a final buffer concentration of 20 mM HEPES pH 7.2, 100 mM KCl, 1 mM DTT, and 10% sucrose to assist with EMSA loading. Protein samples were serially diluted before adding. Mononucleosome concentration was 200 nM in all samples, from which 1 µl (~20 ng) was loaded on gel. EMSAs were ran with 0.5x Tris/Borate/EDTA (TBE) buffer at 4 °C using power cell conditions of 120 V for 45 minutes. Gels were strained with 1:10,000 SybrGold for 10 minutes, then imaged using a Typhoon FLA 7000 gel imager with a 532 nm laser.

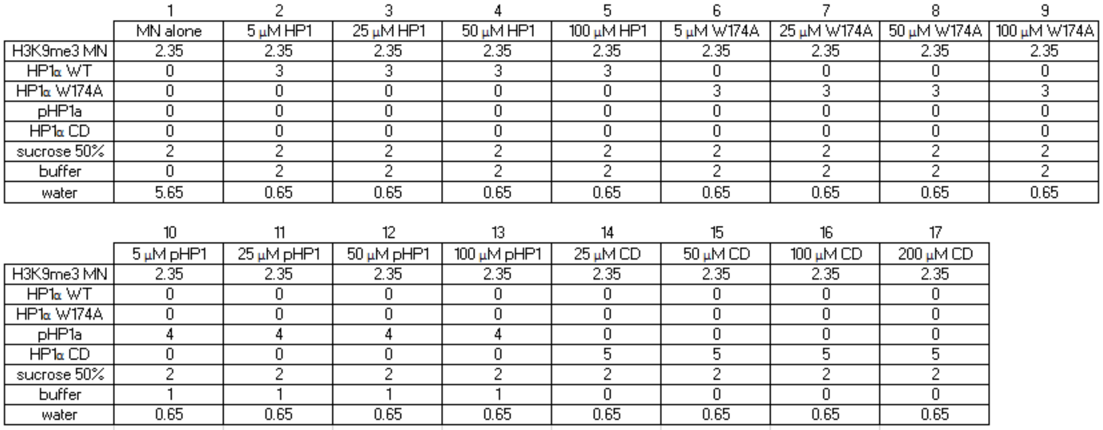
Table for Figure 1A:



	1	2	3	4	5
	MN alone	5 µM HP1	25 µM HP1	50 µM HP1	100 µM HP1
MN	1.33	1.33	1.33	1.33	1.33
HP1α WT	0	3	3	3	3
HP1α W174A	0	0	0	0	0
HP1α CSD	0	0	0	0	0
sucrose 50%	1	1	1	1	1
buffer	7.67	4.67	4.67	4.67	4.67

	6	7	8	9	10	11	12	13
	5 µM W174A	25 µM W174A	50 µM W174A	100 µM W174A	5 µM CSD	25 µM CSD	50 µM CSD	100 µM CSD
MN	1.33	1.33	1.33	1.33	1.33	1.33	1.33	1.33
HP1α WT	0	0	0	0	0	0	0	0
HP1α W174A	3	3	3	3	0	0	0	0
HP1α CSD	0	0	0	0	4	4	4	4
sucrose 50%	1	1	1	1	1	1	1	1
buffer	4.67	4.67	4.67	4.67	3.67	3.67	3.67	3.67

Table for Figure 1B:



	1	2	3	4	5	6	7	8	9
	MN alone	5 µM HP1	25 µM HP1	50 µM HP1	100 µM HP1	5 µM W174A	25 µM W174A	50 µM W174A	100 µM W174A
H3K9me3 MN	2.35	2.35	2.35	2.35	2.35	2.35	2.35	2.35	2.35
HP1α WT	0	3	3	3	3	0	0	0	0
HP1α W174A	0	0	0	0	0	3	3	3	3
pHP1a	0	0	0	0	0	0	0	0	0
HP1α CD	0	0	0	0	0	0	0	0	0
sucrose 50%	2	2	2	2	2	2	2	2	2
buffer	0	2	2	2	2	2	2	2	2
water	5.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65

	10	11	12	13	14	15	16	17
	5 µM pHP1	25 µM pHP1	50 µM pHP1	100 µM pHP1	25 µM CD	50 µM CD	100 µM CD	200 µM CD
H3K9me3 MN	2.35	2.35	2.35	2.35	2.35	2.35	2.35	2.35
HP1α WT	0	0	0	0	0	0	0	0
HP1α W174A	0	0	0	0	0	0	0	0
pHP1a	4	4	4	4	0	0	0	0
HP1α CD	0	0	0	0	5	5	5	5
sucrose 50%	2	2	2	2	2	2	2	2
buffer	1	1	1	1	0	0	0	0
water	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65

Nucleosome binding competition EMSAs

Figure 3B: Binding assays with OCT4 and SOX2 were performed as follows; 20 µl of 300 nM of DNA (147 bp SHL -6 or 177 bp 601 DNA) or MN with SHL -6 (MN -6) site was mixed with nothing, or 3 µM of OCT4 and/or SOX2. Binding buffer was 20 mM Tris pH 7.0 RT, 50 mM NaCl, 10 mM KCl, 1 mM $MgCl_2$, 0.1 mg/ml BSA, and 1 mM DTT. Samples were incubated at room temperature for 1 hour, brought to 5%

sucrose then 4 μl loaded on a 5% polyacrylamide gel and ran in 0.5x TBE buffer at 4 °C at 150 V for 1 hour.

Figure 3C: Binding assays with OCT4, SOX2, and HP1γ were performed as follows; 10 μl of 1.05 μM of MN -6 in binding buffer was mixed with 10 μM of OCT4, 20 μM of SOX2, and 36 μM of HP1γ. Samples were incubated for 30 minutes at room temperature before loading 1.5 μl of mixture on a 5% polyacrylamide gel as described above.

Figure 3D: Binding assays with BRG1 and HP1γ were performed as follows, 10 μl of 250 nM MN -6 in buffer (20 mM HEPES, pH 7.2, 150 mM NaCl, 1 mM TCEP, 2 mM $MgCl_2$, 10% glycerol, +/- 2 mM ATP) was mixed with 846 nM BRG1 and 50 μM of HP1γ. Samples were incubated at 4 °C for 30 minutes before loading 4 μl on a 5% polyacrylamide gel as described above.

Figure 3E: Binding assays with BRG1 and HP1γ were performed as follows, 10 μl of 125 nM MN -6 in buffer (20 mM HEPES, pH 7.2, 150 mM NaCl, 1 mM TCEP, 2 mM $MgCl_2$, 10% glycerol, +/- 2 mM ATP) was mixed with 1.12 μM BRG1 and 100 μM of HP1γ. Samples were incubated at 4 °C for one hour before loading 7 μl on a 5% polyacrylamide gel as described above.

Figure 3F: Binding assays with NFkB and HP1γ were performed as follows, 10 μl of 200 nM SHL -6 DNA or 500 nM MN -6 in buffer (20 mM Tris pH 7.0 RT, 50 mM NaCl, 10 mM KCl, 1 mM $MgCl_2$, 0.1 mg/ml BSA, and 1 mM DTT) was mixed with 500 nM, 1 μM, or 3 μM NFkB and 17 μM of HP1γ. Samples were incubated at 4 °C for 30 minutes before loading 3 μl on a 5% polyacrylamide gel as described above.

MNase digest of OCT4

192 µl of OCT4/DNA was at a concentration of 500 ng/ul DNA and 50 µM GFP-OCT4 (ε488 = 56,000) after purification by size exclusion chromatography. 1 µl of MNase (NEB) was added for every 10 µg of DNA. 20 µl of MNase reaction buffer (NEB) was supplemented. The reaction proceeded for an hour at room temperature, then centrifuged at 21,000g, and finally loaded on a 5% polyacrylamide gel for analysis.

Biotin-Strep bead pulldown

Figure 2D: Invitrogen Dynabeads M-280 Streptavidin were washed with three times with assay buffer (10 mM Tris pH 7.0, 100 mM NaCl, 1 mM EDTA, 1 mM DTT, 0.1 mg/ml BSA). 400 µl binding mixtures of 1 µM HP1α and 4.12 ng/µl mononucleosomes were used with 7.5 µl of magnetic beads. The mixture was incubated at room temperature for 30 minutes, then washed once with 400 µl of assay buffer. The beads were then resuspended in 1x SDS loading dye and boiled at 95 °C for 5 minutes. Samples were loaded onto a 18% polyacrylamide SDS-PAGE gel and ran in SDS-Tris-Glycine at 25 °C using power cell conditions of 200 V for 45 minutes. Gels were stained by Coomassie blue and destained before imaging.

Table for Figure 2D:

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
	HP1α ctrl	CSD ctrl	HP1α 100 mM NaCl	HP1α 100 mM NaCl	HP1α 250 mM NaCl	HP1α 500 mM NaCl	W174A 100 mM NaCl	CSD 100 mM NaCl	CSD 250 mM NaCl	CSD 500 mM NaCl
MN	0	0	0	22	22	22	22	22	22	22
DNA	0	0	4	0	0	0	0	0	0	0
HP1α WT	1.03	0	1.03	1.03	1.03	1.03	0	0	0	0
HP1α W174A	0	0	0	0	0	0	0.89	0	0	0
HP1α CSD	0	1.47	0	0	0	0	0	1.47	1.47	1.47
buffer	398.97	398.53	394.97	376.97	376.97	376.97	377.11	376.53	376.53	376.53
beads	7.5	7.5	7.5	7.5	7.5	7.5	7.5	7.5	7.5	7.5

Monoucleosome remodeling assays

Figure 4: The mononucleosome remodeling relies on the modification of the 601 sequence to include a PstI restriction enzyme cut site (CTGCAG) 22 base pairs into the nucleosomal DNA. H3 K9me3 mononucleosomes made with the 205 bp DNA (+PstI) were remodeled in a 10 µl volume at 3 ng/µl (22.5 nM). The mixture was formulated by adding 2 µl of 5x EpiCypher remodeling buffer (100 mM Tris HCl, pH 7.5, 250 mM KCl, 15 mM MgCl2, 0.05% (w/v) BSA, 0.05% (v/v) Tween 20). 1 µl of a 20 mM ATP stock, 1 µl (20 units) of PstI enzyme (NEB), 1 µl of CHD1 (2.5 µM stock) or 1 µl of ACF (500 nM stock), 1 µl of a range of HP1 stocks (2.5 µM to 250 µM), and filled rest with H_2O. Final conditions had 22.5 nM MN, 2 mM ATP, and 3 mM $MgCl_2$, with HP1 titration spanning 250 nM, 500 nM, 1 µM, 5 µM, 25 µM. The ratio between ATP and $MgCl_2$, PstI and remodeler is crucial toward remodeling efficacy. Reactions were left for 30 minutes at 30 °C, 30 minutes was selected due to its intermediate kinetics (reactions were complete at 60 minutes). After 30 minutes, reactions were quenched by adding 2x quenching buffer (2% SDS, 100 mM EDTA, 20% glycerol and 1:10 Proteinase K (NEB), 2x loading dye) then left at 50 °C for 30 minutes. For Figure 6B, a time course of 15, 30, and 60 minutes was conducted. A target of 10 ng DNA was loaded in each lane of a 5% polyacrylamide gel ran at 150 V for 35 minutes at room temperature. Gels were stained with a 1:10000 SyBr Gold dilution, and imaged with Cy3 excitation/detection. Gel images were analyzed and quantified by ImageJ, using the integrated band intensity for cutting percentages.

Supplemental Figure 3: Native-PAGE remodels were done same as above without the addition of SDS and Proteinase K, and without the 50 °C incubation. CHD1 was titrated from 5 nM, 25 nM, 100 nM, 250 nM, to 500 nM. After remodeling, nucleosomes were kept cold, running electrophoresis at 120 V for one hour at 4 °C.

Nucleosome array remodeling assays

Figure 5: Nucleosome array remodeling was done similar to mononucleosomes, with some minor adjustments. The concentrations used for H3 K9me3 nucleosome arrays were 25 nM (601 site concentration), 66 nM Brg1, 500 nM CHD1, and 1 µl (10 units) of HpaII (NEB) as the restriction enzyme. The HpaII cut site CCGG is present 10 base pairs into the nucleosomal DNA in all 12 nucleosomes. The final $MgCl_2$ concentration was raised to 4 mM from 3 mM to stimulate remodeling while staying below 5 mM to prevent chromatin condensation. Samples were quenched with 2x quench buffer after the designated time periods, incubated at 50 °C, then ran on a 5% polyacrylamide gel at room temperature with 150 V for 40 minutes. Gels were stained with a 1:10000 SyBr Gold dilution, and imaged with Cy3 excitation/detection.

Supplemental Figure 3: This divided remodel assay split the remodel into two sections. The first 30 minutes of remodeling proceeded with 25 µM HP1 for the pre-sample, and without HP1 or extra $MgCl_2$ for the post- samples. After 15 minutes, 1 µl of HpaII was added to all samples, in addition to 25 µM HP1 or 5 mM $MgCl_2$ to designated post- lanes. The reaction proceeded for another 15 minutes in the presence of HpaII before quenching with 2x quench buffer. These samples were then processed the same as above.

Microscopy

The samples were prepared as done for nucleosome array remodeling. 25 nM (nucleosome site) nucleosome arrays unmodified or H3 K9me3 were prepared in remodel buffer and mixed with 25 μM HP1α or 10 mM $MgCl_2$. Samples at room temperature were then stained with 0.1 μM YOYO-1 dye. Microscopy images of chromatin were taken using an Olympus CKX53 with a GFP filter cube.

Spectral analysis

Spectra were processed using Bruker TopSpin 4.0.5 and analyzed using NMR-FAM SPARKY.[49] Peak and integrated signal intensities were quantified using TopSpin 4.0.5.

Structural analysis

Molecular graphics and analyses performed with UCSF ChimeraX, developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco, with support from National Institutes of Health R01-GM129325 and the Office of Cyber Infrastructure and Computational Biology, National Institute of Allergy and Infectious Diseases.[50]

Supplementary figures

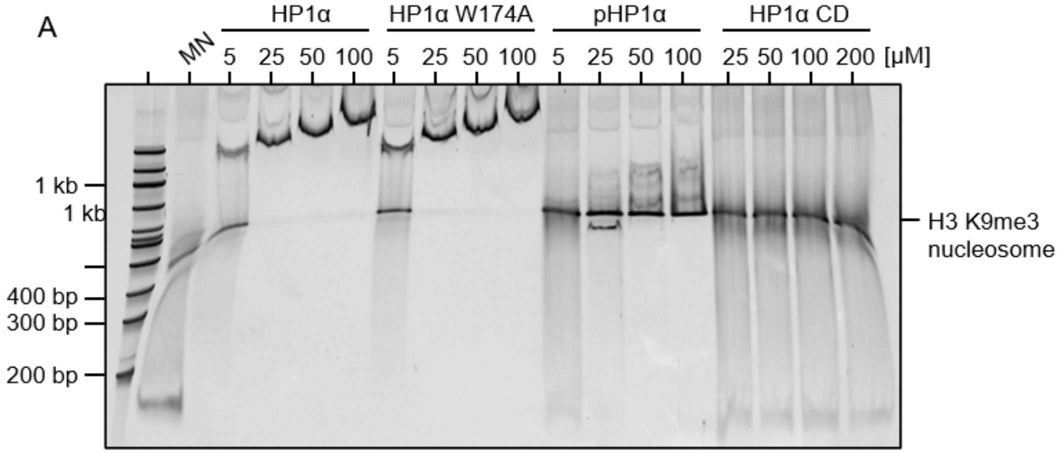


Figure 4.S1. EMSA binding of HP1α and MNs. EMSA of H3 K9me3 MN with HP1α, HP1α W174A, and phosphorylated HP1α ranging from 5 to 100 μM, and HP1α CD ranging from 25 to 200 μM.

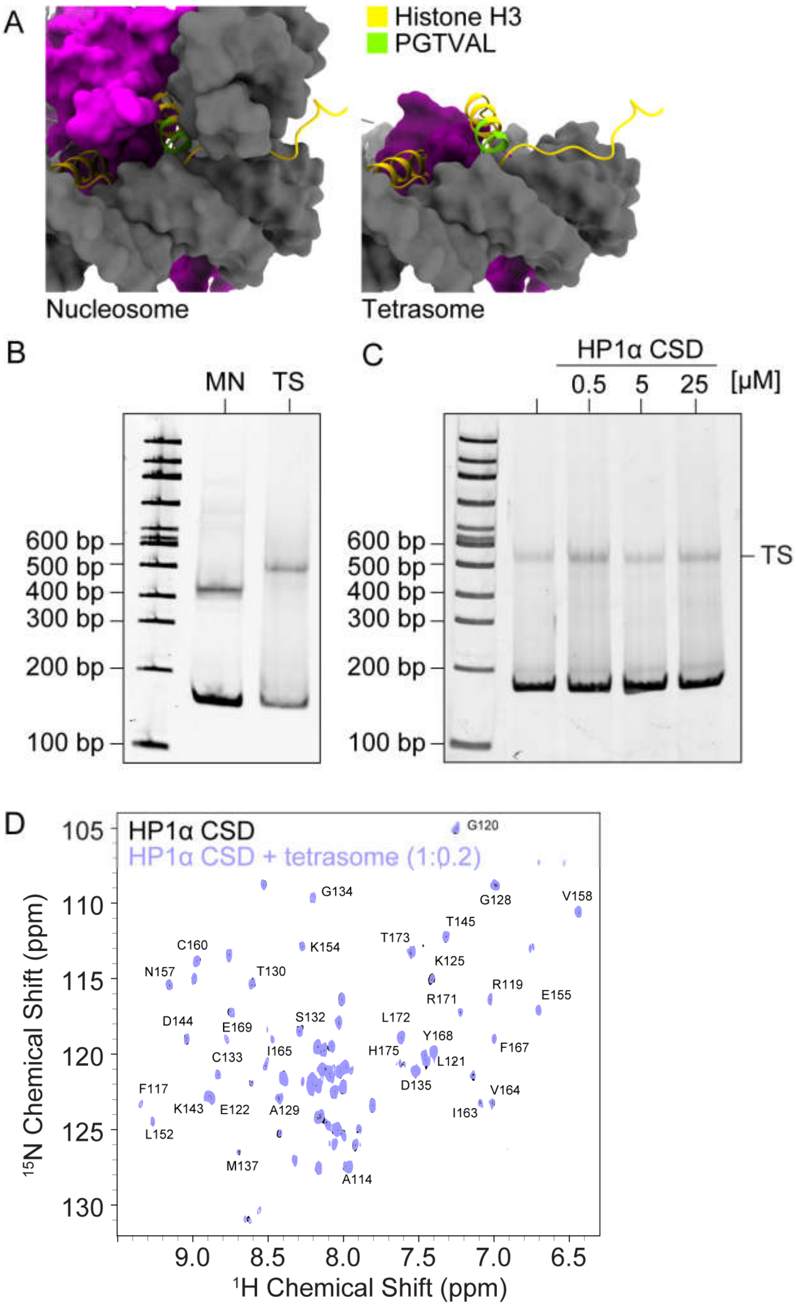


Figure 4.S2. Chromoshadow domain interactions with tetrasomes. A) Comparison of H3 exposure with doubly wrapped DNA in the nucleosome versus singly wrapped DNA in the tetrasome. The tetrasome was artificially created based off of the nucleosome structure PDB:1K5X, by removing the top DNA strand and H2A/H2B histones. B) Native-PAGE gel comparing the tetrasome (TS) to MN. The extended DNA in the TS causes it to run at a higher molecular weight. C) EMSA of TS with varying concentrations of HP1α CSD. D) ^{1}H-^{15}N HSQC of HP1α CSD (black) and HP1α CSD 1:0.2 with TS (purple).

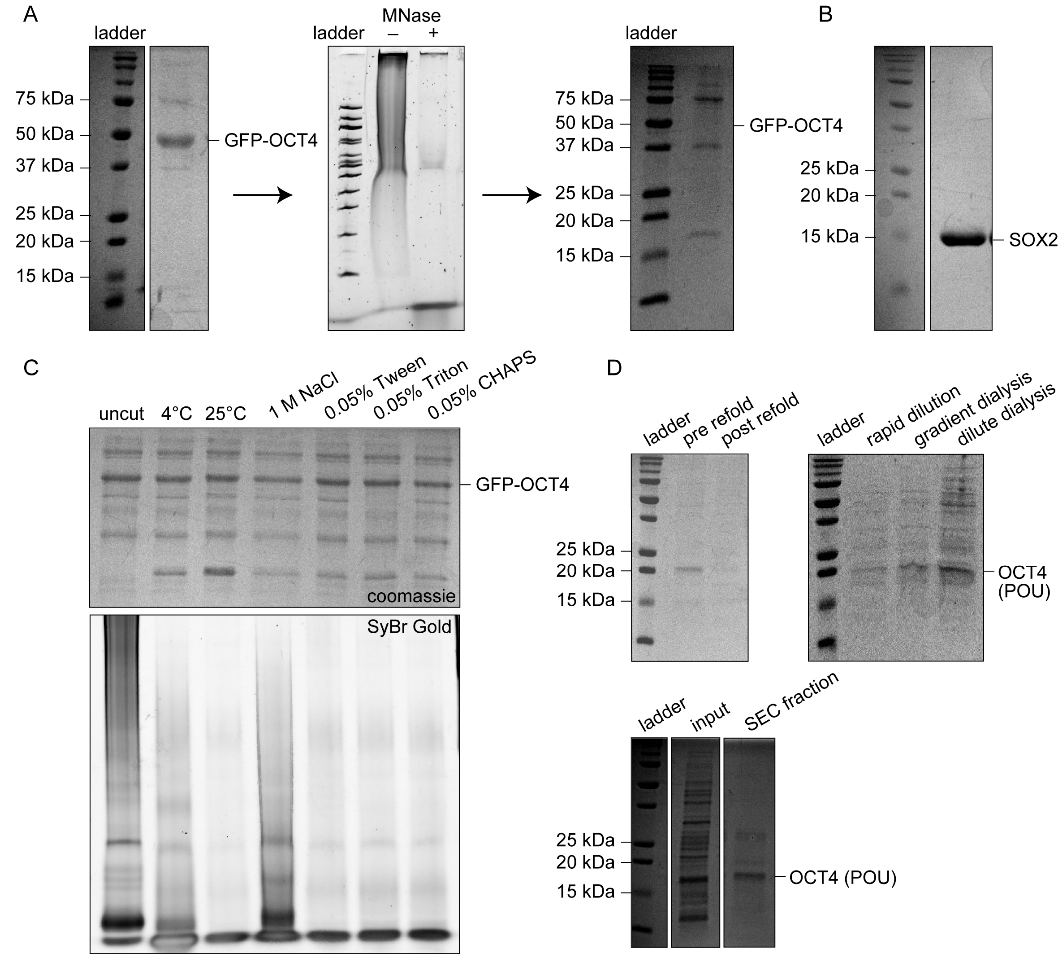


Figure 4.S3. Protein purification of eGFP-OCT4, OCT4 POU and SOX2 HMG. A) SDS-PAGE of eGFP-OCT4 purified from *E.coli.* A polyacrylamide gel stained with SyBr Gold shows MNase treatment of the purified eGFP-OCT4. The following SDS-PAGE of the MNase digested eGFP-OCT4 shows complete loss of protein. B) SDS-PAGE of purified SOX2 HMG domain. C) Above: SDS-PAGE of samples with different additives directly after MNase treatment. eGFP-OCT4 quickly depleted over time after these digestions. Below: SyBr Gold stained polyacrylamide gel of MNase, showing the digestion of large DNA into small pieces at bottom of gel. D) Left: Standard single step refolding protocol causes complete loss of the OCT4 POU domain. Right: Different refolding strategies applied to an equivalent amount of OCT4 POU and concentrated to the same volume. Below: Size-exclusion chromatography fraction of OCT4 POU directly after refolding.

