DYNAMIC MULTI-BODY PROTEIN INTERACTIONS SUGGEST VERSATILE

PATHWAYS FOR COPPER TRAFFICKING

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Aaron Michael Keller

May 2012

DYNAMIC MULTI-BODY PROTEIN INTERACTIONS SUGGEST VERSATILE PATHWAYS FOR COPPER TRAFFICKING

Aaron Michael Keller, Ph. D.

Cornell University 2012

As part of intracellular copper trafficking pathways, the human copper chaperone Hah1 delivers Cu^+ to the Wilson's Disease Protein (WDP) via weak and dynamic protein–protein interactions. WDP contains six homologous metal binding domains (MBDs), connected by flexible linkers, which all can receive Cu^+ from Hah1. The functional roles of the MBD multiplicity in Cu^+ trafficking are not well understood. Building on our previous study of the dynamic interactions between Hah1 and the isolated 4th MBD of WDP, here we study how Hah1 interacts with MBD34, a double-domain WDP construct, using single-molecule fluorescence resonance energy transfer (smFRET) combined with nanovesicle trapping. By alternating the positions of the smFRET donor and acceptor, we systematically probed Hah1–MBD3, Hah1–MBD4, and MBD3–MBD4 interaction dynamics within the multi-domain system. We found that the two interconverting interaction geometries were conserved in both intermolecular Hah1–MBD and intramolecular MBD–MBD interactions. The Hah1–MBD interactions within MBD34 are stabilized by an order of magnitude relative to Hah1 interactions with the isolated single-MBDs, and thermodynamic evidence suggests that Hah1 can interact with both MBDs simultaneously. We proposed 2-body and 3-body interaction models based on previously known structures of homologous proteins to describe both intermolecular Hah1–MBD and intramolecular MBD–MBD interactions. These interaction models and our smFRET results were then used to

formulate and quantify a comprehensive Hah1–MBD34 mechanism. The enhanced interaction stability of Hah1 with the multi-MBD system, the dynamic intramolecular MBD–MBD interactions, and the ability of Hah1 to interact with multiple MBDs simultaneously suggest an efficient and versatile mechanism for the Hah1-to-WDP pathway to transport Cu^+.

BIOGRAPHICAL SKETCH

Aaron Keller grew up in Topeka, Kansas where he earned his ACS certified B.S. degree in Chemistry at Washburn University. As an undergraduate, Aaron's research project involved examining the NMR spectra of cobalt coordination compounds using DFT calculations. Aaron later moved to Ithaca, New York where he would attend graduate school in the Department of Chemistry and Chemical Biology at Cornell University. Working for Professor Peng Chen, he used single-molecule fluorescence microscopy to study weak and dynamic protein–protein interactions involved in the intracellular trafficking of copper. Aaron will begin a postdoctoral appointment with Jim Werner and Jennifer Hollingsworth at Los Alamos National Laboratory where he will perform 3-dimensional single-molecule tracking in live cells using blinking-suppressed quantum dots. While not immersed in research, Aaron enjoys music, art, exploring the wilderness, and spending time with his family and friends.

Dedicated to Dixie and Lilly Anna

ACKNOWLEDGMENTS

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1. Weak and Dynamic Protein Interactions Regulate Copper Homeostasis

Copper is an essential cofactor for many enzymes, but it must be safely transported and regulated in the cell to avoid harmful effects such as oxidative damage.[1-3] In humans, the copper chaperone, Hah1, specifically delivers Cu^+ via weak and dynamic protein interactions to two homologous P_{IB}-type ATPases (Figure 1): the Wilson's Disease Protein (WDP) and the Menkes' Disease Protein (MNK).[4-7] WDP is predominately found in the liver and brain, while MNK is generally found in non-hepatic tissues.[1,8] Under normal physiological conditions both WDP and MNK reside in the trans-Golgi network and use ATP hydrolysis to drive Cu^+ translocation from the cytosol into the Golgi for later incorporation into various copper enzymes; e.g., WDP delivers Cu^+ to ceruloplasmin, an enzyme involved in iron homeostasis, while MNK delivers Cu^+ to tyrosinase, an enzyme involved in the oxidation of phenols.[9,10] Under elevated Cu^+ stress, they relocalize for the export of Cu^+ from the cell (Figure 1).[11-13] Genetic defects in WDP and MNK result in copper toxicity and deficiency disorders respectively.[8,14-16]

Figure 2 presents a schematic of the tertiary structure of WDP and MNK. Both WDP and MNK have *six* cytosolic N-terminal metal binding domains (MBDs, numbered 1–6 starting from the N-terminus) connected by flexible linkers of various lengths. Interestingly, the number of MBDs varies between one and six depending on the organism, with higher organisms generally having more MBDs.[17] The catalytic core of WDP/MNK consists of eight transmembrane helices (TMH) constituting the Cu^+ pump, a nucleotide binding (N-) and phosphotase (P-) domain involved in the binding and hydrolysis of ATP, and an actuator (A-) domain involved in the regulation of Cu^+ translocation.[13,18]

1