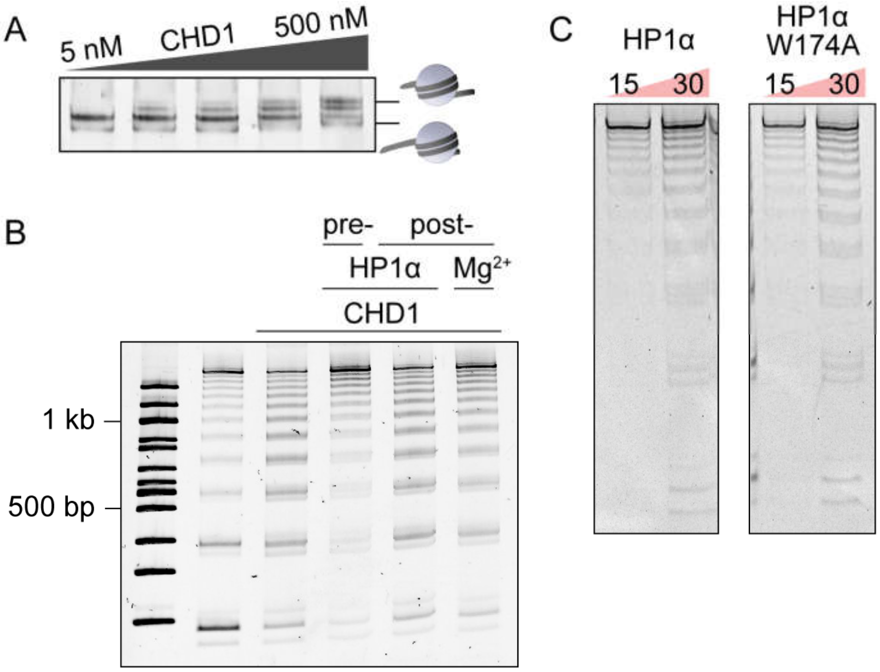


Figure 4.S4. Remodeling assay controls. A) Native gel of the remodeling assay where the DNA overhang modulates the band position on the gel. This gel complements the restriction enzyme-based remodeingl assay by showing the effect of remodeling directly on the nucleosome band. B) Gel of deproteinized remodeling products after nucleosome array remodeling by CHD1. HP1α was either added during the remodeling period (pre-) or after (post-). C) Gel of deproteinized remodeling products after nucleosome array CHD1 remodeling assay using wildtype HP1α or HP1α W174A.

References

[1] aN. Saksouk, E. Simboeck, J. Déjardin, *Epigenetics & Chromatin* **2015**, *8*; bA. Janssen, S. U. Colmenares, G. H. Karpen, **2018**, 24.

[2] Nathaniel A. Hathaway, O. Bell, C. Hodges, Erik L. Miller, Dana S. Neel, Gerald R. Crabtree, *Cell* **2012**, *149*, 1447-1460.

[3] M. Falk, Y. Feodorova, N. Naumova, M. Imakaev, B. R. Lajoie, H. Leonhardt, B. Joffe, J. Dekker, G. Fudenberg, I. Solovei, L. Mirny, **2018**.

[4] L. E. Hall, S. E. Mitchell, R. J. O'Neill, *Chromosome Research* **2012**, *20*, 535-546.

[5] aR. D. Hawkins, G. C. Hon, L. K. Lee, Q. Ngo, R. Lister, M. Pelizzola, L. E. Edsall, S. Kuan, Y. Luu, S. Klugman, J. Antosiewicz-Bourget, Z. Ye, C. Espinoza, S. Agarwahl, L. Shen, V. Ruotti, W. Wang, R. Stewart, J. A. Thomson, J. R. Ecker, B. Ren, *Cell Stem Cell* **2010**, *6*, 479-491; bS. A. Grigoryev, *Journal of Cell Science* **2004**; cN. Z. Zaidan, K. J. Walker, J. E. Brown, L. V. Schaffer, M. Scalf, M. R. Shortreed, G. Iyer, L. M. Smith, R. Sridharan, *Stem Cell Reports* **2018**, *10*, 627-641.

[6] F. G. Westhorpe, A. F. Straight, *Cold Spring Harbor Perspectives in Biology* **2015**, *7*, a015818.

[7] C. R. Vakoc, S. A. Mandat, B. A. Olenchock, G. A. Blobel, *Molecular Cell* **2005**, *19*, 381-391.

[8] K. Muller-Ott, F. Erdel, A. Matveeva, J. P. Mallm, A. Rademacher, M. Hahn, C. Bauer, Q. Zhang, S. Kaltofen, G. Schotta, T. Hofer, K. Rippe, *Molecular Systems Biology* **2014**, *10*, 746-746.

[9] aN. Saksouk, S. Hajdari, Y. Perez, M. Pratlong, C. Barrachina, C. Graber, D. Grégoire, A. Zavoriti, A. Sarrazin, N. Pirot, J.-Y. Noël, L. Khellaf, E. Fabbrizio, E. Julien, F. M. Cammas, *Oncogene* **2020**; bL. Bosch-Presegué, H. Raurell-Vila, J. K. Thackray, J. González, C. Casal, N. Kane-Goldsmith, M. Vizoso, J. P. Brown, A. Gómez, J. Ausió, T. Zimmermann, M. Esteller, G. Schotta, P. B. Singh, L. Serrano, A. Vaquero, *Cell Reports* **2017**, *21*, 2048-2057.

[10] aS. A. Jacobs, *Science* **2002**, *295*, 2080-2083; bP. R. Nielsen, D. Nietlispach, H. R. Mott, J. Callaghan, A. Bannister, T. Kouzarides, A. G. Murzin, N. V. Murzina, E. D. Laue, *Nature* **2002**, *416*, 103-107.

[11] aR. J. Roach, M. Garavís, C. González, G. B. Jameson, V. V. Filichev, T. K. Hale, *Nucleic Acids Research* **2020**, *48*, 682-693; bK. Hiragami-Hamada, S. Soeroes, M. Nikolov, B. Wilkins, S. Kreuz, C. Chen, I. A. De La Rosa-Velázquez, H. M. Zenn, N. Kost, W. Pohl, A. Chernev, D. Schwarzer, T. Jenuwein, M. Lorincz, B. Zimmermann, P. J. Walla, H. Neumann, T. Baubec, H. Urlaub, W. Fischle, *Nature Communications* **2016**, *7*, 11310.

[12] aD. P. Ryan, D. J. Tremethick, *Nucleic Acids Research* **2018**; bM. Lavigne, R. Eskeland, S. Azebi, V. Saint-André, S. M. Jang, E. Batsché, H.-Y. Fan, R. E. Kingston, A. Imhof, C. Muchardt, *PLoS Genetics* **2009**, *5*, e1000769.

[13] A. N. Richart, C. I. W. Brunner, K. Stott, N. V. Murzina, J. O. Thomas, *Journal of Biological Chemistry* **2012**, *287*, 18730-18737.

[14] Y. Liu, S. Qin, M. Lei, W. Tempel, Y. Zhang, P. Loppnau, Y. Li, J. Min, *Journal of Biological Chemistry* **2017**, *292*, 5655-5664.

[15] aB. Ren, E. S. Chen, *Current Genetics* **2019**, *65*, 829-836; bM. A. Dawson, A. J. Bannister, B. Göttgens, S. D. Foster, T. Bartke, A. R. Green, T. Kouzarides, *Nature* **2009**, *461*, 819-822; cS. M. Jang, S. Azebi, G. Soubigou, C. Muchardt, *EMBO reports* **2014**, *15*, 686-694.

[16] R.-S. Nozawa, K. Nagao, H.-T. Masuda, O. Iwasaki, T. Hirota, N. Nozaki, H. Kimura, C. Obuse, *Nature Cell Biology* **2010**, *12*, 719-727.

[17] aV. Ostapcuk, F. Mohn, S. H. Carl, A. Basters, D. Hess, V. Iesmantavicius, L. Lampersberger, M. Flemr, A. Pandey, N. H. Thomä, J. Betschinger, M. Bühler, *Nature* **2018**, *557*, 739-743; bS. Machida, Y. Takizawa, M. Ishimaru, Y. Sugita, S. Sekine, J.-i. Nakayama, M. Wolf, H. Kurumizaka, *Molecular Cell* **2018**, *69*, 385-397.e388; cR. Eskeland, A. Eberharter, A. Imhof, *Molecular and Cellular Biology* **2007**, *27*, 453-465; dS. H. Kwon, L. Florens, S. K. Swanson, M. P. Washburn, S. M. Abmayr, J. L. Workman, *Genes & Development* **2010**, *24*, 2133-2145; eA. V. Emelyanov, A. Y. Konev, E. Vershilova, D. V. Fyodorov, *Journal of Biological Chemistry* **2010**, *285*, 15027-15037; fSamuel P. Rowbotham, L. Barki, A. Neves-Costa, F. Santos, W. Dean, N. Hawkes, P. Choudhary, W. R. Will, J. Webster, D. Oxley, Catherine M. Green, P. Varga-Weisz, Jacqueline E. Mermoud, *Molecular Cell* **2011**, *42*, 285-296; gK. Leopold, A. Stirpe, T. Schalch, *Genes & Development* **2019**, *33*, 565-577.

[18] A. L. Nielsen, *The EMBO Journal* **2002**, *21*, 5797-5806.

[19] aG. D. Bowman, M. G. Poirier, *Chemical Reviews* **2015**, *115*, 2274-2295; bY. Tan, Y. Xue, C. Song, M. Grunstein, *Proceedings of the National Academy of Sciences* **2013**, *110*, 11493-11498.

[20] C. Chronis, P. Fiziev, B. Papp, S. Butz, G. Bonora, S. Sabri, J. Ernst, K. Plath, *Cell* **2017**, *168*, 442-459.e420.

[21] aA. Mattout, A. Biran, E. Meshorer, *Journal of Molecular Cell Biology* **2011**, *3*, 341-350; bY. H. Loh, W. Zhang, X. Chen, J. George, H. H. Ng, *Genes & Development* **2007**, *21*, 2545-2557; cJ. Chen, H. Liu, J. Liu, J. Qi, B. Wei, J. Yang, H. Liang, Y. Chen, J. Chen, Y. Wu, L. Guo, J. Zhu, X. Zhao, T. Peng, Y. Zhang, S. Chen, X. Li, D. Li, T. Wang, D. Pei, *Nature Genetics* **2013**, *45*, 34-42.

[22] A. Soufi, G. Donahue, Kenneth S. Zaret, *Cell* **2012**, *151*, 994-1004.

[23] aN. Singhal, D. Esch, M. Stehling, H. R. Schöler, *BioResearch Open Access* **2014**, *3*, 1-8; bH. W. King, R. J. Klose, *eLife* **2017**, *6*; cJ. S. You, T. K. Kelly, D. D. De Carvalho, P. C. Taberlay, G. Liang, P. A. Jones, *Proceedings of the National Academy of Sciences* **2011**, *108*, 14497-14502.

[24] S. Bilokapic, M. Halic, *Nature Communications* **2019**, *10*.

[25] aS. Bilokapic, M. Strauss, M. Halic, *Nature Structural & Molecular Biology* **2018**, *25*, 101-108; bS. Bilokapic, M. Strauss, M. Halic, *Nature Communications* **2018**, *9*.

[26] aS. Sanulli, M. J. Trnka, V. Dharmarajan, R. W. Tibble, B. D. Pascal, A. L. Burlingame, P. R. Griffin, J. D. Gross, G. J. Narlikar, *Nature* **2019**; bK. K. Sinha, J. D. Gross, G. J. Narlikar, *Science* **2017**, *355*, eaaa3761.

[27] A. K. Michael, R. S. Grand, L. Isbel, S. Cavadini, Z. Kozicka, G. Kempf, R. D. Bunker, A. D. Schenk, A. Graff-Meyer, G. R. Pathare, J. Weiss, S. Matsumoto, L. Burger, D. Schübeler, N. H. Thomä, *Science* **2020**, *368*, 1460-1465.

[28] T. Stormberg, S. Filliaux, H. E. R. Baughman, E. A. Komives, Y. L. Lyubchenko, *Biochim Biophys Acta Gen Subj* **2021**, *1865*, 129934.

[29] G. J. Narlikar, M. L. Phelan, R. E. Kingston, *Mol Cell* **2001**, *8*, 1219-1230.

[30] R. Sundaramoorthy, A. L. Hughes, H. El-Mkami, D. G. Norman, H. Ferreira, T. Owen-Hughes, *eLife* **2018**, *7*.

[31] J. G. Yang, T. S. Madrid, E. Sevastopoulos, G. J. Narlikar, *Nature Structural & Molecular Biology* **2006**, *13*, 1078-1083.

[32] L. Wang, Y. Gao, X. Zheng, C. Liu, S. Dong, R. Li, G. Zhang, Y. Wei, H. Qu, Y. Li, C. D. Allis, G. Li, H. Li, P. Li, *Molecular Cell* **2019**, *76*, 646-659.e646.

[33] aA. R. Strom, A. V. Emelyanov, M. Mir, D. V. Fyodorov, X. Darzacq, G. H. Karpen, *Nature* **2017**, *547*, 241-245; bA. G. Larson, D. Elnatan, M. M. Keenen, M. J. Trnka, J. B. Johnston, A. L. Burlingame, D. A. Agard, S. Redding, G. J. Narlikar, *Nature* **2017**, *547*, 236-240; cF. Erdel, K. Rippe, *Biophysical Journal* **2018**, *114*, 2262-2270; dF. Erdel, A. Rademacher, R. Vlijm, J. Tünnermann, L. Frank, R. Weinmann, E. Schweigert, K. Yserentant, J. Hummert, C. Bauer, S. Schumacher, A. Al Alwash, C. Normand, D.-P. Herten, J. Engelhardt, K. Rippe, *Molecular Cell* **2020**.

[34] C. Her, T. M. Phan, N. Jovic, U. Kapoor, B. E. Ackermann, A. Rizuan, Y. Kim, J. Mittal, G. T. Debelouchina, *bioRxiv* **2022**, 2022.2006.2020.496886.

[35] A. R. Strom, R. J. Biggs, E. J. Banigan, X. Wang, K. Chiu, C. Herman, J. Collado, F. Yue, J. C. Ritland Politz, L. J. Tait, D. Scalzo, A. Telling, M. Groudine, C. P. Brangwynne, J. F. Marko, A. D. Stephens, *eLife* **2021**, *10*, e63972.

[36] S. He, Z. Wu, Y. Tian, Z. Yu, J. Yu, X. Wang, J. Li, B. Liu, Y. Xu, *Science* **2020**, *367*, 875-881.

[37] M. F. Vassallo, N. Tanese, *Proceedings of the National Academy of Sciences* **2002**, *99*, 5919-5924.

[38] J. P. Turowec, J. S. Duncan, A. C. French, L. Gyenis, N. A. St. Denis, G. Vilk, D. W. Litchfield, in *Methods in Enzymology, Vol. 484*, Elsevier, **2010**, pp. 471-493.

[39] J. Yu, K. Vodyanik Ma Fau - Smuga-Otto, J. Smuga-Otto K Fau - Antosiewicz-Bourget, J. L. Antosiewicz-Bourget J Fau - Frane, S. Frane Jl Fau - Tian, J. Tian S Fau - Nie, G. A. Nie J Fau - Jonsdottir, V. Jonsdottir Ga Fau - Ruotti, R. Ruotti V Fau - Stewart, I. I. Stewart R Fau - Slukvin, J. A. Slukvin Ii Fau - Thomson, J. A. Thomson.

[40] B. E. Ackermann, G. T. Debelouchina, *Angew. Chem. Int. Ed.* **2019**, *58*, 6300-6305.

[41] aS. O. Dodonova, F. Zhu, C. Dienemann, J. Taipale, P. Cramer, *Nature* **2020**; bX. Pan, X. Cang, S. Dan, J. Li, J. Cheng, B. Kang, X. Duan, B. Shen, Y.-J. Wang, *Journal of Biological Chemistry* **2016**, *291*, 18353-18369; cS. Li, E. B. Zheng, L. Zhao, S. Liu, *Cell Reports* **2019**, *28*, 2689-2703.e2684; dK. Echigoya, M. Koyama, L. Negishi, Y. Takizawa, Y. Mizukami, H. Shimabayashi, A. Kuroda, H. Kurumizaka, *Scientific Reports* **2020**, *10*.

[42] D. Esch, J. Vahokoski, M. R. Groves, V. Pogenberg, V. Cojocaru, H. vom Bruch, D. Han, H. C. A. Drexler, M. J. Araúzo-Bravo, C. K. L. Ng, R. Jauch, M. Wilmanns, H. R. Schöler, *Nature Cell Biology* **2013**, *15*, 295-301.

[43] M. D. Simon, F. Chu, L. R. Racki, C. C. de la Cruz, A. L. Burlingame, B. Panning, G. J. Narlikar, K. M. Shokat, *Cell* **2007**, *128*, 1003-1012.

[44] G. T. Debelouchina, K. Gerecht, T. W. Muir, *Nature Chemical Biology* **2017**, *13*, 105-110.

[45] aA. Flaus, T. J. Richmond, *Journal of Molecular Biology* **1998**, *275*, 427-441; bP. N. Dyer, R. S. Edayathumangalam, C. L. White, Y. Bao, S. Chakravarthy, U. M. Muthurajan, K. Luger, in *Methods in Enzymology, Vol. 375*, Elsevier, **2003**, pp. 23-44.

[46] B. Fierz, C. Chatterjee, R. K. McGinty, M. Bar-Dagan, D. P. Raleigh, T. W. Muir, *Nature Chemical Biology* **2011**, *7*, 113-119.

[47] P. T. Lowary, J. Widom, *Journal of Molecular Biology* **1998**, *276*, 19-42.

[48] E. A. Morrison, L. Baweja, M. G. Poirier, J. Wereszczynski, C. A. Musselman, *Nucleic Acids Research* **2021**, *49*, 4750-4767.

[49] W. Lee, M. Tonelli, J. L. Markley, *Bioinformatics* **2015**, *31*, 1325-1327.

[50] E. F. Pettersen, T. D. Goddard, C. C. Huang, E. C. Meng, G. S. Couch, T. I. Croll, J. H. Morris, T. E. Ferrin, *Protein Sci* **2021**, *30*, 70-82.

Acknowledgements

Chapter 4 is work to be contributed to a publication by Nesreen Elathram, Evan Clark, and Galia T. Debelouchina.

Chapter V

Targetable Tetrazine-Based Dynamic Nuclear Polarization Agents for Biological Systems

Targetable Tetrazine-Based Dynamic Nuclear Polarization Agents for Biological Systems**

Byung Joon Lim[+], Bryce E. Ackermann[+], and Galia T. Debelouchina*[a]

Dynamic nuclear polarization (DNP) has shown great promise as a tool to enhance the nuclear magnetic resonance signals of proteins in the cellular environment. As sensitivity increases, the ability to select and efficiently polarize a specific macromolecule over the cellular background has become desirable. Herein, we address this need and present a tetrazine-based DNP agent that can be targeted selectively to proteins containing the unnatural amino acid (UAA) norbornene-lysine. This UAA can be introduced efficiently into the cellular milieu by genetic means. Our approach is bio-orthogonal and easily adaptable to any protein of interest. We illustrate the scope of our methodology and investigate the DNP transfer mechanisms in several biological systems. Our results shed light on the complex polarization-transfer pathways in targeted DNP and ultimately pave the way to selective DNP-enhanced NMR spectroscopy in both bacterial and mammalian cells.

Dynamic nuclear polarization (DNP) has had a significant impact on solid-state nuclear magnetic resonance investigations of complex biological systems.[1,2] With now routine signal enhancements in the 100-fold range, DNP has aided the in vitro structural studies of amyloid polymers,[3,4] membrane proteins,[5,6] and biological material such as collagen, bones and tissues.[7,8] Improvements in DNP instrumentation, sample preparation, and polarization agent design have also allowed NMR spectroscopists to turn their attention to the cellular environment, where the structural analysis of endogenous concentrations of biological macromolecules has come within reach.[3,9–11] In a typical DNP experiment, the sample is doped with millimolar amounts of polarization agents, small molecules that contain two stable unpaired-electron spins. The sample is cryoprotected, frozen to ≈100 K, and subjected to continuous microwave irradiation during magic angle spinning (MAS). The generated electron spin polarization is transferred to the nuclear spins, resulting in significant enhancement of the NMR signals. Concurrently, multidimensional NMR experiments are performed, and structural constraints with much improved sensitivity are obtained. Although this strategy has been applied to most in vitro and cellular DNP experiments to date, it suffers from one major limitation: the polarization transfer is nonselective and results in enhancements for all molecules in the sample, including the cellular background.[10,11] To overcome this, several biochemical strategies can be employed including overexpression of the protein of interest[12,13] or the introduction of recombinant isotopically labeled protein into cells by electroporation.[11] These approaches, however, require extensive perturbation of the cellular milieu and are not applicable to the structural studies of endogenous concentrations of cellular components. Alternatively, the design of targeted DNP agents has been explored, with selectivity achieved either through cysteine chemistry or protein–ligand interactions.[14–20] The applications of these strategies, however, are limited by the reducing cellular environment or the requirement for a known synthetically accessible ligand of the target protein. We therefore sought to develop a general targeting strategy that could be applied to any protein of interest in a cysteine-independent manner, and that was compatible with DNP-enhanced NMR studies of proteins in the cellular interior.

Our targeting approach is based on the bio-orthogonal chemical reactions between tetrazines and strained cycloalkenes.[21,22] These reactions display superb efficiency and selectivity in the cellular milieu and are routinely used to label cellular proteins with optical probes.[23–26] Here, we harness their beneficial properties to attach DNP agents to a variety of proteins (Figure 1 A). To this end, we started by coupling a tetrazine moiety to TOTAPOL, a stable, nitroxide-based DNP agent[27] (see scheme and Figure S1 in the Supporting Information). The TOTAPOL–tetrazine biradical (TTz) displays characteristic nitroxide hyperfine splitting in its EPR spectra and can be efficiently conjugated to strained cycloalkenes such as norbornene. Next, we synthesized the unnatural amino acid (UAA) norbornene-lysine[25] and introduced it at position 6 in the sequence of the model protein ubiquitin through amber suppression (Figures 1 and S2).

Amber suppression is a molecular biology technique that enables the introduction of UAAs into living cells by genetic means through reassignment of the amber (TAG) stop codon.[28] The UAA is incorporated at the desired TAG position through an orthogonal tRNA and aminoacyl-tRNA synthetase (aaRS) pair introduced into the cell through a separate plasmid construct. Although amber suppression is often used for optical and EPR experiments,[29–31] our application required concurrent ^{13}C,^{15}N labeling of the targeted protein and thus presented additional challenges for protein expression. After testing several plasmid vectors and constructs, we chose the pULTRA vector, in which expression of the tRNA/aaRS pair is under the

---

[a] B. J. Lim,+ B. E. Ackermann,+ Dr. G. T. Debelouchina
Department of Chemistry and Biochemistry, University of California, San Diego
9500 Gilman Dr., La Jolla, CA 92093 (USA)
E-mail: gdebelouchina@ucsd.edu

[+] These authors contributed equally to this work.

[**] A previous version of this manuscript has been deposited on a preprint server (https://doi.org/10.1101/740530).

Supporting information and the ORCID identification numbers for the authors of this article can be found under https://doi.org/10.1002/cbic.201900609.

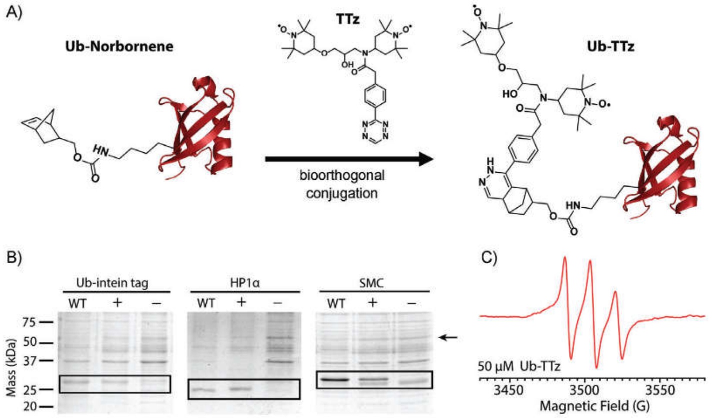


Figure 1. A) A general polarization-agent-targeting strategy based on the bio-orthogonal reaction between tetrazine and norbornene. B) Incorporation of norbornene-lysine into three different proteins by amber suppression. The bands compare the expression of unmodified ^{13}C,^{15}N-labeled proteins (WT), the expression of ^{13}C,^{15}N-labeled proteins containing the UAA (+) and a control sample where the UAA was not added to the media and no protein expression is expected (−). C) 9 GHz EPR spectrum of 50 µm ubiquitin–TTz.

control of a *lacI* promoter,[32] and combined it with a tRNA/aaRS sequence optimized for the incorporation of norbornene-lysine.[24] Using this strategy, we observed high yields of isotopically labeled and modified proteins, not only for ubiquitin but also for several other systems of various sizes and complexity, including heterochromatin protein 1 (HP1α) and the structural maintenance of chromosomes (SMC) protein (Figure 1 B).

Armed with these reagents, we performed the biorthogonal reaction to produce protein–TTz conjugates. The reaction between tetrazine and norbornene was robust, with high efficiency under both denaturing and native conditions. After successful protein–TTz conjugation, excess TTz was removed by dialysis, and the expected 1:1 ratio of protein to TTz was confirmed by LC-MS and EPR (Figures 1 C, S3, and S4).

We next determined the efficiency of TTz as a polarization source in comparison to standard DNP sample preparation, in which excess polarization agent is evenly distributed throughout a glassy matrix. As a direct measure of the enhancement, we used the signal in the presence of microwaves divided by the signal in the absence of irradiation. Under these conditions, we measured an enhancement of 24 for 1 mm ubiquitin–TTz; this is 68 % of the enhancement provided by dispersed TOTAPOL despite the 15-fold difference in radical concentration (Figure 2 A). The enhancement observed in our conjugated system is comparable to or higher than the enhancements reported for other targeted DNP approaches,[14] and it was sufficient to record a 2D ^{13}C,^{13}C spectrum of just 100 µg of ubiquitin (1 mm) in less than a day (Figure 2 B). Comparison to a ubiquitin spectrum obtained with 15 mm dispersed TOTAPOL did not reveal any significant differences in the linewidths of the two samples (Figure S5). Overall, our linewidths (1.5–2 ppm for resolved peaks) appear to be similar to the linewidths obtained in other frozen protein solutions, for which significant surface exposure to the solvent can lead to inhomogeneous line broadening.[11,33–35] Many of the correlations in our spectrum could be assigned based on the recently published low-temperature assignments of ubiquitin samples enhanced with dispersed AMUPol.[11]

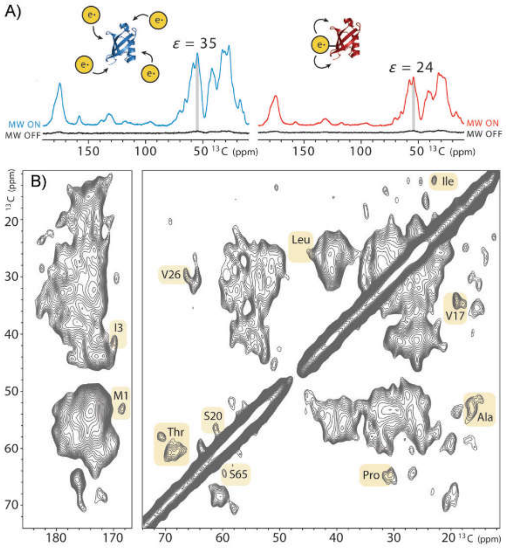


Figure 2. A) Comparison of the DNP enhancements measured in a sample containing 1 mm wild-type ubiquitin with 15 mm dispersed TOTAPOL (left) and 1 mm ubiquitin–TTz (right). Spectra were acquired at 600 MHz ^{1}H Larmor frequency, 12 kHz MAS, 100 K. B) DNP-enhanced 2D ^{13}C,^{13}C CORD correlation[36] spectrum of 1 mm ubiquitin–TTz acquired at 24 kHz.

DNP enhancements depend on many experimental parameters including polarization agent concentration, interscan delay, and MAS frequency.[1,27,37,38] As these parameters can provide valuable information regarding the DNP transfer pathways, we investigated them in more detail. First, we compared conjugated samples at 1 mM protein (i.e., 1 mM TTz) and dilute samples at $\approx$50 μM protein and TTz (Figure 3). Similarly to previous reports,[19] the measured enhancements at lower concentration were smaller, in the 6–10 range for different proteins. It should be noted, however, that dispersed TOTAPOL or TTz samples with similar concentration of biradicals yielded no enhancement. Therefore, conjugation significantly improves DNP transfers by bringing the polarization source and target into close proximity with each other. Under these conditions, we did not observe evidence for significant depolarization or bleaching due to the attached radical (Figure S6).[39,40] The amount of protein in these samples is in the 5–10 μg range and is close to the endogenous amounts of protein present in 10^{6} mammalian cells that can easily fit into an NMR rotor.[11,41]

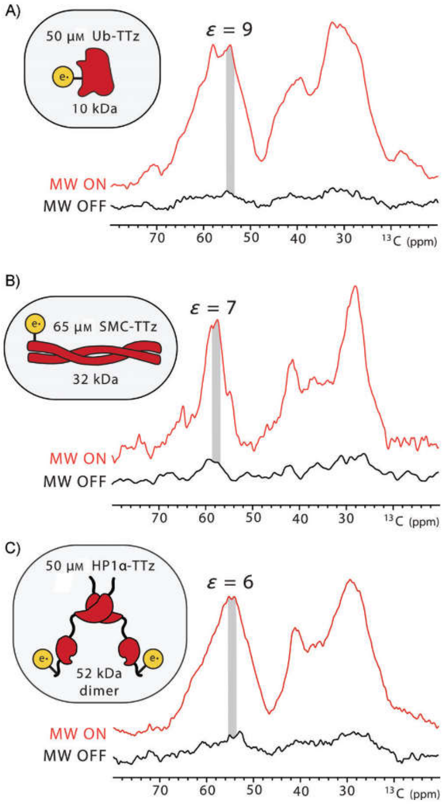


Figure 3. DNP enhancements for low concentrations of A) Ub–TTz; B) SMC–TTz; and C) HP1α–TTz.

Intrigued by the concentration dependence of the enhancements in our targeted systems, we also measured polarization buildup curves that report on the distance dynamics of polarization transfers (Figures 4 and S7).[42,43] We compared the buildup behavior of a sample prepared with 15 mM dispersed TOTAPOL and samples containing 1 mM or 50 μM Ub–TTz. Whereas the dispersed sample displayed a monoexponential buildup curve with a 5.8 s time constant, the behavior of the Ub–TTz samples appeared more complex. In particular, the buildup curve for the 50 μM Ub–TTz sample is consistent with two events, one with a short buildup time, and another event characterized by a much longer time constant. Considering the low concentration of Ub–TTz in this sample, we interpret the first buildup event as the fast polarization transfer from TTz to the directly conjugated protein ($\approx$3 nm in size), whereas the second step reflects long-distance transfer from one conjugate to another ($\approx$50 nm distance). Furthermore, our observations also suggest that at $\approx$1 mM Ub–TTz, where the distance between biradicals is $\approx$15 nm, the buildup curve is at the turning point between mono- and bi-exponential behavior. In this case, each protein is polarized by its own source as well as TTz moieties on neighboring proteins, thus resulting in larger enhancements and buildup times of the order of 8–12 s. As noted before,[19] the majority of targeted DNP studies published so far have been performed at concentrations of 1 mM or above, and, thus, most likely, the reported enhancements reflect a significant intermolecular polarization transfer component.

Finally, we explored the effects of MAS frequency (Figures S8 and S9). Previous studies have reported a complicated relationship between MAS frequency and DNP enhancements due to a time-dependent modulation of the electron and nuclear di-

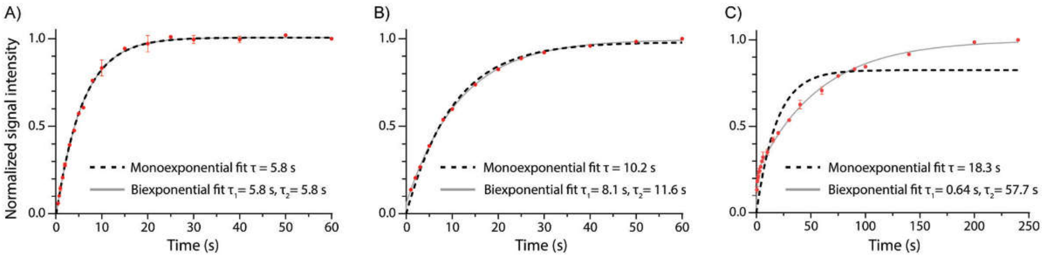


Figure 4. Polarization buildup curves for A) 1 mM wild-type ubiquitin with 15 mM dispersed TOTAPOL; B) 1 mM Ub–TTz; and C) 50 μM Ub–TTz.

polar couplings.[37,44–46] In our case, increasing the MAS frequency from 12 to 24 kHz reduced the enhancements in both dispersed and conjugated samples, although the reduction was much less pronounced for Ub–TTz at low concentration. For example, the signal intensities for the Cα region decreased by 44 % for the dispersed sample, 49 % for the 1 mM Ub–TTz sample, and 15 % for the 50 μM Ub–TTz sample. These results suggest that tethering the biradical to the protein of interest suppresses the unfavorable MAS relationship, presumably due to a less pronounced dependence on long-distance polarization transfers that might be affected by MAS, as predicted by ab initio simulations of MAS DNP enhancements.[47]

To further demonstrate the scope of our targeted bio-orthogonal approach, we extended our studies to a more complex setting. In particular, we targeted isotopically labeled bacterial lysates containing overexpressed ubiquitin with the norbornene-lysine UAA. As TOTAPOL is susceptible to reduction,[48] we used dialysis to remove any endogenous thiols (Figure 5). After incubation with TTz, we again dialyzed the lysate to remove unreacted TTz, and confirmed complete TTz-to-protein conjugation by LC-MS (Figure S10). We performed biochemical assays to estimate the amount of ubiquitin in this sample, which was ≈5 % of the total protein and was present at ≈40 μM in the DNP rotor (Figure S10). Using the background-subtraction approach introduced by Viennet et al.,[18] we estimate that the selective enhancement for conjugated ubiquitin is five times greater than the enhancement of the lysate background. Our overall ubiquitin enhancement in this setting is ≈10, which is on a par with the enhancements we measured for similar concentrations of purified protein–TTz conjugates. The ratio of specific to background enhancement in lysates can be tuned by the interscan delay, where shorter delays favor the targeted protein due to the close proximity of the paramagnetic species (Figure S11). The specificity can be improved even further by coupling targeted DNP with selective isotopic-labeling strategies and direct protein expression in which only the protein of interest is isotopically labeled and enhanced.[49]

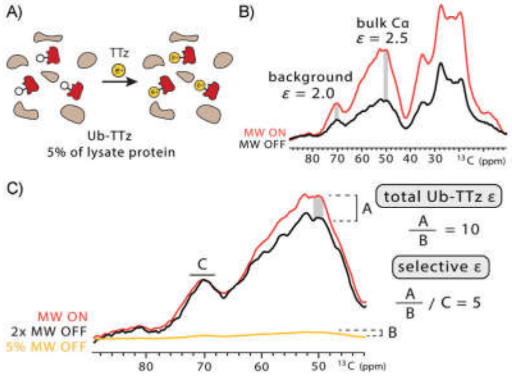


Figure 5. A) Targeted DNP strategy for bacterial lysates. B) Overall signal enhancement of the ^{13}C-labeled lysate. C) Strategy to determine the selective enhancement of Ub–TTz in the lysate. The microwave (MW) off signal was scaled to the microwave on signal so that the glycerol solvent peaks matched in intensity. The leftover enhancement was denoted as *A*. Then, the microwave off signal was scaled down to reflect the 5 % abundance of Ub–TTz in the sample, and the intensity of this rescaled spectrum was denoted as *B*. The ratio of *A* to *B* produced the total enhancement on ubiquitin. The selective enhancement is equal to the total ubiquitin enhancement divided by the background enhancement.

In summary, our tetrazine-based agents represent a general and efficient strategy to target DNP to a variety of protein systems both in vitro and in complex biological settings. As tools for amber suppression exist for different organisms, our approach is applicable to virtually any protein in bacterial or mammalian cells where tetrazine-based reagents have exhibited good permeability and superb reactivity and selectivity.[23,26,50] In this study, we chose to couple tetrazine to TOTAPOL due to the ease of synthesis of the biradical; however, our strategy is compatible with other popular DNP agents such as AMUPol[51] or future biradicals that might be more stable in the cellular milieu. Our results also demonstrate that much is still to be learned about the polarization transfer mechanisms in targeted DNP. For example, we show that at low protein concentrations, the intramolecular short-range polarization transfer can be decoupled from the intermolecular long-range transfer mechanisms. We have also explored the relationship between enhancement and MAS frequency, and expect that targeted DNP will outperform conventional DNP strategies as improvements in instrumentation bring DNP into the fast MAS regime. A thorough understanding of the DNP transfer mechanisms in conjugated systems will be essential for the selective signal enhancement of proteins of various sizes and shapes and to prevent polarization "leakage" into the cellular background. We are also excited at the prospect of combining our targeted DNP approach with other developments in the field, such as higher magnetic fields,[52,53] time domain DNP experiments[54] and optical approaches[10] for the comprehensive description of protein structures in cells.

Acknowledgements

We would like to thank Dr. Ivan Sergeyev, Justin Lee, and Dr. Andy Borovik for help with EPR data acquisition, Dr. Peter Schultz, Dr. Minseob Koh, Dr. Carsten Schultz and Dr. Jan-Erik Hoffmann for help with amber suppression, and Brysa Alvarado and Melinda Serrato for assistance with TOTAPOL synthesis. B.A. was supported by NIH Molecular Biophysics Training Grant T32 GM008326. This work utilized the Biotechnology Research Center for NMR Molecular Imaging of Proteins at UCSD, supported by NIH grant P41 EB002031.

Conflict of Interest

The authors declare no conflict of interest.

Keywords: amber suppression · bioconjugation · NMR spectroscopy · protein chemistry · structural biology

[1] A. S. Lilly Thankamony, J. J. Wittmann, M. Kaushik, B. Corzilius, *Prog. Nucl. Magn. Reson. Spectrosc.* **2017**, *102–103*, 120–195.
[2] K. Jaudzems, T. Polenova, G. Pintacuda, H. Oschkinat, A. Lesage, *J. Struct. Biol.* **2019**, *206*, 90–98.
[3] K. K. Frederick, V. K. Michaelis, B. Corzilius, T. C. Ong, A. C. Jacavone, R. G. Griffin, S. Lindquist, *Cell* **2015**, *163*, 620–628.
[4] G. T. Debelouchina, M. J. Bayro, A. W. Fitzpatrick, V. Ladizhansky, M. T. Colvin, M. A. Caporini, C. P. Jaroniec, V. S. Bajaj, M. Rosay, C. E. Macphee, M. Vendruscolo, W. E. Maas, C. M. Dobson, R. G. Griffin, *J. Am. Chem. Soc.* **2013**, *135*, 19237–19247.
[5] L. B. Andreas, A. B. Barnes, B. Corzilius, J. J. Chou, E. A. Miller, M. Caporini, M. Rosay, R. G. Griffin, *Biochemistry* **2013**, *52*, 2774–2782.
[6] M. Kaplan, S. Narasimhan, C. de Heus, D. Mance, S. van Doorn, K. Houben, D. Popov-Celeketic, R. Damman, E. A. Katrukha, P. Jain, W. J. C. Geerts, A. J. R. Heck, G. E. Folkers, L. C. Kapitein, S. Lemeer, P. M. P. van Bergen en Henegouwen, M. Baldus, *Cell* **2016**, *167*, 1241–1251.
[7] W. Y. Chow, R. Li, I. Goldberga, D. G. Reid, R. Rajan, J. Clark, H. Oschkinat, M. J. Duer, R. Hayward, C. M. Shanahan, *Chem. Commun.* **2018**, *54*, 12570–12573.
[8] T. Azaïs, S. Von Euw, W. Ajili, S. Auzoux-Bordenave, P. Bertani, D. Gajan, L. Emsley, N. Nassif, A. Lesage, *Solid State Nucl. Magn. Reson.* **2019**, *102*, 2–11.
[9] W. N. Costello, Y. Xiao, K. K. Frederick, *Methods Enzymol.* **2019**, *615*, 373–406.
[10] B. J. Albert, C. Gao, E. L. Sesti, E. P. Saliba, N. Alaniva, F. J. Scott, S. T. Sigurdsson, A. B. Barnes, *Biochemistry* **2018**, *57*, 4741–4746.
[11] S. Narasimhan, S. Scherpe, A. Lucini Paioni, J. van der Zwan, G. E. Folkers, H. Ovaa, M. Baldus, *Angew. Chem. Int. Ed.* **2019**, *58*, 12969–12973; *Angew. Chem.* **2019**, *131*, 13103–13107.
[12] M. Renault, S. Pawsey, M. P. Bos, E. J. Koers, D. Nand, R. Tommassen-van Boxtel, M. Rosay, J. Tommassen, W. E. Maas, M. Baldus, *Angew. Chem. Int. Ed.* **2012**, *51*, 2998–3001; *Angew. Chem.* **2012**, *124*, 3053–3056.
[13] K. Yamamoto, M. A. Caporini, S. C. Im, L. Waskell, A. Ramamoorthy, *Biochim. Biophys. Acta Biomembr.* **2015**, *1848*, 342–349.
[14] R. Rogawski, A. E. McDermott, *Arch. Biochem. Biophys.* **2017**, *628*, 102–113.
[15] M. Kaushik, T. Bahrenberg, T. V. Can, M. A. Caporini, R. Silvers, J. Heiliger, A. A. Smith, H. Schwalbe, R. G. Griffin, B. Corzilius, *Phys. Chem. Chem. Phys.* **2016**, *18*, 27205–27218.
[16] M. A. Voinov, D. B. Good, M. E. Ward, S. Milikisiyants, A. Marek, M. A. Caporini, M. Rosay, R. A. Munro, M. Ljumovic, L. S. Brown, V. Ladizhansky, A. I. Smirnov, *J. Phys. Chem. B* **2015**, *119*, 10180–10190.
[17] E. A. van der Cruijsen, E. J. Koers, C. Sauvee, R. E. Hulse, M. Weingarth, O. Ouari, E. Perozo, P. Tordo, M. Baldus, *Chem. Eur. J.* **2015**, *21*, 12971–12977.
[18] T. Viennet, A. Viegas, A. Kuepper, S. Arens, V. Gelev, O. Petrov, T. N. Grossmann, H. Heise, M. Etzkorn, *Angew. Chem. Int. Ed.* **2016**, *55*, 10746–10750; *Angew. Chem.* **2016**, *128*, 10904–10908.
[19] R. Rogawski, I. V. Sergeyev, Y. Li, M. F. Ottaviani, V. Cornish, A. E. McDermott, *J. Phys. Chem. B* **2017**, *121*, 1169–1175.
[20] I. Marin-Montesinos, D. Goyard, E. Gillon, O. Renaudet, A. Imberty, S. Hediger, G. De Paepe, *Chem. Sci.* **2019**, *10*, 3366–3374.
[21] N. K. Devaraj, R. Weissleder, S. A. Hilderbrand, *Bioconjugate Chem.* **2008**, *19*, 2297–2299.
[22] M. L. Blackman, M. Royzen, J. M. Fox, *J. Am. Chem. Soc.* **2008**, *130*, 13518–13519.
[23] H. Wu, N. K. Devaraj, *Acc. Chem. Res.* **2018**, *51*, 1249–1259.
[24] J. E. Hoffmann, T. Plass, I. Nikic, I. V. Aramburu, C. Koehler, H. Gillandt, E. A. Lemke, C. Schultz, *Chem. Eur. J.* **2015**, *21*, 12266–12270.
[25] K. Lang, L. Davis, J. Torres-Kolbus, C. J. Chou, A. Deiters, J. W. Chin, *Nat. Chem.* **2012**, *4*, 298–304.
[26] T. Plass, S. Milles, C. Koehler, J. Szymanski, R. Mueller, M. Wiessler, C. Schultz, E. A. Lemke, *Angew. Chem. Int. Ed.* **2012**, *51*, 4166–4170; *Angew. Chem.* **2012**, *124*, 4242–4246.
[27] C. S. Song, K. N. Hu, C. G. Joo, T. M. Swager, R. G. Griffin, *J. Am. Chem. Soc.* **2006**, *128*, 11385–11390.
[28] L. Wang, A. Brock, B. Herberich, P. G. Schultz, *Science* **2001**, *292*, 498–500.
[29] E. G. Evans, G. L. Millhauser, *Methods Enzymol.* **2015**, *563*, 503–527.
[30] M. J. Schmidt, A. Fedoseev, D. Bucker, J. Borbas, C. Peter, M. Drescher, D. Summerer, *ACS Chem. Biol.* **2015**, *10*, 2764–2771.
[31] S. Tyagi, E. A. Lemke, *Curr. Opin. Struct. Biol.* **2015**, *32*, 66–73.
[32] A. Chatterjee, S. B. Sun, J. L. Furman, H. Xiao, P. G. Schultz, *Biochemistry* **2013**, *52*, 1828–1837.
[33] G. T. Debelouchina, M. J. Bayro, P. C. van der Wel, M. A. Caporini, A. B. Barnes, M. Rosay, W. E. Maas, R. G. Griffin, *Phys. Chem. Chem. Phys.* **2010**, *12*, 5911–5919.
[34] A. H. Linden, W. T. Franks, U. Akbey, S. Lange, B. J. van Rossum, H. Oschkinat, *J. Biomol. NMR* **2011**, *51*, 283–292.
[35] A. B. Siemer, K. Y. Huang, A. E. McDermott, *PLoS One* **2012**, *7*, e47242.
[36] G. Hou, S. Yan, J. Trebosc, J. P. Amoureux, T. Polenova, *J. Magn. Reson.* **2013**, *232*, 18–30.
[37] F. Mentink-Vigier, U. Akbey, Y. Hovav, S. Vega, H. Oschkinat, A. Feintuch, *J. Magn. Reson.* **2012**, *224*, 13–21.
[38] H. Takahashi, C. Fernandez-de-Alba, D. Lee, V. Maurel, S. Gambarelli, M. Bardet, S. Hediger, A. L. Barra, G. De Paepe, *J. Magn. Reson.* **2014**, *239*, 91–99.
[39] F. Mentink-Vigier, S. Paul, D. Lee, A. Feintuch, S. Hediger, S. Vega, G. De Paepe, *Phys. Chem. Chem. Phys.* **2015**, *17*, 21824–21836.
[40] K. R. Thurber, R. Tycko, *J. Chem. Phys.* **2014**, *140*, 184201.
[41] J. R. Wisniewski, M. Y. Hein, J. Cox, M. Mann, *Mol. Cell. Proteomics* **2014**, *13*, 3497–3506.
[42] P. C. van der Wel, K. N. Hu, J. Lewandowski, R. G. Griffin, *J. Am. Chem. Soc.* **2006**, *128*, 10840–10846.
[43] A. C. Pinon, J. Schlagnitweit, P. Berruyer, A. J. Rossini, M. Lelli, E. Socie, M. X. Tang, T. Pham, A. Lesage, S. Schantz, L. Emsley, *J. Phys. Chem. C* **2017**, *121*, 15993–16005.
[44] F. Mentink-Vigier, U. Akbey, H. Oschkinat, S. Vega, A. Feintuch, *J. Magn. Reson.* **2015**, *258*, 102–120.
[45] B. Corzilius, L. B. Andreas, A. A. Smith, Q. Z. Ni, R. G. Griffin, *J. Magn. Reson.* **2014**, *240*, 113–123.
[46] M. Rosay, L. Tometich, S. Pawsey, R. Bader, R. Schauwecker, M. Blank, P. M. Borchard, S. R. Cauffman, K. L. Felch, R. T. Weber, R. J. Temkin, R. G. Griffin, W. E. Maas, *Phys. Chem. Chem. Phys.* **2010**, *12*, 5850–5860.
[47] F. A. Perras, M. Pruski, *J. Chem. Phys.* **2018**, *149*, 154202.
[48] K. M. McCoy, R. Rogawski, O. Stovicek, A. E. McDermott, *J. Magn. Reson.* **2019**, *303*, 115–120.
[49] E. Luchinat, L. Banci, *Acc. Chem. Res.* **2018**, *51*, 1550–1557.
[50] J. W. Chin, *Nature* **2017**, *550*, 53–60.
[51] C. Sauvée, M. Rosay, G. Casano, F. Aussenac, R. T. Weber, O. Ouari, P. Tordo, *Angew. Chem. Int. Ed.* **2013**, *52*, 10858–10861; *Angew. Chem.* **2013**, *125*, 11058–11061.
[52] P. Fricke, D. Mance, V. Chevelkov, K. Giller, S. Becker, M. Baldus, A. Lange, *J. Biomol. NMR* **2016**, *65*, 121–126.
[53] A. B. Barnes, E. Markhasin, E. Daviso, V. K. Michaelis, E. A. Nanni, S. K. Jawla, E. L. Mena, R. DeRocher, A. Thakkar, P. P. Woskov, J. Herzfeld, R. J. Temkin, R. G. Griffin, *J. Magn. Reson.* **2012**, *224*, 1–7.
[54] K. O. Tan, C. Yang, R. T. Weber, G. Mathies, R. G. Griffin, *Sci. Adv.* **2019**, *5*, eaav6909.

Manuscript received: October 6, 2019
Accepted manuscript online: November 19, 2019
Version of record online: January 21, 2020

Supplementary methods

Materials and equipment

All reagents and solvents were purchased from Sigma Aldrich (St. Louis, MO), Thermo-Fisher Scientific (including Fisher Scientific and Acros Organics, Waltham, MA) and used without further purification unless otherwise noted. Isotopically enriched reagents and solvents for NMR were purchased from Cambridge Isotope Laboratories (Tewksbury, MA). Thin layer chromatography analyses were carried out using aluminum-backed, precoated silica gel plates (Merck TLC silica gel 60 F254) from EMD Millipore (Billerica, MA). Column chromatography was performed using Acros Organics Silica gel (0.035-0.070 mm, 60 Å). ^{1}H and ^{13}C NMR of small molecules were performed on a Varian Vx 500 MHz spectrometer. Primer synthesis and DNA sequencing were performed by Integrated DNA Technologies (Coralville, IA) and Genewiz (South Plainfield, NJ), respectively. PCR was performed using a Bio-Rad T100 thermocycler (Hercules, CA), while purification of the DNA relied on kits from Biomiga (San Diego, CA) and New England BioLabs (Ipswich, MA). Dialysis kits were purchased from Thermo Fisher Scientific and protein concentrators were obtained from Sartorius (Goettingen, Germany). Reverse-phase (RP) HPLC was performed on a 2545 Binary Gradient Module Waters system equipped with a 2484 UV/vis detector from Waters Corporation (Milford, MA). For prep-scale RP-HPLC purification, we relied on a Waters XBridge BEH C18 19 mm x 250 mm, 10 µm particle size column, while for analytical measurements we used a Waters Symmetry 300 C18 4.6 mm x 150 mm, 5 µm particle size column. HPLC solvent A contained 100% H_2O + 0.1% trifluoroacetic acid (TFA), while solvent B contained 100% acetonitrile + 0.1 % TFA. HR-MS and LC-ESI-TOF MS analysis was conducted on an Agilent 6230 TOF-MS with Jet Stream ESI. Gel images were acquired using a camera and light box from Fotodyne Incorporated (Hartland, WI). Protein and fluorophore absorbances were measured using a Nanodrop One Spectrophotometer by Thermo-Fisher Scientific.

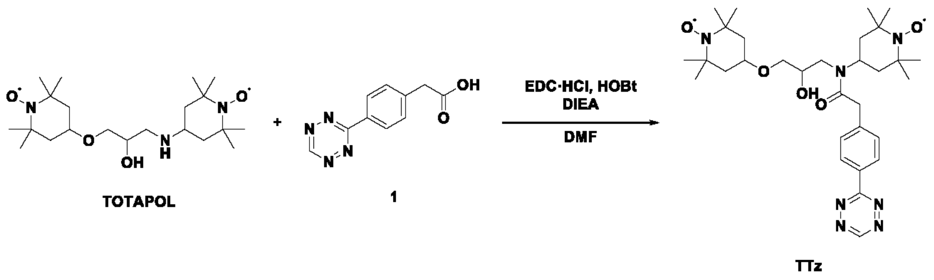
Synthesis of polarization agents and UAAs

Synthesis of norbornene lysine

Norbornene lysine was synthesized according to the protocol reported by Plass et al.[1] The final product was analyzed by solution NMR spectroscopy and high-resolution mass spectrometry (HR-

MS). ^{1}H NMR (500 MHz, DMSO-d_6): 7.15 (0.5H, t, *J* = 5.5 Hz), 7.10 (0.5H, t, *J* = 5.5 Hz), 6.16 (0.5H, dd, *J* = 5.6, 3.0 Hz), 6.11-6.07 (1H, m), 5.92 (0.5H, dd, *J* = 5.6, 2.8 Hz), 4.02–3.97 (m, 0.5H), 3.87–3.81 (m, 0.5H), 3.69–3.64 (m, 0.5H), 3.51–3.48 (2H, m), 3.00-2.90 (2H, m), 2.82–2.75 (2H, m), 2.65 (0.5H s), 2.33–2.25 (0.5H, m), 1.80–1.66 (3H, m), 1.62–1.53 (0.5H, m), 1.35-1.20 (3.5H, m), 1.10–1.1 (1H, m), 0.45 (0.5H, ddd, *J* = 11.5, 4.1, 2.6 Hz). ^{13}C NMR (125 MHz, DMSO-d_6): δ 171.24, 156.45, 156.40, 137.40, 136.89, 136.26, 132.28, 67.68, 67.11, 51.96, 49.03, 44.70, 43.49, 43.24, 41.78, 41.15, 38.22, 37.94, 29.77, 29.12, 28.97, 28.76, 21.70. HR-MS (ESI): calculated m/z for $[C_{15}H_{25}N_2O_4]^+$ ($[M+H]^+$) is 297.1736, experimentally determined value is 297.1809.

Synthesis of TOTAPOL-tetrazine (TTz)



TOTAPOL was synthesized according to the protocol reported by Song et al.[2] Compound **1** was synthesized according to the protocol reported by Yang et al.[3] Compound **1** (100 mg, 0.463 mmol), *N*-(3-dimethylaminopropyl)-*N'*-ethylcarbodiimide hydrochloride (EDC-HCl) (177 mg, 0.925 mmol), 1-hydroxybenzotriazole (HOBt) (125 mg, 0.925 mmol), and *N,N*-diisopropylethylamine (DIEA) (161 mL, 0.925 mmol) were dissolved in dry *N,N*-dimethylformamide (DMF) (5 mL, 0.1 M). TOTAPOL (222 mg, 0.555 mmol) was then added to the solution and the reaction mixture was stirred for 6 hours at room temperature. After completion, the reaction was analyzed by thin layer chromatography (dichloromethane:methanol = 10:1, R_f = 0.38), and ethyl acetate (20 mL) was added to the solution. The organic layer was washed with water (3 times) and brine, then dried over anhydrous sodium sulfate. The crude mixture was filtered, and the solvent was evaporated under reduced pressure. The final product was purified by silica flash column chromatography using ethyl acetate:hexanes (1:2 to 2:1 gradient) as pink solid (yield: 62 %). HR-MS (ESI): calculated m/z for [$C_{31}H_{47}N_7O_5Na^+$] ($[M+Na]^+$) is 620.3531, measured m/z is 620.3525. HPLC trace and EPR spectrum are shown in **SI Fig 1**.

Preparation of plasmid constructs

pULTRA tRNA^{Pyl}/PylRS^{AF} plasmid

This plasmid encodes the tRNA (tRNA^{Pyl}) and the tRNA synthetase (PylRS^{AF}) for norbornene lysine incorporation by amber suppression. After initial tests and issues with ^{13}C label dilution in M9 media, we designed a construct where the pULTRA vector backbone from Chatterjee et al.[4] was combined with the tRNA^{Pyl}/PylRS^{AF} insert sequence from Plass et al.[1] The new plasmid was constructed using the NEBuilder® HiFi DNA assembly. Protein expression in this construct is controlled by an IPTG-inducible *lacI* promoter.

Forward (vector): ATGGACAAAAAACCGCTGAA

Reverse (vector): TTACAGGTTCGTAGAGATCC

Forward (PylRS): GGATCTCTACGAACCTGTAATAAGCGGCCGCGTTTAAACGG

Reverse (PylRS): TTCAGCGGTTTTTTGTCCATGCGGCCGCACCTCCT

Protein sequence (PylRS):
MDKKPLNTLISATGLWMSRTGTIHKIKHHEVSRSKIYIEMACGDHLVVNNSRSSRTARALRHHKYRKTCKRCR
VSDEDLNKFLTKANEDQTSVKVKVVSAPTRTKKAMPKSVARAPKPLENTEAAQAQPSGSKFSPAIPVSTQES
VSVPASVSTSISSISTGATASALVKGNTNPITSMSAPVQASAPALTKSQTDRLEVLLNPKDEISLNSGKPFRELES
ELLSRRKKDLQQIYAEERENYLGKLEREITRFFVDRGFLEIKSPILIPLEYIERMGIDNDTELSKQIFRVDKNFCLRP
MLAPNLANYLRKLDRALPDPIKIFEIGPCYRKESDGKEHLEEFTMLNFCQMGSGCTRENLESIITDFLNHLGIDF
KIVGDSCMVFGDTLDVMHGDLELSSAVVGPIPLDREWGIDKPWIGAGFGLERLLKVKHDFKNIKRAARSESY
YNGISTNL*

WT Ubiquitin plasmid

WT ubiquitin was used as a control in samples prepared with dispersed TOTAPOL. Here we used the pTXB1-Ub-NpuDnaE-His_6 construct described previously.[5] This construct encodes the native ubiquitin sequence, followed by the NpuDnaE intein and a His_6 tag. The His_6 tag is useful in amber suppression (see below) as it allows the purification of the full-length protein from prematurely truncated polypeptides, while the intein enables the traceless removal of the tag without the use of proteases.[6-7]

K6X Ubiquitin plasmid

This construct contained the amber stop codon (TAG) in place of the lysine codon (AAG) at position 6 of the ubiquitin sequence. Site-directed mutagenesis was performed using the NEB Q5 kit, the pTXB1-Ub-NpuDnaE-His$_6$ plasmid and the following primers:

Forward: GATCTTCGTG***TAG***ACTCTGACTG

Reverse: TGCATATGTATATCTCCTTC

Protein sequence:

MQIFV*TLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGCLSYETEILTVEYGLLPIGKIVEKRIECTVYSVDNNGNIYTQPVAQWHDRGEQEVFEYCLEDGSLIRATKDHKFMTVDGQMLPIDEIFERELDLMRVDNLPNIKIATRKYLGKQNVYDIGVERDHNFALKNGFIASA_AFNHHHHHH*_

K6X Ubiquitin-His$_6$ plasmid

This plasmid was prepared from the pTXB1-K6X-Ub-NpuDnaE-His$_6$ construct where the NpuDnaE sequence was deleted using the Q5 mutagenesis kit from New England Biolabs. This produced plasmid pTXB1-K6X-Ub-His$_6$. We used this construct for experiments in bacterial lysates where ubiquitin could be pulled from the lysate if desired using the His$_6$ tag. We removed the intein from the construct to avoid premature intein hydrolysis and to increase the lifetime of the protein in the lysate.

Forward: CATCATCATCATCATCATTAAA

Reverse: ACCACCTCTGAGACG

Sequence:

MQIFV*TLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRGG_HHHHHH*_

WT His$_6$-TEV-HP1α plasmid

The HP1α gene was excised from a GST HP1α plasmid provided by Naoko Tanese[8] (Addgene plasmid # 24074 ; http://n2t.net/addgene:24074 ; RRID:Addgene_24074) and cloned into the pET vector backbone of a 2BT MacroLab plasmid (generously provided by Dr. Kevin Corbett) using NEBuilder® HiFi DNA Assembly Master Mix. The resulting plasmid contained a His6-TEV-HP1α construct.

Protein sequence:

MKSSHHHHHHENLYFQSGKKTKRTADSSSSEDEEEYVVEKVLDRRVVKGQVEYLLKWKGFSEEHNTWEPEK
NLDCPELISEFMKKYKKMKEGENNKPREKSESNKRKSNFSNSADDIKSKKKREQSNDIARGFERGLEPEKIIGA
TDSCGDLMFLMKWKDTDEADLVLAKEANVKCPQIVIAFYEERLTWHAYPEDAENKEKETAKS*

His$_6$-TEV-HP1α-2X plasmid

A TAG codon was inserted at position 2 of HP1α using the NEBuilder® HiFi DNA Assembly Master Mix.

Forward: CAATCCGGA***TAG***AAGAAAACCAAGCGGACA

Reverse: TTTTCTT***CTA***TCCCGGATTGGAAGTACAGGTT

Protein sequence:

MKSSHHHHHHENLYFQSG*KKTKRTADSSSSEDEEEYVVEKVLDRRVVKGQVEYLLKWKGFSEEHNTWEPE
KNLDCPELISEFMKKYKKMKEGENNKPREKSESNKRKSNFSNSADDIKSKKKREQSNDIARGFERGLEPEKIIG
ATDSCGDLMFLMKWKDTDEADLVLAKEANVKCPQIVIAFYEERLTWHAYPEDAENKEKETAKS*

WT SMC-His$_6$ plasmid

This construct was purchased from GeneArt Gene Synthesis by ThermoFisher Scientific. The encoding region contains WT Pyrococcus yayanosii SMC (345-468, 694-814) with a SGGS linker and a C-terminal His$_6$ tag. The gene was transferred into the pET vector backbone of the 2BT MacroLab plasmid using NEBuilder® HiFi DNA Assembly Master Mix.

Protein sequence:

MTKGAIVRWGKRKEKLIEEIRAREEERNALVCRLGEIDRTFAVAREEFDTVVKELEEARKSLYEGEARIKRAEEE
KERLKAEILTGEARLPGLRERAENLRRLVEEKRAEISELERRLSSITSKRSGGSLESQSFELRIKLSDLEKELELARKD
LEKVLAEERAVREEIEVAKRRINELDTLIERERGELAKLRGRIERLERKRDKLKKALENPEARELTEKIRAVEKEIAA
LREELSRVEGKLEGLESKGGHHHHHH*

SMC V6X-His$_6$ plasmid

Site-directed mutagenesis by HiFi was used to introduce the TAG codon at position 6 of the SMC sequence.

Forward: GCCATT***TAG***CGTTGGGGTAAACGTAAAGA

Reverse: CCCCAACG***CTA***AATGGCACCTTTGGTCATA

Protein sequence:

MTKGAI*RWGKRKEKLIEEIRAREEERNALVCRLGEIDRTFAVAREEFDTVVKELEEARKSLYEGEARIKRAEEEKERLKAEILTGEARLPGLRERAENLRRLVEEKRAEISELERRLSSITSKRSGGSLESQSFELRIKLSDLEKELELARKDLEKVLAEERAVREEIEVAKRRINELDTLIERERGELAKLRGRIERLERKRDKLKKALENPEARELTEKIRAVEKEIAALREELSRVEGKLEGLESKGGHHHHHH*

Protein expression and purification

K6X Ubiquitin

The Ub-K6X-NpuDnaE-His$_6$ plasmid was co-transformed with the pULTRA-tRNA^{Pyl}/PylRS^{AF} plasmid into BL21 (DE3) competent cells. The cells were grown on an LB agar plate with the presence of ampicillin (100 µg/mL) and spectinomycin (50 µg/mL). After overnight incubation, a colony was selected and inoculated into 5 mL LB media containing antibiotics. After overnight incubation at 37 °C, 1 mL of the preculture was inoculated into 1 L of M9 minimal media supplemented with ^{15}N-labeled ammonium chloride, ^{13}C-labeled glucose and antibiotics. The culture was grown at 37 °C with shaking until the OD$_{600}$ reached 0.3 – 0.5. At this point, 10 mL of 0.1 M norbornene lysine stock solution (in 0.5 M NaOH, 1 mM final) was added to the culture and incubated further for 0.5 – 1 h with OD$_{600}$ reaching the 0.5-0.7 range. 1 mL of 0.5 M IPTG (0.5 mM final) was added to the culture and the protein was expressed at 37 °C for 4 hours. After expression, the cells were harvested by centrifugation at 5,000 ×*g* and resuspended in 20 mL of cold lysis buffer (50 mM phosphate, 300 mM NaCl, 5 mM imidazole, 1 mM TCEP, Pierce protease inhibitor tablet, pH 7.5). The cells were lysed by sonication and the lysate was clarified by centrifugation at 30,000 ×*g*. The supernatant was incubated with 2 mL of Ni-NTA beads for one hour, washed with lysis buffer containing increasing amounts of imidazole (5, 20, and 50 mM imidazole, 20 mL each), and eluted with 12 mL of lysis buffer containing 250 mM imidazole. These steps were performed at 4 °C to avoid premature cleavage of the NpuDnaE intein purification tag. At this point, 120 μL of 1M TCEP (final 10 mM) was added to the elute and incubated at room temperature overnight to cleave the NpuDnaE-His_6 intein and to generate ubiquitin without a His_6 tag. The sample was then loaded onto a semi-prep C18 column and purified by RP-HPLC. The purified protein was lyophilized and dissolved in 1 mL of 6M guanidine hydrochloride. The solution was dialyzed against 10 mM Tris·KCl buffer (10 mM KCl, pH 7.5) to refold the ubiquitin-norbornene construct. The final sample was analyzed by analytical RP-HPLC and ESI-MS, as well as by solution NMR to confirm the correct fold of the protein (**SI Fig. 2**).

WT Ubiquitin

In the case of WT ubiquitin, the pTXB1-Ub-NpuDnaE-His_6 plasmid was transformed alone into BL21 (DE3) competent cells. Expression and purification were performed as described above except without the addition of spectinomycin and norbornene lysine.

HP1α-2X

BL21(DE3)-Rosetta cells were co-transformed with the HP1α and pUltra plasmids. Pre-cultures were grown at 37 °C overnight under ampicillin and spectinomycin antibiotic selection in LB. Pre-cultures were pelleted in the morning, LB was removed and the pellet was resuspended into ^{13}C, ^{15}N enriched M9 media. Cells were grown at 37 °C until OD600 = 0.8, then induced with 0.5 mM IPTG and 1 mM norbornene lysine. Cells were collected 20 hours later by centrifuging at 5,000 *xg*. The cell pellet was resuspended in lysis buffer (1x PBS, pH 7.4, 300 mM NaCl, 10% glycerol, 7.5 mM imidazole, Roche protease inhibitor tablet) and sonicated at 4 °C. Cell debris were cleared by centrifugation for 30 minutes at 30,000 *xg*. Approximately 5 mL of Ni-NTA resin (per 1 L culture) was added to the lysate and rotated at 4 °C for one hour. The lysate-resin mixture was loaded into a plastic Bio-Rad Econo column and washed with buffer (1x PBS, pH 7.4, 300 mM NaCl, 10% glycerol, 7.5 mM imidazole). Elution buffer (20 mM HEPES, pH 7.2, 300 mM KCl, 400 mM imidazole, 1 mM DTT) was then poured over the column. The eluant was adjusted to 6 M guanidine-hydrochloride and pH 3.0, then applied to a Waters XBridge BEH C18 prep-size RP-HPLC column and fractions were collected using a gradient of 10%-60% solvent B. HP1α eluted around 40% solvent B and due to protein retention on the column, multiple runs were required to collect the protein. The elution product was lyophilized, then refolded by initial resuspension in 6M guanidinium, 20 mM HEPES, pH 7.2, 75 mM KCl and 1 mM TCEP, followed by overnight dialysis against 20 mM HEPES, pH 7.2, 75 mM KCl. Dialysis ensured the removal of TCEP which is necessary prior to reaction with TTz. The final sample was analyzed by analytical RP-HPLC and ESI-MS.

WT HP1α

The purification of WT HP1α followed the protocol described above except without the addition of spectinomycin and norbornene-lysine. Additionally, TCEP was kept in the sample to maintain long-term stability.

SMC-V6X

Protein expression and cell lysis followed the protocol for HP1α-2X. After clearing the lysate by centrifugation at 30,000 *xg* for 30 min, the lysate supernatant was mixed with Ni-NTA resin, rotated for 1 hr at 4 °C and loaded into a plastic Bio-Rad Econo column. The resin was washed with buffer containing 1x PBS, pH 7.4, 300 mM NaCl, 10% glycerol, 7.5 mM imidazole, and protein elution was performed with buffer containing 20 mM Tris, pH 7.4, 50 mM NaCl, 400 mM imidazole, 1 mM DTT. The elute was adjusted to 6 M guanidine-hydrochloride and pH 3.0, then loaded onto a Waters XBridge BEH C18 prep-size RP-HPLC column. Fractions were collected using a gradient of 10%-60% solvent B (100% ACN + 0.1% TFA). SMC eluted around 40% solvent B. The product was lyophilized, then refolded by initial resuspension in 6M guanidinium, 20 mM Tris, pH 7.4, 50 mM NaCl and 1 mM TCEP. Overnight dialysis against 20 mM Tris, pH 7.4, 50 mM NaCl was performed. Dialysis ensured the removal of TCEP which is necessary prior to reaction with TTz. The final sample was analyzed by analytical RP-HPLC and ESI-MS.

WT SMC

The purification of WT SMC followed the protocol described above except without the addition of spectinomycin and norbornene-lysine.

Conjugation of TTz to proteins

Ub-TTz

Lyophilized K6X ubiquitin was dissolved in 1 mL of 6 M guanidine hydrochloride buffer (10 mM Tris-HCl, 10 mM KCl) to a final concentration of 100 μM. TTz was dissolved in DMSO to prepare a 5 mM stock solution and 100 μL of the stock were added to the ubiquitin solution for a final TTz concentration of 500 μM. The mixture was mixed gently overnight and the completion of the reaction was confirmed by the absence of a K6X ubiquitin peak in the analytical RP-HPLC trace (**SI Fig. 3**). The solution was then dialyzed against 10 mM Tris, 10 mM KCl, pH 7.5 buffer to refold the protein and remove excess TTz. The complete removal of unbound TTz was confirmed by analytical RP-HPLC and the final conjugated Ub-TTz product was confirmed by ESI-MS (**SI Fig. 4**). Since tetrazine-norbornene conjugates have absorbance at A280, we determined protein concentration in the following way. Serial dilutions of WT ubiquitin and Ub-TTz were loaded on an SDS-PAGE gel and the density of the bands and the A280 absorbance of the samples were compared. Thus, it was determined that the same amount of Ub-TTz has 8-fold higher absorbance at A280 compared to WT ubiquitin, producing an extinction coefficient of 10,240 $cm^{-1}\cdot M^{-1}$ for Ub-TTz vs 1280 $cm^{-1}\cdot M^{-1}$ for WT ubiquitin.

HP1α-TTz

Refolded HP1α-2X was dialyzed against 20 mM HEPES, pH 7.2, 75 mM KCl and diluted to 20 μM. 200 μM TTz from a 50 mM stock was added to the HP1α solution. The mixture was gently rotated for one hour until reaction completion was confirmed by the retention time shift of the HP1α-2X RP-HPLC peak (**SI Fig. 3)**. The solution was then concentrated to both concentrate HP1α and remove excess TTz. We found that TTz is completely removed after several buffer exchanges (**SI Fig. 3 & 4**). The final concentration of HP1α-TTz was determined by SDS-PAGE gel analysis as performed for Ub-TTz.

SMC-TTz

Refolded SMC-V6X was dialyzed against 20 mM Tris, pH 7.4, 50 mM NaCl and diluted to 25 μM. 250 μM TTz from a 50 mM stock was added to the SMC solution. The mixture was gently rotated for three hours until reaction completion was confirmed by the retention time shift of the SMC-V6X RP-HPLC peak (**SI Fig. 3)**. The solution was then concentrated to both concentrate SMC and remove excess TTz (**SI Fig. 3 & 4**). The final concentration of SMC-TTz was determined by SDS-PAGE gel analysis as done performed Ub-TTz.

DNP sample preparation

Ub-TTz

Refolded Ub-TTz in 10 mM Tris, 10 mM KCl was concentrated and the protonated buffer was exchanged several times with deuterated buffer with the same composition (10 mM Tris-d_{11}, 10 mM KCl, pD 7.5 in D_2O). The final concentration of protein was 15 mg/mL (1.66 mM) and 97% deuteration of the solvent was achieved (D_2O:H_2O = 97:3). The sample was diluted with deuterated buffer to the desired protein concentration and 9 μL of the sample were combined with 1 μL of glycerol-d_8, and transferred into a 1.9 mm DNP zirconia rotor.

WT ubiquitin with 15 mM TOTAPOL

Refolded WT ubiquitin in 10 mM Tris, 10 mM KCl, pH 7.5 was concentrated and the protonated buffer was exchanged several times with 60% glycerol-d_8 in 90% deuterated buffer with the same composition (glycerol-d_8:D_2O:H_2O = 60:4:36). The final concentration of protein was 11.2 mg/mL (1.28 mM). The sample was diluted with the deuterated buffer (final 10.5 mg/mL) and 9.5 μL of the sample were combined with 0.5 μL of 300 mM TOTAPOL in 90% deuterated buffer with the same composition, and transferred into a 1.9 mm DNP zirconia rotor.

HP1α-TTz

HP1α-TTz was concentrated and the protonated buffer was exchanged multiple times into 20 mM HEPES, 50 mM KCl, pD 7.2 in 97% D_2O. The final concentration of protein was 200 μM. The sample was then diluted to 50 μM in deuterated buffer and 10% (^{13}C-depleted 0.05%) glycerol-d_8 and transferred into a 1.9 mm DNP zirconia rotor.

SMC-TTz

SMC-TTz was concentrated and the protonated buffer was exchanged multiple times into 20 mM Tris, 50 mM NaCl, pD 7.4 in 97% D_2O. The final concentration of protein was 250 μM. The sample was diluted to 50 μM in deuterated buffer and 10% (^{13}C-depleted 0.05%) glycerol-d_8 and transferred into a 1.9 mm DNP zirconia rotor.

Solution NMR experiments

Experiments were performed using a Bruker Avance III 600 MHz (^{1}H Larmor frequency) NMR spectrometer equipped with a triple resonance probe. Chemical shifts were referenced to the TMS frequency. The HSQC spectra were acquired using the standard Bruker fhsqcf3gpph pulse sequence. The HSQC spectrum was acquired with 16 scans while using the following parameters: 298 K, 2048 points and 122 ms acquisition in the direct dimension, 256 points and 58 ms acquisition in the indirect dimension, interscan delay of 1.5 s, ^{1}H center frequency of 4.7 ppm, ^{15}N center frequency of 118 ppm.

MAS DNP NMR experiments

Experiments were performed on a 600 MHz Bruker DNP NMR spectrometer equipped with a NEO console and a 395 GHz gyrotron for high-power microwave irradiation. All spectra were recorded using a triple resonance (^{1}H, ^{13}C, ^{15}N) 1.9 mm MAS Bruker probe. The magic angle was set with KBr and the ^{13}C chemical shift was referenced to adamantane[9] (40.49 ppm) at room temperature or the rotor silicon plug ^{13}C peak (3.2 ppm) at low temperature. Sample temperature was maintained at 100 K, at either 12 or 24 kHz. Temperature was confirmed by the T1 of a KBr standard.[10] The largest signal enhancements were recorded at microwave power of 5 W.

1D ^{1}H-^{13}C CP experiments for enhancement determination had the following parameters: 100 K, 12 or 24 kHz MAS, 5 s interscan delay, 20 ms acquisition, 2048 points, ^{13}C center frequency set at 100 ppm. Scans were tailored for each sample to achieve sufficient signal-to-noise in the microwave off spectrum. For CP, a 500 µs contact time was used with the Hartmann-Hahn matching condition satisfied by setting the ^{13}C power to 66 kHz and optimizing on the ^{1}H channel. 100 kHz spinal64 decoupling[11] was performed during acquisition. DNP build-ups were measured with this 1D ^{1}H-^{13}C CP experiment but with varied interscan delays until saturation.

The 2D ^{13}C -^{13}C CORD experiment used the CORDxy4 pulse sequence[12], with the following parameters: 100 K, 24 kHz MAS, 64 scans, 5 s interscan delay, 1536 points and 15 ms acquisition in the direct dimension, 288 points and 4 ms acquisition in the indirect dimension, ^{13}C center frequency of 100 ppm. The CORD conditions were set to 24 kHz ($1*\omega_r$) and 12 kHz ($0.5*\omega_r$). For CP, a 500 µs contact time was used, together with 50 ms of CORD mixing. 100 kHz spinal64 decoupling was performed during acquisition. To process the spectrum, a GM window function was used with 75 Hz line broadening in each dimension.

DNP of bacterial lysates

Sample preparation

The K6X Ubiquitin-His$_6$ plasmid and the pULTRA-tRNA^{Pyl}/PylRS^{AF} plasmid were co-transformed into BL21 (DE3) competent cells. We followed the expression protocol for K6X ubiquitin as described above with the exception that we only used 200 mL of M9 media enriched with ^{13}C-glucose and ^{15}N-labeled ammonium chloride. After the end of the expression period, the cells were lysed by sonication in 5 mL lysis buffer containing 50 mM phosphate, 300 mM NaCl, 5 mM imidazole, 1 mM TCEP, Pierce protease inhibitor tablet, pH 7.5. After lysis, the solution was dialyzed overnight against buffer containing 20 mM phosphate and 150 mM NaCl, pH 7.2. This step was performed to remove endogenous thiol molecules as they can reduce the lifetime of nitroxide-based polarization agents.[13] At this point, 50 uL of 100 mM TTz in DMSO (final 1 mM) were added to the dialyzed solution and the mixture was mixed gently. After 8 hours, 50 µL of the reaction mixture was removed and diluted 20 times in deuterated phosphate buffer with the same salt composition as above. The diluted solution was concentrated using a 3,000 MWCO concentration device and the buffer was exchanged several times to achieve 95% deuteration level. Complete removal of unbound TTz was confirmed by the disappearance of the TTz peak in the analytical RP-HPLC trace. To prepare the sample for DNP, 9 μL of the sample were combined with 1 μL of glycerol-d_8 and the solution was transferred into a 1.9 mm zirconia rotor.

To determine the amount of ubiquitin in the lysate, we resolved the protein content of the lysate sample by SDS-PAGE gel electrophoresis (**SI Fig. 10**). The intensities of all visible bands were quantified using the ImageJ analysis software and the intensity of the ubiquitin band was divided by the combined intensity of all the other protein bands in the gel, resulting in ~5% ubiquitin content. Furthermore, the intensity of the ubiquitin band was compared to the intensities of known amounts of purified ubiquitin loaded on the same gel. This gave an estimate of the absolute amount of ubiquitin in the lysate DNP sample, which was ~ 50 μg.

Finally, to determine the efficiency of the TTz-norbornene reaction in this context, we used Ni-NTA affinity column purification to remove ubiquitin from the lysate. The purified ubiquitin was then subjected to LC-MS, confirming the presence of substantial amounts of Ub-TTz (**SI Fig. 10**).

Analysis of the DNP signal enhancements

One-dimensional ^{1}H-^{13}C cross-polarization experiments were acquired with and without microwaves. After applying line broadening of 100 Hz and polynomial baseline correction, the signal intensity at the largest Cα peak (54.4 ppm for Ub, 57.4 ppm for SMC, 54.7 ppm for HP1α) was measured for the on and off states. The on to off ratio is reported as the enhancement ε .

Quantification of the selective DNP enhancements in lysates

To calculate the enhancement of Ub-TTz in lysates, we followed a method previously described by Viennet et al.[14] We began by measuring the peak intensity at 54.4 ppm for the ON signal (6.11×10^9) and OFF signal (2.44×10^9). The OFF signal was then multiplied by two (4.88×10^9) to match the intensity of the glycerol signals from each spectrum. The glycerol peaks arise from the sample solvent and we assume that their enhancement represents the non-selective enhancement of the lysate background. The difference between the ON signal and 2x the OFF signal was 1.22×10^9 and represents the enhancement on ubiquitin. Since ubiquitin is only 5% of the total protein in the sample, we assumed that only 5% of the microwave OFF signal is due to ubiquitin (1.22×10^8). Therefore, the "true" enhancement of ubiquitin in the sample is the ratio of 1.22×10^9 and

1.22 x 10^8 , i.e. it is equal to 10. We divided this by two to account for the two-fold enhancement on the background glycerol signals, which then produces a selective ubiquitin enhancement equal to 5.

Electron paramagnetic resonance spectroscopy

EPR spectra were acquired on a 9 GHz EMX Bruker EPR spectrometer at room temperature. We recorded the EPR spectra on samples taken directly from the DNP NMR rotors and placed in quartz capillaries. The sample length in the capillary spanned the whole length of the microwave cavity so that the concentrations of radicals in all samples could be compared to a 50 and 100 µM TOTAPOL standards.

Calculation of the distance between polarization agents at various concentrations

We followed the application of the Wigner-Seitz radius as done previously by Pinon et al.[15]

$$d = 2\left(\frac{3}{4\pi C N_A}\right)^{1/3}$$

Here d is distance between the biradical polarization agents, C is the concentration of polarization agents, N_A is Avogadro's number.

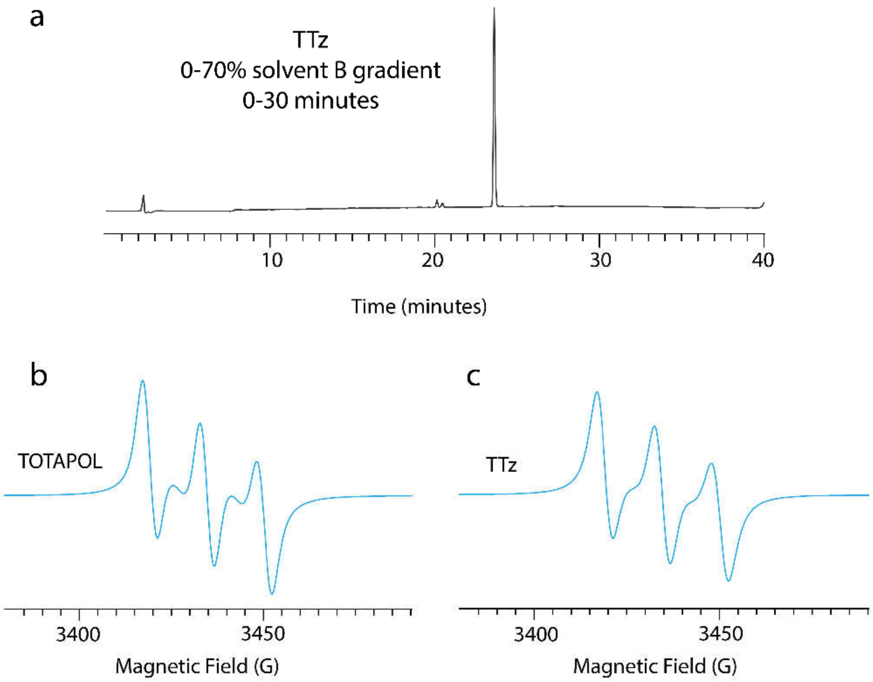
Analysis of the polarization buildup curves

Peak intensities at the 54.4 ppm Cα peak were measured using TopSpin 4.0.5. Using GraphPad Prism version 8.2, data were fit to either a monoexponential association model, $I = I_0 + \left(1 - e^{-\frac{t}{\tau}}\right)$ or a biexponential association model, $I = I_0 + a\left(1 - e^{-\frac{t}{\tau_1}}\right) + b\left(1 - e^{-\frac{t}{\tau_2}}\right)$ where a and b are fit parameters that account for the fraction of fast and slow components and τ_1 and τ_2 are the build-up time constants of those components. Here I_0 = 0.

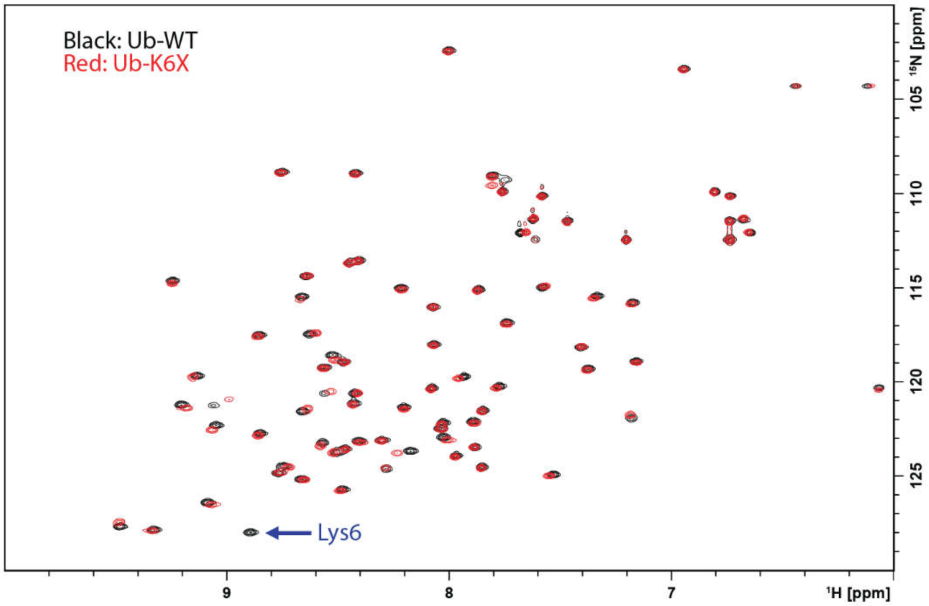
References

1. Plass, T.; Milles, S.; Koehler, C.; Szymanski, J.; Mueller, R.; Wiessler, M.; Schultz, C.; Lemke, E. A., Amino acids for Diels-Alder reactions in living cells. *Angew Chem Int Ed Engl* **2012,** *51* (17), 4166-70.
2. Song, C. S.; Hu, K. N.; Joo, C. G.; Swager, T. M.; Griffin, R. G., TOTAPOL: A biradical polarizing agent for dynamic nuclear polarization experiments in aqueous media. *J Am Chem Soc* **2006,** *128* (35), 11385-11390.
3. Yang, J.; Karver, M. R.; Li, W.; Sahu, S.; Devaraj, N. K., Metal-catalyzed one-pot synthesis of tetrazines directly from aliphatic nitriles and hydrazine. *Angew Chem Int Ed Engl* **2012,** *51* (21), 5222-5.
4. Chatterjee, A.; Sun, S. B.; Furman, J. L.; Xiao, H.; Schultz, P. G., A versatile platform for single- and multiple-unnatural amino acid mutagenesis in Escherichia coli. *Biochemistry* **2013,** *52* (10), 1828-37.
5. Shah, N. H.; Dann, G. P.; Vila-Perello, M.; Liu, Z. H.; Muir, T. W., Ultrafast Protein Splicing is Common among Cyanobacterial Split Inteins: Implications for Protein Engineering. *Journal of the American Chemical Society* **2012,** *134* (28), 11338-11341.
6. Debelouchina, G. T.; Gerecht, K.; Muir, T. W., Ubiquitin utilizes an acidic surface patch to alter chromatin structure. *Nat Chem Biol* **2017,** *13* (1), 105-110.
7. Batjargal, S.; Walters, C. R.; Petersson, E. J., Inteins as Traceless Purification Tags for Unnatural Amino Acid Proteins. *Journal of the American Chemical Society* **2015,** *137* (5), 1734-1737.
8. Vassallo, M. F.; Tanese, N., Isoform-specific interaction of HP1 with human TAFII130. *Proc Natl Acad Sci U S A* **2002,** *99* (9), 5919-24.
9. Morcombe, C. R.; Zilm, K. W., Chemical shift referencing in MAS solid state NMR. *J Magn Reson* **2003,** *162* (2), 479-86.
10. Thurber, K. R.; Tycko, R., Measurement of sample temperatures under magic-angle spinning from the chemical shift and spin-lattice relaxation rate of 79Br in KBr powder. *J Magn Reson* **2009,** *196* (1), 84-7.
11. Fung, B. M.; Khitrin, A. K.; Ermolaev, K., An improved broadband decoupling sequence for liquid crystals and solids. *J Magn Reson* **2000,** *142* (1), 97-101.
12. Hou, G.; Yan, S.; Trebosc, J.; Amoureux, J. P.; Polenova, T., Broadband homonuclear correlation spectroscopy driven by combined R2(n)(v) sequences under fast magic angle spinning for NMR structural analysis of organic and biological solids. *J Magn Reson* **2013,** *232,* 18-30.
13. McCoy, K. M.; Rogawski, R.; Stovicek, O.; McDermott, A. E., Stability of nitroxide biradical TOTAPOL in biological samples. *J Magn Reson* **2019,** *303,* 115-120.
14. Viennet, T.; Viegas, A.; Kuepper, A.; Arens, S.; Gelev, V.; Petrov, O.; Grossmann, T. N.; Heise, H.; Etzkorn, M., Selective Protein Hyperpolarization in Cell Lysates Using Targeted Dynamic Nuclear Polarization. *Angew Chem Int Ed Engl* **2016,** *55* (36), 10746-50.
15. Pinon, A. C.; Schlagnitweit, J.; Berruyer, P.; Rossini, A. J.; Lelli, M.; Socie, E.; Tang, M. X.; Pham, T.; Lesage, A.; Schantz, S.; Emsley, L., Measuring Nano- to Microstructures from Relayed Dynamic Nuclear Polarization NMR. *J Phys Chem C* **2017,** *121* (29), 15993-16005.

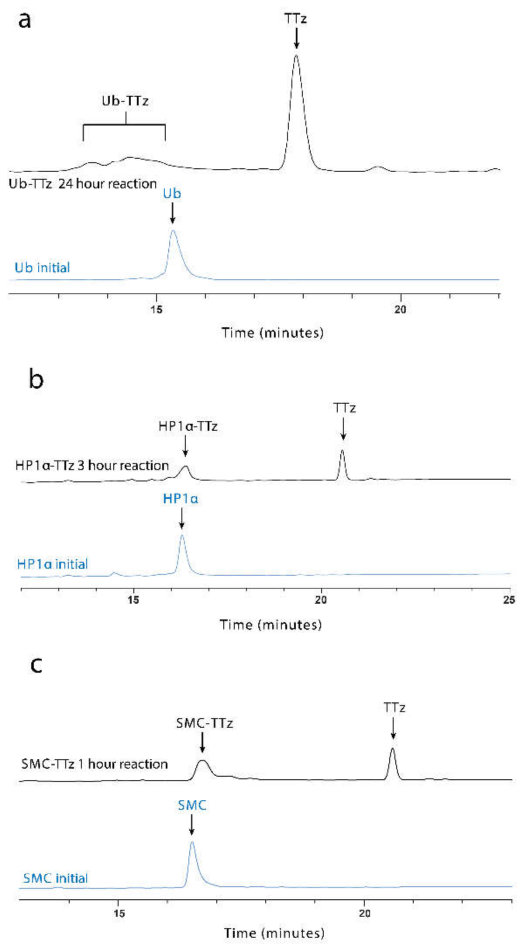
Supplementary figures



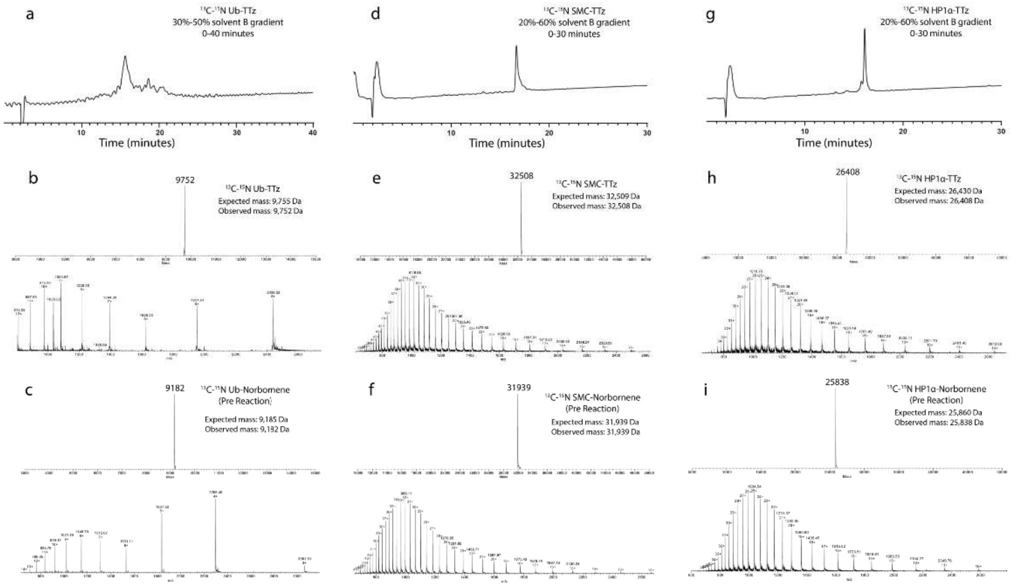
SI Figure 1: Analysis of the TOTAPOL-tetrazine (TTz) polarization agent. (a) RP-HPLC trace at 214 nm absorbance of purified TTz. (b) A 9 GHz solution EPR spectrum of TOTAPOL, and c) TTz. Spectra were acquired in DMSO-d_6. TTz displays the characteristic nitroxide spectral features.



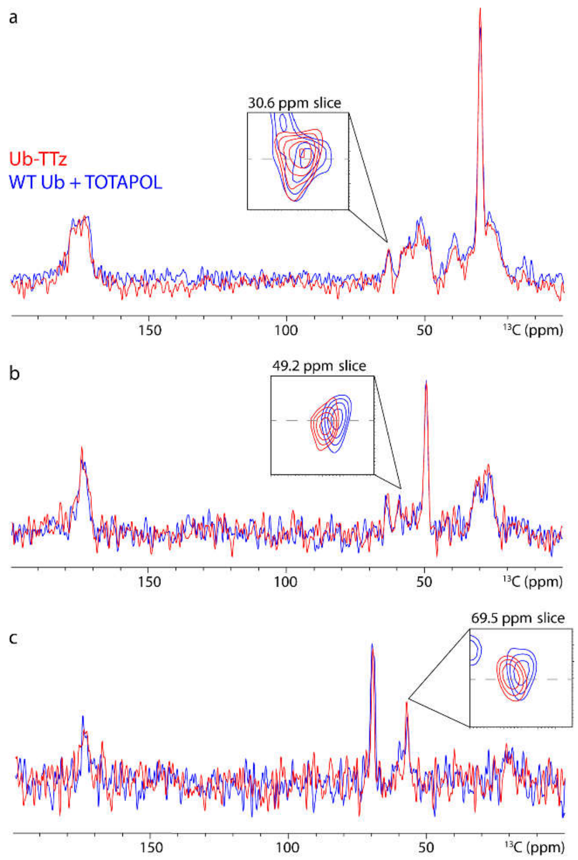
SI Figure 2: ^{1}H-^{15}N HSQC spectra of wild-type ubiquitin (black) and ubiquitin incorporating norbornene-lysine at position 6 (Ub-K6X) (red). Since the UAA is supplied at natural abundance, the Lys6 cross-peak is missing from the spectrum. The UAA causes only minor local perturbations of the ubiquitin fold.



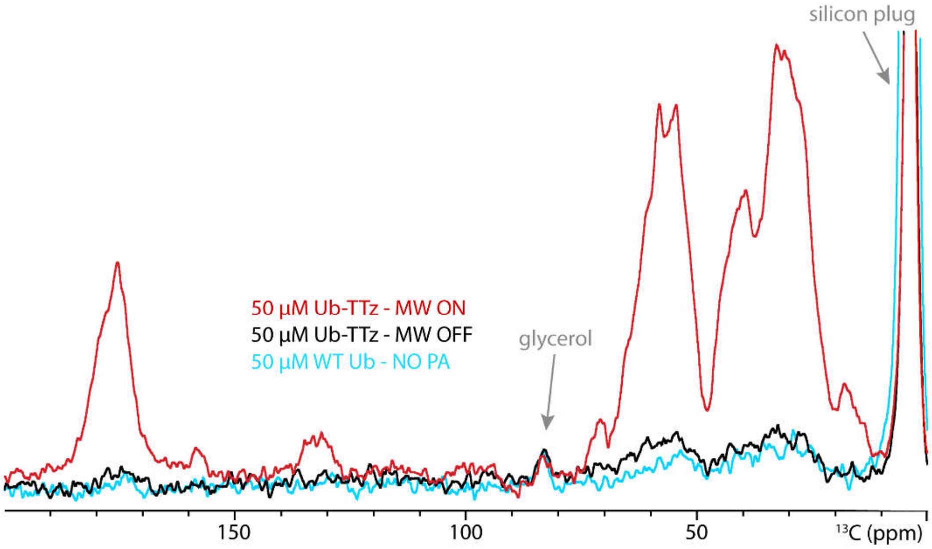
SI Figure 3: Protein-TTz reactions monitored by analytical RP-HPLC for (a) Ub-TTz, (b) HP1α -TTz, and (c) SMC-TTz. The conjugation reaction results in a shift of the retention time and peak broadening due to multiple isomer products with the same mass. Unreacted TTz has strong absorbance at 214 nm and 280 nm, which was useful in monitoring TTz removal from the final DNP samples.



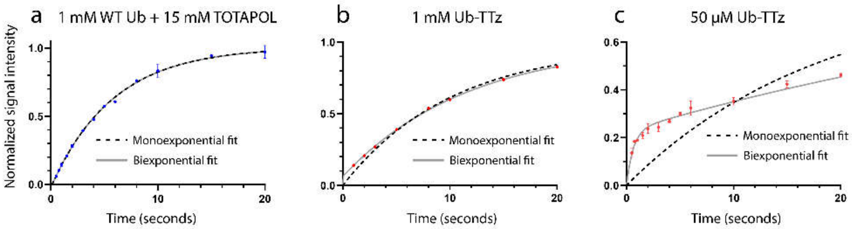
SI Figure 4: Purified ^{13}C,^{15}N-labeled protein-TTz samples analyzed by RP-HPLC and ESI-TOF-MS. (a), (d) and (g) show the RP-HPLC traces for purified Ub-TTz, SMC-TTz and HP1α-TTz respectively. (b), (e), and (h) show the ESI-MS of the pure conjugated proteins, while (c), (f) and (i) show the ESI-MS of the proteins before TTz conjugation.



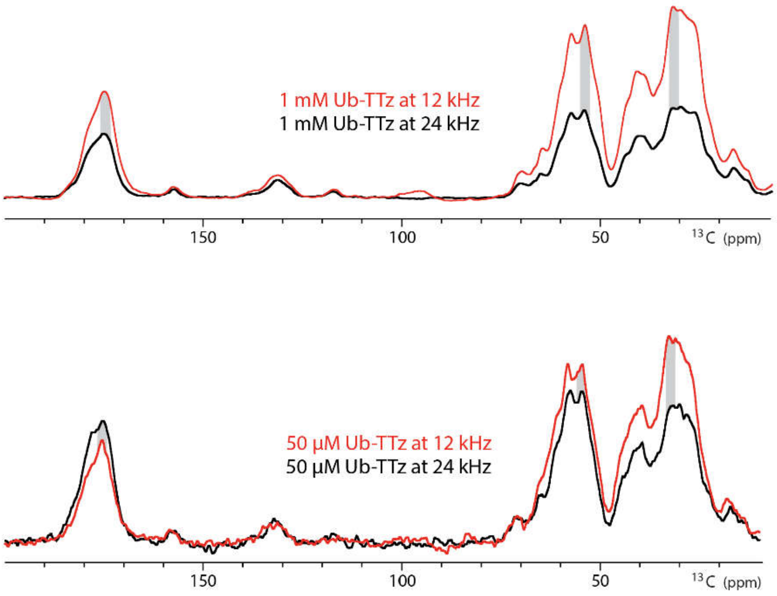
SI Figure 5: Comparison of the linewidths in 2D DARR spectra of 1 mM wild-type ubiquitin prepared with 15 mM dispersed TOTAPOL (blue) and 1 mM Ub-TTz (red). 1D slices were chosen for isolated peaks (inset) at 30.6 ppm (a), 49.2 ppm (b) and 69.5 ppm (c). Despite the relatively low signal-to-noise, the linewidths in the direct dimension appear similar. Thus, we do not see evidence for line-broadening in the Ub-TTz sample. Both spectra were acquired with 100 µg of protein, at 12 kHz MAS frequency, 32 scans, an interscan delay of 3.5 s, 488 indirect dimension points (6.8 ms) for a total acquisition time of 14 hr.



SI Figure 6: Comparison of the signal intensity in 50 μM ubiquitin samples including a wild-type ubiquitin sample without any polarization agent (blue), and a Ub-TTz sample in the presence (red) and absence of microwaves (black). Spectra were acquired at 12 kHz MAS, 4096 scans, interscan delay of 5 s, and a total acquisition time of 5.5 hours. Under these conditions, we do not observe evidence for depolarization or quenching/bleaching due to the attached radical.



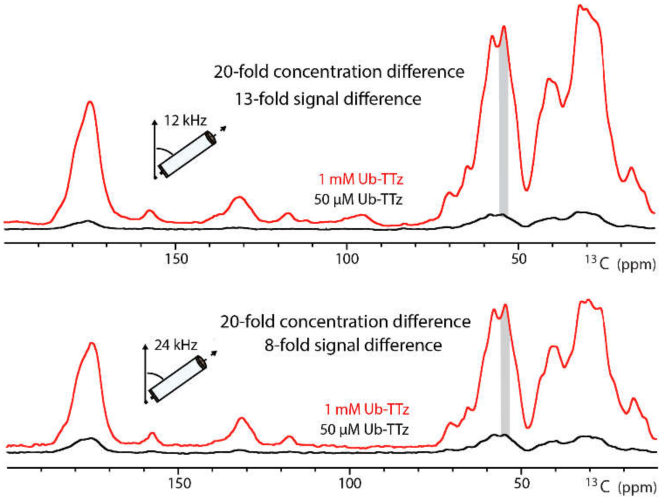
SI Figure 7: DNP build-up curves identical to Figure 4 in the main text with zoomed axes for (a) 1 mM WT Ub + 15 mM TOTAPOL , (b) 1 mM Ub-TTz, and (c) 50 μM Ub-TTz. At this scale, the initial build-up clearly depicts a bimodal increase in the DNP signal for 50 μM Ub-TTz. Error bars correspond to the standard deviation of two independent measurements of the enhancements.



MAS dependence based on MW ON signal intensity

Sample	Ratio of signal intensity 24 kHz : 12 kHz		
	Cβ, Cγ (31 ppm)	Cα (54 ppm)	C′ (175 ppm)
50 μM Ub-TTz	0.69	0.85	1.23
1 mM Ub-TTz	0.47	0.51	0.61
1 mM WT Ub + 15 mM TOTAPOL	0.46	0.56	0.78

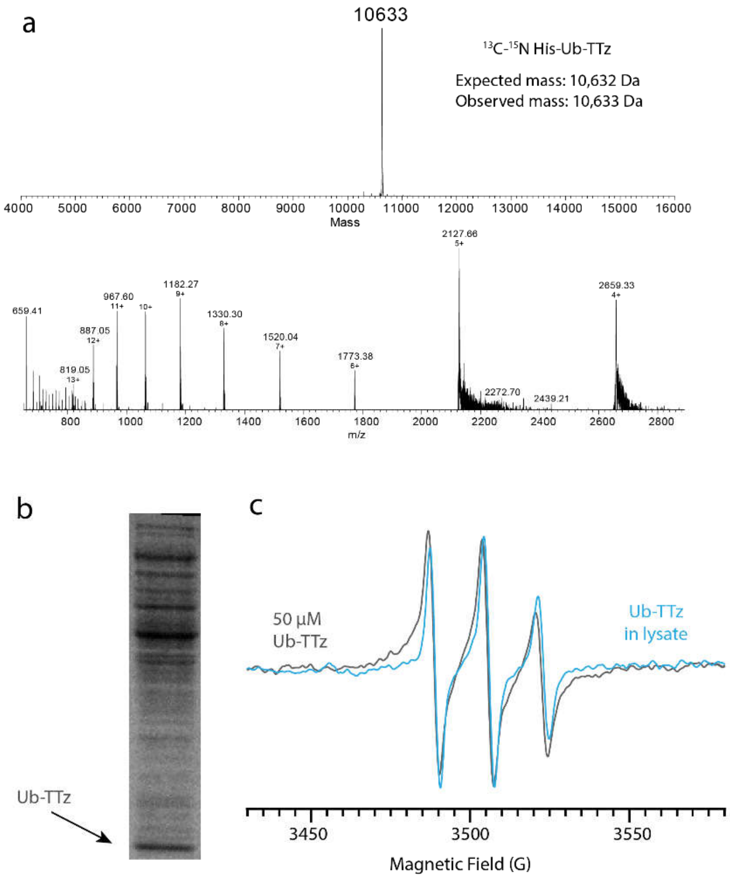
SI Figure 8: Comparison of the DNP-enhanced (microwave on) signal intensities as a function of MAS frequency for 1 mM Ub-TTz (top) and 50 μM Ub-TTz (bottom). The table summarizes the ratio of signal intensity at 24 kHz to 12 kHz MAS for different regions of the spectrum.



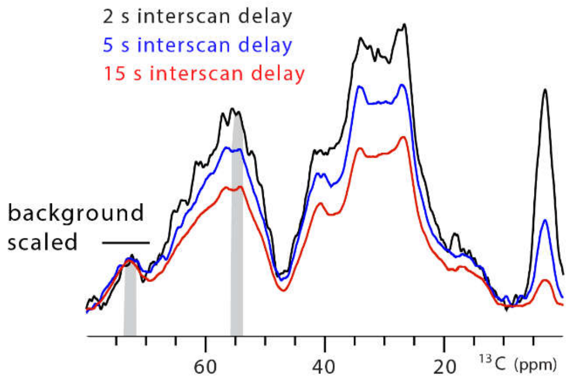
Comparative MAS dependence based on MW ON signal intensity

Samples	Fold difference Ub concentration	Fold difference MW ON @ 12 kHz MAS	Fold difference MW ON @ 24 kHz MAS
1 mM Ub-TTz : 50 μM Ub-TTz	20	13.15	7.90
1 mM WT Ub : 50 μM Ub-TTz	20	13.81	9.02
1 mM WT Ub : 1 mM Ub-TTz	–	1.05	1.1

SI Figure 9: Comparison of absolute enhanced signal intensities for samples containing different concentrations of ubiquitin-TTz or ubiquitin with dispersed TOTAPOL. The purpose of this comparison is to illustrate that smaller microwave on/off enhancements do not necessarily translate to less NMR signal. For example, the 20-fold concentration difference between the 50 μM and 1 mM Ub-TTz samples would imply that a 20-fold difference in NMR signal intensities is to be expected. However, the measured absolute signal difference is much smaller and decreases at higher MAS. The comparison also shows nearly identical signal intensities for the 1 mM Ub-TTz sample and the 1 mM ubiquitin with 15 mM dispersed TOTAPOL sample, despite their very different enhancements. We attribute this effect to radical-induced depolarization in the absence of microwaves which is expected to be more prominent in samples containing higher radical concentrations. Depolarization leads to artificially lower signals in the microwave off spectra and thus skews the reported on/off enhancements.



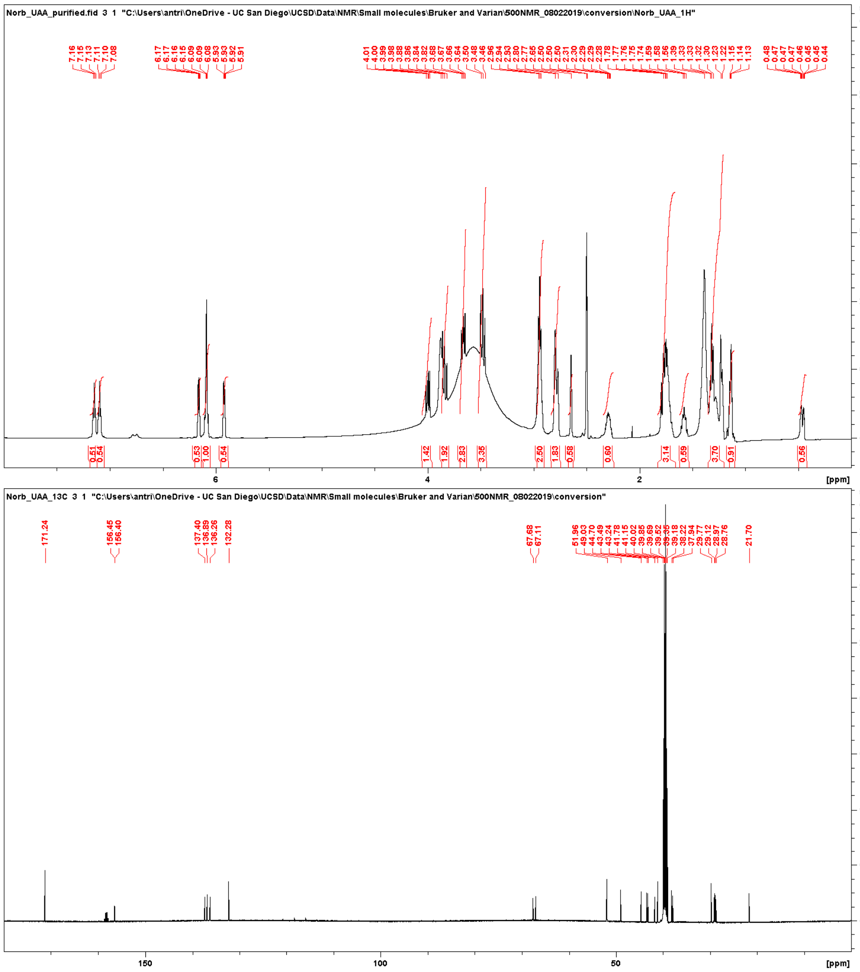
SI Figure 10: Analysis of Ub-TTz in lysate samples. Ub-TTz in lysates contained a His$_6$ tag enabling the affinity column purification of the protein from the lysate and the LC-MS analysis of the conjugation efficiency. (a) ESI-TOF MS of Ub-TTz purified from the lysate. (b) SDS-PAGE analysis of the lysate sample. Analysis of the protein band intensities allowed us to quantify the percentage of ubiquitin present in the lysate (5% of the total protein content and present at 40 μM). (c) 9 GHz EPR of the lysate sample confirmed that Ub-TTz is present at the expected 40 μM concentration.



Ub-TTz selectivity variation by interscan delay

Sample	Ratio of signal intensity Cα : glycerol		
	2 s	5 s	15 s
Ub-TTz in lysate	4.03	3.47	2.59

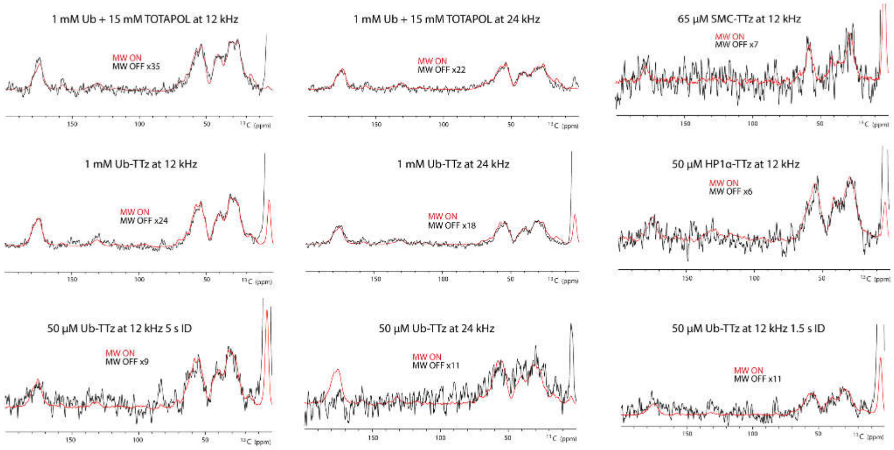
SI Figure 11: The selective DNP enhancement of ubiquitin in lysates can be tuned as a function of the interscan delay. Overlaid spectra were acquired with 256 scans and interscan delays of 2 seconds, 5 seconds and 15 seconds. The table summarizes the ratio of protein signal (Cα) to the background glycerol signals.



SI Figure 12: ^{1}H and ^{13}C NMR spectra of norbornene-lysine.

Enhancements

Sample	Concentration	12 kHz MAS (ε)	24 kHz MAS (ε)
Ub-TTz	50 μM	11 (1.5 s ID) | 9 (5 s ID)	11
	1 mM	24	18
WT Ub + 15 mM TOTAPOL	1 mM	35	22
SMC-TTz	65 μM	7	–
HP1α-TTz	50 μM	6	–



SI Figure 13: Summary of the raw data and experimental parameters used to measure the DNP enhancements reported throughout the manuscript. ID = interscan (recycle) delay.

Acknowledgements

Chapter 5, in full, is a reprint of the material as it appears in Targetable tetrazine-based dynamic nuclear polarization agents for biological systems. 2020. Lim, Byung Joon*, Ackermann, Bryce E.*, Debelouchina, Galia T. ChemBioChem. The dissertation author was the co-primary investigator and co-author of this material.

Chapter VI

A comparative study of nitroxide-based biradicals for dynamic nuclear polarization in cellular environments

A comparative study of nitroxide-based biradicals for dynamic nuclear polarization in cellular environments

Bryce E. Ackermann[1], Byung Joon Lim[1], Sirish Narayanan[1], Galia T. Debelouchina[1*]

[1]*Department of Chemistry and Biochemistry, University of California San Diego, La Jolla, CA 92093, USA*

**Corresponding author: gdebelouchina@ucsd.edu*

Abstract

Dynamic nuclear polarization (DNP) is a powerful tool to enhance the NMR signals of molecules by transferring polarization from unpaired electron spins to nuclei through microwave irradiation. The resulting signal enhancements can enable the analysis of samples that have previously been intractable by NMR spectroscopy, including proteins, nucleic acids, and metabolites in cells. To carry out DNP, the sample is doped with a polarization agent, a small molecule containing two stable nitroxide radicals. DNP applications in cells, however, present significant challenges as nitroxides are often susceptible to the reducing cellular environment. Here, we introduce a novel polarization agent, POPAPOL, that exhibits increased lifetimes under reducing conditions. We also compare its bioresistance and DNP performance with three popular, commercially available polarization agents. Our work indicates that pyrrolidine-based nitroxides can outperform piperidine-based nitroxides in cellular environments, and that future polarization agent designs must carefully balance DNP performance and stability for cellular applications.

Keywords: dynamic nuclear polarization, in-cell NMR, nitroxide radicals, TOTAPOL, AMUPol, AsymPolPok

The living cell maintains an internal environment that dictates the structure, dynamics, and interactions of its macromolecules. Understanding these environmental effects has presented a tremendous challenge to structural biologists who have typically studied biological macromolecules in isolation and in purified form. NMR spectroscopy, which works under physiological conditions and is nonperturbative to living systems, can be a powerful tool to capture these environmental effects in cells. Solution NMR spectroscopy has already been used to describe the structure of small proteins in the complex cellular milieu while solid-state magic angle spinning (MAS) NMR has been able to characterize the properties of membrane proteins in intact cells or cell envelopes.[1] However, the low sensitivity of NMR experiments limits the types of interactions and processes that can be studied in the cellular environment, and requires large amounts of the protein of interest, which is typically overexpressed or electroporated into the cell.[1b, 1d] One possible solution to this problem is dynamic nuclear polarization (DNP), a methodology that takes advantage of the inherently higher polarization of electron spins in magnetic fields to enhance the sensitivity of NMR experiments.[2]

MAS DNP NMR is typically performed by doping a highly deuterated sample with millimolar quantities of a polarization agent (PA) that contains two unpaired electron spins. The most common PAs for biological systems, TOTAPOL, AMUPol and the recently introduced AsymPolPok, contain stable nitroxide radicals and work through a polarization transfer mechanism called the cross effect (CE).[3] The sample is also cryoprotected with a glassing agent, cooled to cryogenic temperatures, rotated at the magic angle (54.7°), and irradiated with microwaves to transfer the large spin polarization of the unpaired electrons to nearby nuclei. Multidimensional solid-state NMR experiments are performed concurrently with microwave irradiation to collect structural data with enhanced sensitivity. Typical enhancements are in the 10 – 100 range, depending on the sample properties, the choice of polarization agent, the experimental conditions, and the strength of the external magnetic field.

MAS DNP NMR has enabled the structural characterization of numerous complex biological systems that typically suffer from low sensitivity. This has included amyloid fibrils, membrane proteins, bacterial secretion proteins, chromatin polymers, nucleic acids, cell envelopes, and plant cell walls.[4] Several recent studies have also demonstrated that DNP can be applied to bacterial and mammalian cells,[5] however, detailed structural studies of proteins inside cells have been limited by the unfavorable properties of nitroxide radicals which are highly susceptible to the reducing cellular environment.[6] This problem is particularly severe for bacterial cells, less so for mammalian cells where the intracellular concentration of radicals can also be boosted by electroporation.[5b, 6] Nevertheless, the low bioresistance of nitroxide PAs necessitates rapid sample preparation, limits the types of cells and time-dependent biological processes that can be studied, and hampers the exploration of PA-targeting DNP approaches that may be required to reduce the unwanted background NMR signal of cellular components.[6a, 7] PA lifetimes can be extended by adding oxidizing agents to regenerate radicals, but this has had limited success in actualizing better signal enhancements.[6] Alternatively, the stability problem can be partially overcome by saturating the cells with PAs (15-30 mM).[5a, 5b] This strategy, however, can generate a large pool of partially reduced PAs and deteriorate DNP performance through paramagnetic relaxation effects. Here, we compare the bioresistance of popular PAs such as TOTAPOL, AMUPol and AsymPolPok with a focus on their stability in mammalian cells.[3] We also design and synthesize a novel PA, 1-(PROXYL-3-oxy)-3-(PROXYL-3-amino)propan-2-ol (POPAPOL), which exhibits reasonable DNP enhancements and improved lifetime under reducing conditions. This work allows us to establish common principles and guidelines that can be used to design efficient bioresistant PAs for cellular applications.

Previous literature has indicated that the stability of nitroxide radicals under reducing conditions can be improved if the six-membered piperidine nitroxide structure (i.e. TEMPO) is substituted with a five-membered pyrrolidine ring (i.e. PROXYL).[8] This substitution alone increases the lifetime of the nitroxide radical in ascorbic acid by 60-fold.[8a] Inspired by this work, we designed and synthesized POPAPOL, a PROXYL version of the first successful PA TOTAPOL[3a] **(Scheme S1 and Fig. S1)**. While TOTAPOL is no longer the best polarization agent for DNP, its facile synthesis could easily be adapted to POPAPOL and allowed us to directly assess the effect of the pyrrolidine nitroxides on the bioresistance and performance of the PA. Throughout this study, we documented the performance of POPAPOL alongside commercial PAs that bear either dual six-membered nitroxides (TOTAPOL, AMUPol) or an asymmetric split of one 6- and one 5-membered nitroxide (AsymPolPOK) **(Fig. 1)**. We first assessed the

DNP performance of our PA pool in an optimal *in vitro* environment (60% d_8-glycerol, 30% D_2O, 10% H_2O) with 10-15 mM PA. We recorded signal enhancements ($\varepsilon_{on}/_{off}$) of $^{13}C,^{15}N$-proline at 14.1 T and 12 kHz MAS **(Fig. 1A).** Our results in these conditions match previously published comparisons between TOTAPOL, AMUPol and AsymPolPOK.[3b, 3c] Importantly, the enhancements of 32 for POPAPOL and 36 for TOTAPOL are comparable. While these two biradicals have different dipolar coupling and exchange interactions due to the slightly different distance between the two unpaired electron spins, the flexible propanol linker likely dampens the effects on the DNP enhancement under these conditions.[3c, 9]

While most *in vitro* DNP studies are performed with high concentrations of PA (10-15 mM), the DNP performance at low concentrations might be more relevant for experiments in cells. For example, a recent report on the delivery of PAs into mammalian cells by electroporation determined that only 1 mM of active PA successfully reached the cell interior.[5b] Here, *in vitro* DNP experiments with PAs at concentrations of 1 mM yielded reduced enhancements and some notable differences **(Fig. 1B).** For example, the enhancements of 54 for AMUPol and 46 for AsymPolPOK are much closer than in the high concentration condition **(Fig. 1A).** The improved efficacy of AsymPolPOK is likely due to its larger electron dipolar and J-exchange interactions which lead to a shorter polarization buildup time,[3c] a property that would be particularly advantageous for DNP at low concentrations. Similarly, the reversal between POPAPOL (ε=16) and TOTAPOL (ε=10) may be explained by the shorter electron-electron distance of POPAPOL. The enhancements we have reported so far do not take into account depolarization, a phenomenon that artificially reduces the intensity of nuclear signals in the presence of a PA.[10] To address the depolarization effect, we compared the enhanced spectra of all PAs with the same spectrum of 100 mM ^{13}C,^{15}N-proline collected without microwave irradiation **(Fig. 1C)**. In this case, AsymPolPOK presents the best true signal enhancement, while POPAPOL performs closer to AMUPol than the microwave on/off enhancements suggest. These results are consistent with previous observations that AMUPol is particularly susceptible to depolarization effects.[3c, 10]

Our *in vitro* DNP experiments demonstrate that POPAPOL performs equally as well or better than TOTAPOL depending on the conditions, and that AsymPolPOK gives the highest signal-to-noise ratios. Next, we assessed PA reduction resistance in ascorbic acid, an intracellular metabolite often used as a reducing agent to evaluate radical stability.[8a] PA radical lifetime was measured by EPR spectroscopy at 5-minute intervals upon the addition of 1 mM or 2 mM ascorbic acid to 1 mM PA **(Fig. 1D).** Since each PA has two unpaired electron spins, 2 mM ascorbic acid represents a stoichiometric condition where two of the PAs were completely reduced within the first time point (AMUPol and TOTAPOL). We therefore used 1 mM ascorbic acid to better compare the differences between the PAs at the early time points. As expected, POPAPOL had the slowest rate of reduction with a decay constant (τ) of 8.7 minutes compared to the 3.2 minutes of TOTAPOL in the 1 mM ascorbic acid condition **(Table S1).** AsymPolPOK, which has one six- and one five-membered nitroxide, showed an intermediate decay rate.

In order to assess the stability of the PAs in more relevant reducing environments, we next performed experiments in HEK293 cell lysates.[2] While lysates are expected to have a lower reduction capacity, they offer a similar composition to the cellular interior and capture the complexity of different reducing agents and biomolecular interactions.[4b, 11] Our lysate reduction assay was performed by incubating 1 mM PA in lysates generated by 3 freeze-thaw cycles **(Fig. 2A).** Remarkably, the reduction decay behavior in lysates, which contain a mix of different reducing agents, deviates from the trends observed in ascorbic acid. In particular, POPAPOL and AsymPolPOK exhibit similar reduction rates, while AMUPol performs better than TOTAPOL. The improved stability of AMUPol and AsymPolPOK suggests that the spirocyclohexyl rings neighboring the nitroxide provide more protection in cellular lysate than their methyl counterparts.[12] Visualization of the EPR spectra during the experimental timecourse, however, also shows progression towards monoradicals, which is more dramatic in AMUPol and AsymPolPOK **(Fig. 2B).** Since the presence of two dipolar coupled unpaired electron spins is essential for DNP via the CE, we estimated the population of biradical, monoradical, and fully reduced PAs for each symmetric PA **(Fig. S2)**.[13] Our analysis shows that at 30 minutes, 62%, 45%, and 21% of the PA molecules are in a biradical state for POPAPOL, AMUPol, and TOTAPOL, respectively. These values are likely overestimates since some radical loss already occurs in the first 10 min when the sample and the instrument are prepared for measurement.[3a] However, they still highlight the rapid loss of active PA during sample preparation (**Fig. S3**). The timescale of this process is important since previous reports on intact HEK293 cell preparation for DNP required 30 minutes to complete electroporation, washing, and slow-freezing, indicating that a significant loss of DNP-competent PA occurs during sample handing.[5b]

To understand how total EPR signal loss and conversion of the biradical to the monoradical form affect enhancements, we performed DNP experiments in lysates as a function of incubation time (**Fig. 2C,D**). To this end, 1.25 mM PA in DNP juice (20% d_8-glycerol, 70% D_2O, 10% H_2O, 1x PBS) was added to HEK293 cells, followed by centrifugation into the sample rotor. The samples were subjected to three freeze-thaw cycles inside the DNP probe to lyse the cells and achieve a more accurate zero time-point (**Fig. S4**). We recorded a time course by removing the sample from the probe for 5, 15, and 45 minutes and keeping it at room temperature in between time points (**Fig. 2C**). Measurements of the absolute DNP enhancement and signal as a function of time indicate that AsymPolPOK clearly outperforms all other PAs in cell lysates. Even after 45 min of incubation, AsymPolPOK exhibits an enhancement of 32, which is higher than the initial enhancement of any other PA. Interestingly, however, normalized data demonstrate that the DNP enhancements of AsymPolPOK and POPOPOL decrease at a similar rate (**Fig. S5**). Therefore, the outstanding performance of AsymPolPOK is likely due to its inherently higher polarization ability at low concentrations, which gives it an advantage at the zero time point. The poor performance of AMUPol in this experiment is also noteworthy. This is likely due to the low concentration of PA used in the study (1.25 mM initial concentration) which is far from optimal for AMUPol, as well as the rapid reduction of the six-membered ring nitroxides in the first few minutes of the time course.

In summary, our experiments lead to the following conclusions, summarized in **Table 1**: 1) Among the four nitroxide-based PAs that we tested, the best performer for DNP-enhanced NMR spectroscopy in cell environments appears to be AsymPolPOK.

This PA has the best balance of features for bioresistance and enhancement, including one five-membered ring nitroxide and strong dipolar coupling and exchange interaction between the unpaired electron spins that allows optimal polarization transfer, even at low concentrations. 2) The use of five-membered ring nitroxide building blocks can significantly improve the bioresistance of PAs. While pyrrolidine-based nitroxides may not necessarily display the best DNP performance *in vitro*, their longer lifetimes in cells may still allow them to outperform popular *in vitro* spirocyclohexyl-bearing PAs such as AMUPol. Therefore, as DNP experiments in cells become more popular, future PA design should focus not only on optimizing the enhancement and solubility of the radical, but also take into consideration its stability in the cellular milieu. In addition to five-membered ring nitroxides, there are other design principles that can be employed to achieve this goal. For example, EPR experiments have shown that substitution of the methyl groups on the nitroxide with diethyl moieties can further improve bioresistance in biological settings.[8a, 8b, 14] Alternatively, PAs can also be constructed with closed conformation spirocyclohexyl rings.[12a] Trityl-radical PAs may also be an avenue toward efficient DNP at high magnetic fields,[6b] however, more work is required to understand their capabilities in biological settings.[15] Additionally, triradical PAs can be used to favor a population of polarization agents with at least two active radicals.[16]

Finally, it is also important to consider the ability of the PAs to cross the cellular membrane. Published literature has shown that AMUPol is not cell permeable and that TOTAPOL has a high tendency to interact with bacterial membranes.[5b, 6b] Using an HPLC-based assay, we also confirmed that POPAPOL and AsymPolPOK are not efficient at entering mammalian cells (**Fig. S6**). Therefore, in addition to enhancement and stability, future designs for in cell applications must also take into account cell permeability and solubility of the PAs. As cellular environments place considerable demands on the desired properties of DNP PAs, we hope that this study will inform the design and use of PAs as they become tuned for functional purposes beyond pure signal enhancements.

Acknowledgements

We thank Prof. Snorri Sigurdsson and Dr. Thomas Halbritter for generously providing AsymPolPOK, Prof. Cornelis Murre's lab for help with cell culture experiments, and Dr. Xuemei Huang for assistance with EPR measurements. This work was funded by NIH R35 GM138382 and NSF MRI CHE 2019066 grants to G.T.D., and NIH T32 GM008326 fellowship to B.E.A.

Conflict of Interest

The authors declare no conflict of interest.

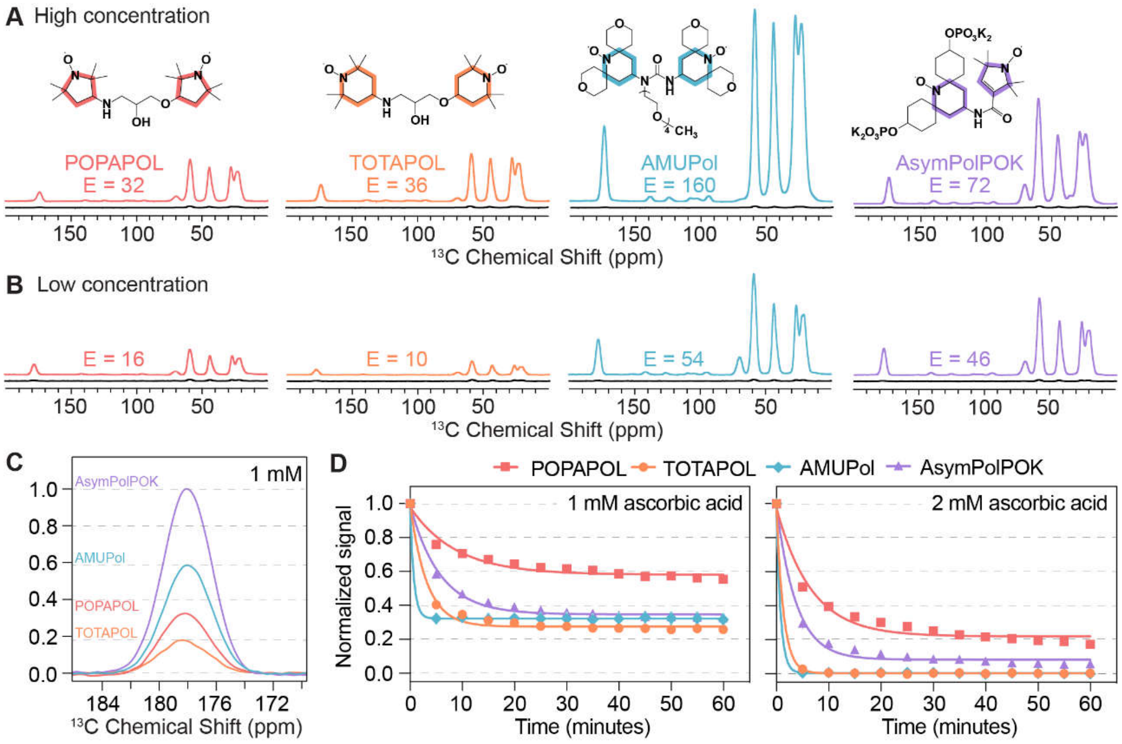


Figure 1. Polarization agent performance *in vitro*. (**A**) DNP enhancements of 0.25M ^{13}C,^{15}N-labeled proline with 10 or 15 mM PA. (**B**) DNP enhancements of 0.1M ^{13}C,^{15}N-labeled proline with 1 mM PA. (**C**) Absolute enhancement with 1 mM PA shown by scaling the carbonyl region to the same microwave off spectrum. Relative intensities are shown on the y-axis, normalized to the highest peak. (**D**) PA (1 mM) reduction time course in 1 or 2 mM ascorbic acid, recorded by EPR. E – enhancement. Standard error bars are shown, n = 3.

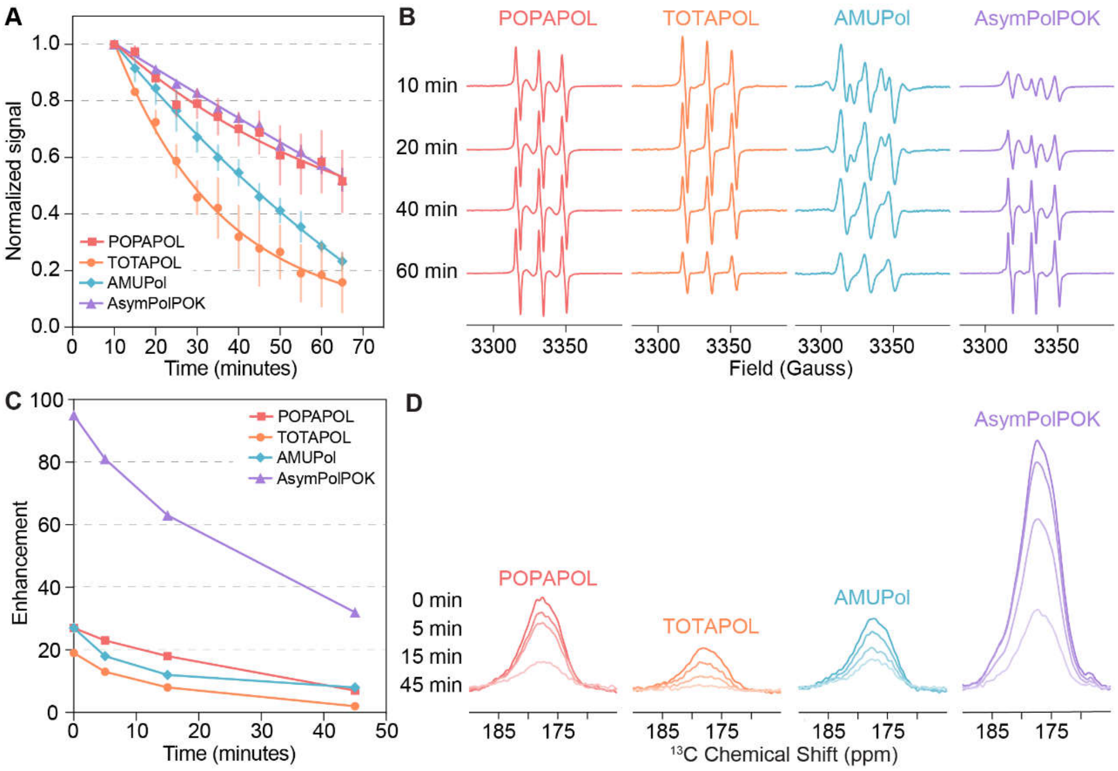
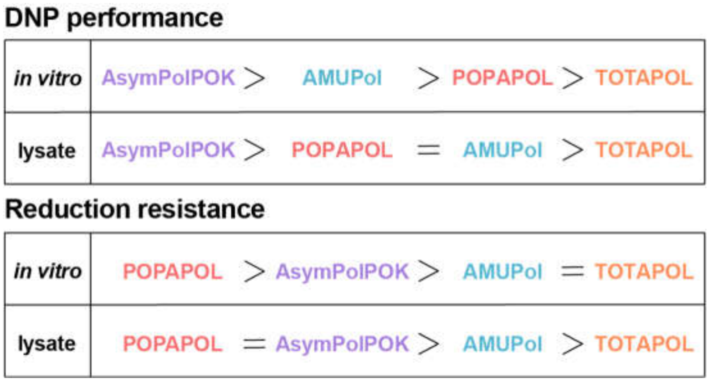


Figure 2. Polarization agent performance in HEK293T cell lysates. (**A**) 1 mM PA was incubated in lysates generated by three freeze-thaw cycles. First time point was recorded 10 minutes after final freeze-thaw. (**B**) EPR spectra from the lysate time course. (**C**) DNP enhancements of lysates with 1.25 mM PA. The enhancement was measured on the aliphatic signals at different time points after incubation at 298 K. (**D**) Raw carbonyl signals of the DNP-enhanced cell lysate samples at different time points. Samples contain the same number of cells. Standard error bars are shown, n = 3.

Table 1. Summary of PA performance in vitro and in cell lysates.



DNP performance

in vitro	AsymPolPOK > AMUPol > POPAPOL > TOTAPOL
lysate	AsymPolPOK > POPAPOL = AMUPol > TOTAPOL

Reduction resistance

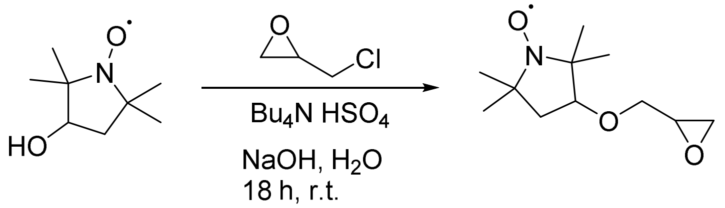
in vitro	POPAPOL > AsymPolPOK > AMUPol = TOTAPOL
lysate	POPAPOL = AsymPolPOK > AMUPol > TOTAPOL

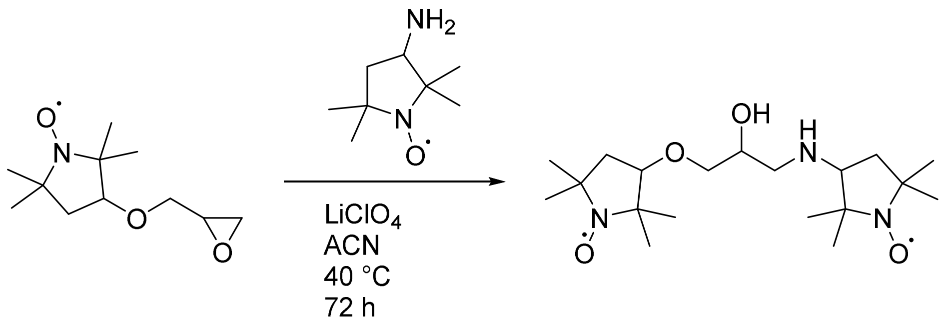
Materials and Methods

Materials and equipment

All reagents and solvents were purchased from Sigma Aldrich (St. Louis, MO), Thermo-Fisher Scientific (including Fisher Scientific and Acros Organics, Waltham, MA) and used without further purification unless otherwise noted. AsymPolPOK was received as a gift from the Sigurdsson Lab, AMUPol was purchased from CortecNet (Brooklyn, NY), and TOTAPOL was synthesized as described in Ref. [3a]. Isotopically enriched reagents and solvents for NMR were purchased from Cambridge Isotope Laboratories (Tewksbury, MA). TLC analyses were carried out using aluminum-backed, precoated silica gel plates (Merck TLC silica gel 60 F254) from EMD Millipore (Billerica, MA). Column chromatography was performed on Acros Organics Silica gel (0.035-0.070 mm, 60 Å). Reverse-phase (RP) HPLC was performed on a 2545 Binary Gradient Module Waters system equipped with a 2484 UV/vis detector from Waters Corporation (Milford, MA). Analytical measurements used a Waters Symmetry300 C18 4.6 mm x 150 mm, 5 µm particle size column. HPLC solvent A contained 100% H_2O + 0.1% trifluoroacetic acid (TFA), while solvent B contained 100% acetonitrile + 0.1% TFA. HR-MS analysis was conducted on an Agilent 6230 TOFMS with Jet Stream ESI. Quartz Q-.9x1.1 capillary tubes used for EPR spectroscopy were purchased from SP-Wilmad-Glass (Vineland, NJ). 1.9 mm ZrO_2 MAS rotors for DNP were purchased from CortecNet (Brooklyn, NY).

Scheme 1: Synthesis of POPAPOL





POPAPOL was synthesized similar to the previously published TOTAPOL synthesis scheme.[3a] Briefly, tetrabutylammonium hydrogen sulfate (5.61 mg) and epichlorohydrin (161.64 mg) were dissolved in 50% NaOH solution at room temperature. PROXYL-4-OH (65.40 mg) was added to the solution and the reaction mixture was stirred overnight. The organic layer was extracted three times, dried with $NaSO_4$, filtered and evaporated *in vacuo*. Silica column chromatography was performed using hexanes, and a pale orange liquid was obtained as the product (62.0 mg, 70% yield), 4-(2,3-Epoxypropoxy)-2,2,5,5-tetramethyl-1-pyrrolidin-1-oxyl [4-(2,3-Epoxypropoxy)-PROXYL].

[4-(2,3-Epoxypropoxy)-PROXYL] (49.0 mg) and $LiClO_4$ (29.19 mg) were dissolved in acetonitrile and combined with PROXYL-4-NH (35.95 mg). The mixture was stirred for 72 hours at 40 °C, then loaded directly onto silica column for chromatography with dichloromethane, and repeated twice for purity. A yellow liquid was obtained as the product (30.4 mg, 20.7% yield), 1-(PROXYL-3-oxy)-3-(PROXYL-3-amino)propan-2-ol [POPAPOL]. The product was confirmed by HR-MS(ESI), expected mass for $[C_{19}H_{38}N_3O_4]^+$ $[M+H]^+$ is 372.2857 Da, experimentally determined mass is 372.2859 Da. The RP-HPLC trace and EPR spectrum are shown in **Fig. S1**.

MAS DNP NMR setup

Experiments were performed on an Ascend^{TM} Bruker 600 MHz (^{1}H Larmor frequency) spectrometer equipped with an AVANCE NEO Bruker console and coupled to a 395 GHz Bruker gyrotron for high-power microwave irradiation. All spectra were recorded using a triple resonance (^{1}H, ^{13}C, ^{15}N) 1.9 mm MAS Bruker probe (C.MASDVT600W2 BL1.9 X/Y/H DNP). The magic angle was set with KBr and the ^{13}C chemical shift was referenced to the downfield peak of the rotor silicon plug at 3.2 ppm at low temperature. Sample temperature was maintained at 100 K with a 12 kHz spining rate using a 1.9 mm zirconia rotor. Temperature was determined by the T_1 of a KBr standard.[17]

1D ^{1}H-^{13}C cross polarization experiments for enhancement determination had the following parameters: 100 K, 12 kHz MAS rate, 5 s interscan delay, 9 ms acquisition, 900 points, and ^{13}C center frequency set at 100 ppm. For cross polarization, a 500 μs contact time was used with the Hartmann-Hahn matching condition satisfied by setting the ^{13}C power to 66 kHz and optimizing on the ^{1}H channel. 100 kHz spinal64 decoupling was performed during acquisition. Optimal enhancements were obtained at 125 mA (AMUPol/AsymPolPOK) or 105 mA (TOTAPOL/POPAPOL) microwave irradiation. Spectra were processed with an exponential window and 30 Hz line broadening.

MAS DNP NMR *in vitro*

Samples for high concentration enhancement measurements were prepared with 250 mM ^{13}C, ^{15}N-Proline in DNP juice (60% d8-glycerol, 30% D_2O, 10% H_2O), and either 15 mM (TOTAPOL/POPAPOL) or 10 mM (AMUPol/AsymPolPOK) PA. The concentration was optimized to give the best enhancement for the particular PA. Samples for low concentration enhancement measurements were prepared with 100 mM ^{13}C, ^{15}N-Proline in DNP juice (60% d8-glycerol, 30% D_2O, 10% H_2O) and 1 mM PA. Enhancements were determined by scaling of the microwave on and microwave off spectra.

EPR *in vitro*

Samples for reduction rate measurements were prepared with 1 mM PA in either 1 mM or 2 mM ascorbic acid solution. EPR was performed with a 9.35 GHz EMXplus-Xband CW spectrometer with consistent acquisition settings of 2 mW microwave power at 298 K, acquiring spectra every 5 minutes with a starting point at 5 minutes.

Cell culture

HEK293 cells were grown in 60 mm or 15 cm culture dishes in DMEM (Gibco), 10% FBS, and 1% Pen-Strep at 37 °C in 5% CO_2. Cells were processed at 90% confluency by detachment with ATV trypsin (Life Technologies) and growth media. Cell suspensions were centrifuged at 400 rpm to pellet. Media were removed and the cell pellet was used to passage cells or for experiments. Cells were counted manually using a phase contrast hemacytometer (Hausser Scientific).

HPLC cell entry experiments

HEK293 cell entry assays were performed with 6 million cells. Cells were removed from confluent 15 cm cell dishes with ATV solution and pelleted at 400 rpm.

6 million cells were washed once with 1 mL PBS, pelleted, then resuspended in PBS to 40 µl with 15 mM PA. The starting quantity of PA was determined by pelleting the cells directly after mixing and removing 10 µl for HPLC. The cell slurry was then placed on a rotator for 60 minutes at 22 °C. Again, 10 µl was removed for the final timepoint and analyzed by RP-HPLC with a 30 minute 0-70% acetonitrile gradient. The A280 chromatogram peaks were integrated to determine the difference in supernatant concentration of PA at 0 and 60 minutes.

EPR in HEK293 cell lysates

Reduction rate assays in HEK293 cell lysates were performed with 3 million cells. Cells were removed from confluent 60 mm cell dishes with ATV solution and pelleted at 400 rpm. Cells were resuspended in PBS and spun down again into a 20 µl pellet. PA was added to the pellet to reach 1 mM (final density at 133,000 cells/µl). The cell pellet was subjected to three freeze-thaw cycles with 1-minute intervals. The resulting lysate was pipetted into a capillary tube for EPR. The time course was started 10 minutes after completing the freeze-thaws due to necessary time to prepare the sample and calibrate the instrument. Time points were recorded every five minutes. EPR spectrometer settings were the same as for the *in vitro* experiments.

MAS DNP NMR of HEK293 cell lysates

DNP reduction assays in HEK293 cell lysates were performed with 3 million cells. Cells were removed from confluent 60 mm cell dishes with ATV solution and pelleted at 400 rpm. Cells were resuspended in PBS and spun down again into a 20 µl pellet. The pellet was mixed with 20 µl solution of PBS-based DNP juice (20% d8-glycerol, 70% D_2O, 10% H_2O, 1xPBS) with 2.5 mM PA (final PA concentration 1.25

mM, final cell density ~300,000/μl). The mixture was spun down at 400xg directly into the rotor. The freeze-thaw cycle, which was done inside the DNP stator, began after 15 minutes after the PA was added to the lysate. Enhancements were determined by scaling of the microwave on and microwave off spectra for the aliphatic peaks upfield of the glycerol peaks.

Calculation of biradical fractions

The fraction of PA that was in a biradical, monoradical or completely reduced state, was estimated using the follow equation set,[13] where I(t) is the normalized double integral of the EPR at a given time point t. The three population fractions are normalized to unity at t = 0, in order to always sum up to one.

$$Biradical\ (t) = I(t)^2$$

$$Monoradical\ (t) = 2I(t) * (1 - I(t))$$

$$No\ radical\ (t) = (1 - I(t))^2$$

The equation set can be understood as follows: I(t) is the normalized data graphed in **Fig. 2A** where it represents the signal of all active nitroxides in the sample. The probability of finding two active nitroxides on the same PA is then $I(t)^2$. In that case, 1 - I(t) represents the fractions of radicals that are reduced to hydroxylamine, and the probability of having two reduced nitroxides on the same PA is $(1 - I(t))^2$. And finally, the probability of having one reduced and one active nitroxide in a PA is equal to 2I(t) * (1 - I(t)). These equations assume that both nitroxides on the same PA are reduced with equivalent kinetics and can therefore be applied only to symmetric PAs such as TOTAPOL, AMUPol and POPAPOL.

Radical reduction rate analysis

The reduction rates in ascorbic acid as presented in **Table S1** were obtained as follows: The double integral of the EPR spectra was determined using the Bruker EPR software Xenon or EasySpin.[18] Using GraphPad Prism version 8.2, time course data were fit to a single-phase decay model, where P is the plateau and τ is the decay constant.

$$I = e^{-\frac{t}{\tau}} * (1 - P) + P$$

Spectral analysis

NMR spectra were processed and analyzed using Bruker TopSpin 4.0.5. EPR spectra were processed and integrated using Xenon or EasySpin.[18]

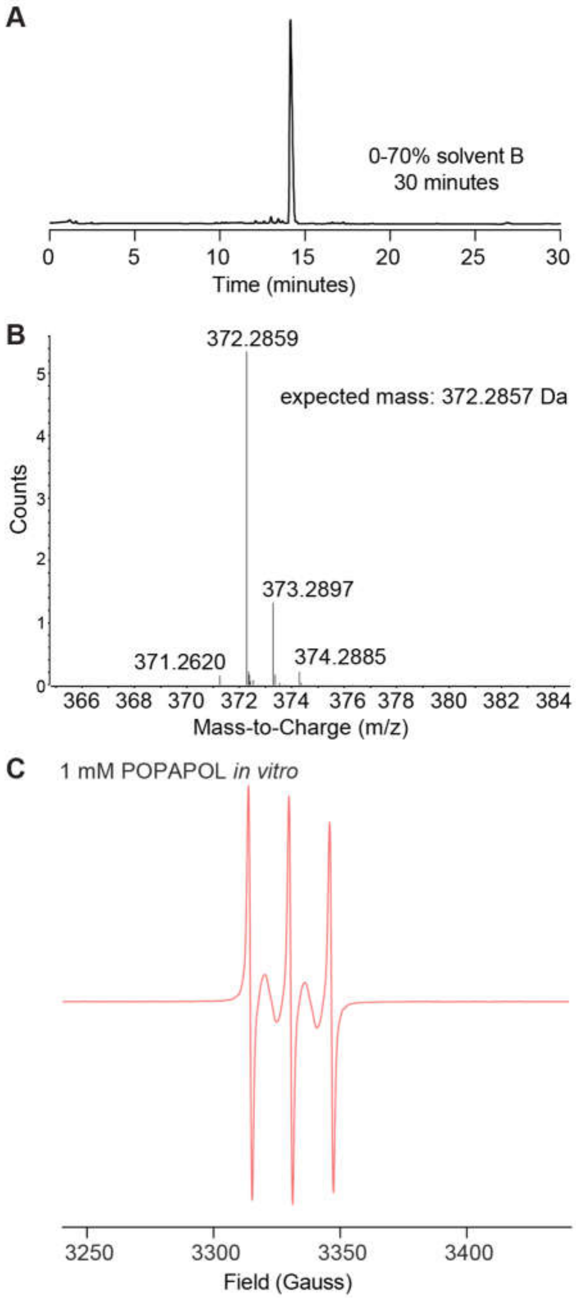


Figure S1. Characterization of POPAPOL. A) HPLC A280 trace of POPAPOL with 0-70% acetonitrile/TFA gradient over 30 minutes. B) HR-ESI-TOF mass spectrometry of POPAPOL. Expected mass for $[C_{19}H_{38}N_3O_4]^+$ $[M+H]^+$ is 372.2857 Da, experimentally determined mass is 372.2859 Da. C) EPR spectrum of 1 mM POPAPOL dissolved in water.

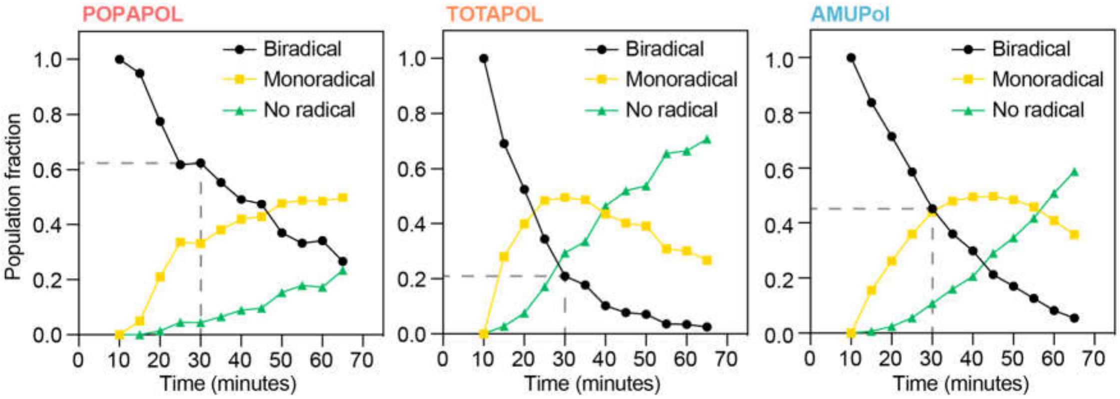


Figure S2. Calculation of the population of biradicals, monoradicals, and fully reduced PAs in HEK293T cell lysates. This analysis is only performed on symmetric biradicals due to the assumption that each nitroxide isreduced with equivalent kinetics. The gray dotted lines highlight the estimated fraction of biradicals remaining at 30 minutes.

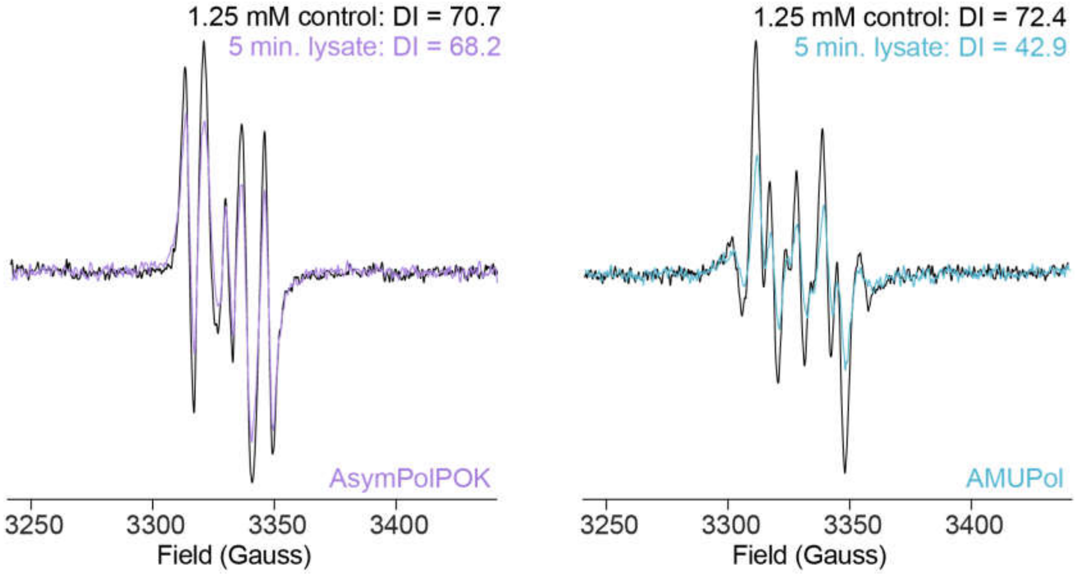


Figure S3. EPR spectra of 1.9 mm DNP zirconia rotors filled with 1.25 mM PA in water (black) or 1.25 mM PA in cell lysates, demonstrating the amount of signal that is lost in the five minutes necessary to prepare the sample and start the EPR measurement.

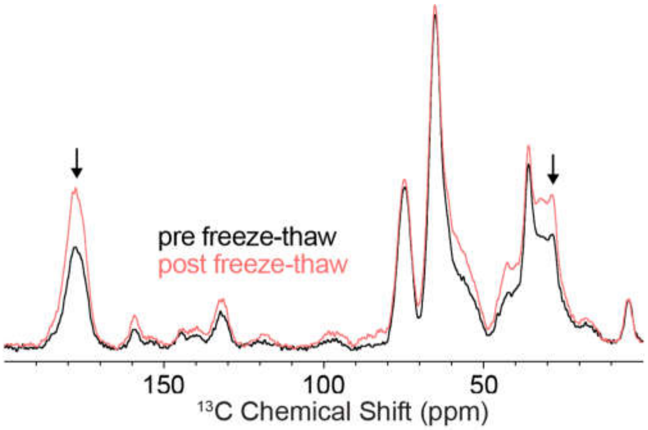


Figure S4. DNP-enhanced MAS NMR spectra of cell lysates before and after one freeze-thaw cycle. Samples contain 1 mM POPAPOL. The freeze-thaw procedure results in better dispersion of PA and cellular components, leading to higher enhancement compared to intact cells.

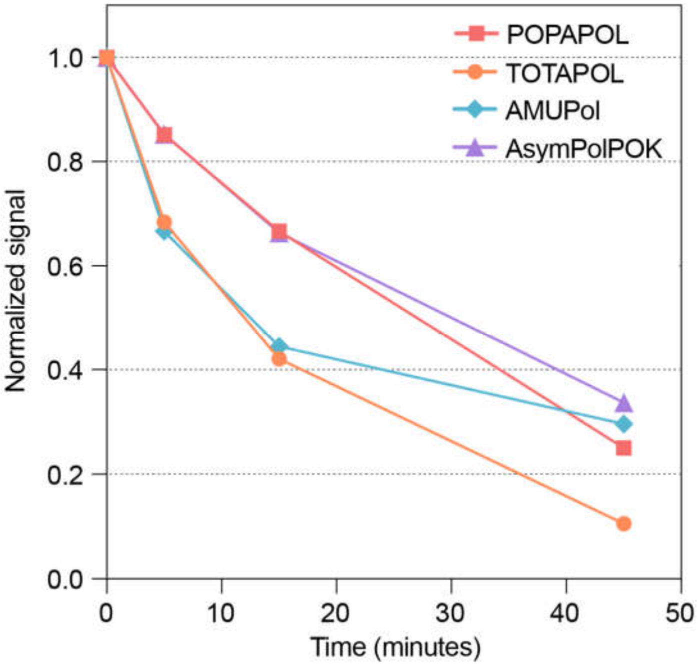


Figure S5. DNP enhancements of cell lysates as a function of reduction time, normalized to the maximum enhancement at the initial time point.

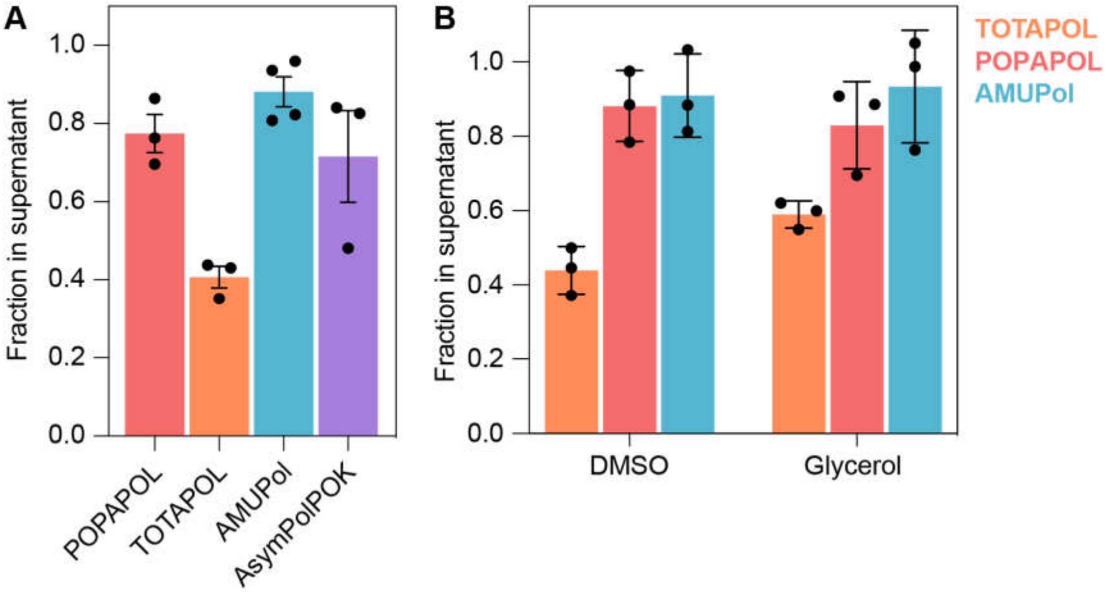
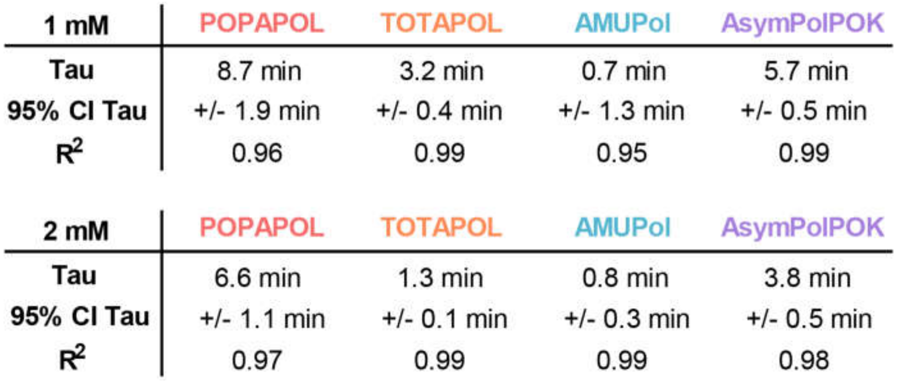


Figure S6. HPLC assay to determine the propensity of PAs to interact with and/or enter cells. A 6 million HEK293T cells were incubated with 15 mM PA, the cells were spun down gently and the amount of PA left in the supernatant was determined by analytical RP-HPLC. A) Fraction of PA remaining in the supernatant after incubation in PBS. B) Fraction of PA remaining in the supernatant after incubation in PBS with 10% DMSO or 15% glycerol.

Table S1. Fitting parameters for the reduction of PAs in ascorbic acid solutions as shown in Fig. 1D. CI – confidence interval.



1 mM	POPAPOL	TOTAPOL	AMUPol	AsymPolPOK
Tau	8.7 min	3.2 min	0.7 min	5.7 min
95% CI Tau	+/- 1.9 min	+/- 0.4 min	+/- 1.3 min	+/- 0.5 min
R^2	0.96	0.99	0.95	0.99

2 mM	POPAPOL	TOTAPOL	AMUPol	AsymPolPOK
Tau	6.6 min	1.3 min	0.8 min	3.8 min
95% CI Tau	+/- 1.1 min	+/- 0.1 min	+/- 0.3 min	+/- 0.5 min
R^2	0.97	0.99	0.99	0.98

References:

[1] aE. Luchinat, L. Banci, *Curr Opin Struct Biol* **2022**, *74*, 102374; bF. X. Theillet, A. Binolfi, B. Bekei, A. Martorana, H. M. Rose, M. Stuiver, S. Verzini, D. Lorenz, M. van Rossum, D. Goldfarb, P. Selenko, *Nature* **2016**, *530*, 45-50; cM. Kaplan, S. Narasimhan, C. de Heus, D. Mance, S. van Doorn, K. Houben, D. Popov-Celeketic, R. Damman, E. A. Katrukha, P. Jain, W. J. C. Geerts, A. J. R. Heck, G. E. Folkers, L. C. Kapitein, S. Lemeer, P. M. P. van Bergen En Henegouwen, M. Baldus, *Cell* **2016**, *167*, 1241-1251 e1211; dJ. E. Kent, L. M. Fujimoto, K. Shin, C. Singh, Y. Yao, S. H. Park, S. J. Opella, G. V. Plano, F. M. Marassi, *Biophys J* **2021**, *120*, 453-462.

[2] D. A. Hall, D. C. Maus, G. J. Gerfen, S. J. Inati, L. R. Becerra, F. W. Dahlquist, R. G. Griffin, *Science* **1997**, *276*, 930-932.

[3] aC. Song, K. N. Hu, C. G. Joo, T. M. Swager, R. G. Griffin, *J Am Chem Soc* **2006**, *128*, 11385-11390; bC. Sauvee, M. Rosay, G. Casano, F. Aussenac, R. T. Weber, O. Ouari, P. Tordo, *Angew Chem Int Ed Engl* **2013**, *52*, 10858-10861; cF. Mentink-Vigier, I. Marin-Montesinos, A. P. Jagtap, T. Halbritter, J. van Tol, S. Hediger, D. Lee, S. T. Sigurdsson, G. De Paepe, *J Am Chem Soc* **2018**, *140*, 11013-11019.

[4] aT. Biedenbander, V. Aladin, S. Saeidpour, B. Corzilius, *Chem Rev* **2022**, *122*, 9738-9794; bN. Elathram, B. E. Ackermann, G. T. Debelouchina, *J Magn Reson Open* **2022**, *10-11*.

[5] aS. Narasimhan, S. Scherpe, A. Lucini Paioni, J. van der Zwan, G. E. Folkers, H. Ovaa, M. Baldus, *Angew Chem Int Ed Engl* **2019**, *58*, 12969-12973; bR. Ghosh, Y. Xiao, J. Kragelj, K. K. Frederick, *J Am Chem Soc* **2021**, *143*, 18454-18466; cB. J. Albert, C. Gao, E. L. Sesti, E. P. Saliba, N. Alaniva, F. J. Scott, S. T. Sigurdsson, A. B. Barnes, *Biochemistry* **2018**, *57*, 4741-4746; dH. Takahashi, I. Ayala, M. Bardet, G. De Paepe, J. P. Simorre, S. Hediger, *J Am Chem Soc* **2013**, *135*, 5105-5110.

[6] aR. Ghosh, R. Dumarieh, Y. Xiao, K. K. Frederick, *J Magn Reson* **2022**, *336*, 107150; bK. M. McCoy, R. Rogawski, O. Stovicek, A. E. McDermott, *J Magn Reson* **2019**, *303*, 115-120.

[7] B. J. Lim, B. E. Ackermann, G. T. Debelouchina, *Chembiochem* **2020**, *21*, 1315-1319.

[8] aJ. T. Paletta, M. Pink, B. Foley, S. Rajca, A. Rajca, *Org Lett* **2012**, *14*, 5322-5325; bA. P. Jagtap, I. Krstic, N. C. Kunjir, R. Hansel, T. F. Prisner, S. T. Sigurdsson, *Free Radic Res* **2015**, *49*, 78-85; cS. Bleicken, T. E. Assafa, H. Zhang, C. Elsner, I. Ritsch, M. Pink, S. Rajca, G. Jeschke, A. Rajca, E. Bordignon, *ChemistryOpen* **2019**, *8*, 1057-1065.

[9] aG. Mathies, M. A. Caporini, V. K. Michaelis, Y. Liu, K. N. Hu, D. Mance, J. L. Zweier, M. Rosay, M. Baldus, R. G. Griffin, *Angew Chem Int Ed Engl* **2015**, *54*, 11770-11774; bR. Harrabi, T. Halbritter, F. Aussenac, O. Dakhlaoui, J. van Tol, K. K. Damodaran, D. Lee, S. Paul, S. Hediger, F. Mentink-Vigier, S. T. Sigurdsson, G. De Paepe, *Angew Chem Int Ed Engl* **2022**, *61*, e202114103.

[10] F. Mentink-Vigier, S. Paul, D. Lee, A. Feintuch, S. Hediger, S. Vega, G. De Paepe, *Phys Chem Chem Phys* **2015**, *17*, 21824-21836.

[11] K. K. Frederick, V. K. Michaelis, B. Corzilius, T. C. Ong, A. C. Jacavone, R. G. Griffin, S. Lindquist, *Cell* **2015**, *163*, 620-628.

[12] aK. Tagami, A. Equbal, I. Kaminker, B. Kirtman, S. Han, *Solid State Nucl Magn Reson* **2019**, *101*, 12-20; bG. Stevanato, G. Casano, D. J. Kubicki, Y. Rao, L. Esteban Hofer, G. Menzildjian, H. Karoui, D. Siri, M. Cordova, M. Yulikov, G. Jeschke, M. Lelli, A. Lesage, O. Ouari, L. Emsley, *J Am Chem Soc* **2020**, *142*, 16587-16599.

[13] K. Singewald, M. J. Lawless, S. Saxena, *J Magn Reson* **2019**, *299*, 21-27.

[14] G. Karthikeyan, A. Bonucci, G. Casano, G. Gerbaud, S. Abel, V. Thome, L. Kodjabachian, A. Magalon, B. Guigliarelli, V. Belle, O. Ouari, E. Mileo, *Angew Chem Int Ed Engl* **2018**, *57*, 1366-1370.
[15] L. Chen, L. Wu, X. Tan, A. Rockenbauer, Y. Song, Y. Liu, *J Org Chem* **2021**, *86*, 8351-8364.
[16] W. M. Yau, K. R. Thurber, R. Tycko, *J Magn Reson* **2014**, *244*, 98-106.
[17] K. R. Thurber, R. Tycko, *J Magn Reson* **2009**, *196*, 84-87.
[18] S. Stoll, A. Schweiger, *J Magn Reson* **2006**, *178*, 42-55.

Acknowledgements

Chapter 6, in full, is a reprint of the material submitted in A comparative study of nitroxide-based biradicals for dynamic nuclear polarization in cellular environments. 2022. Ackermann, Bryce E. Lim, Byung Joon, Narayanan, Sirish, Debelouchina, Galia T. The dissertation author was the primary investigator and author of this material.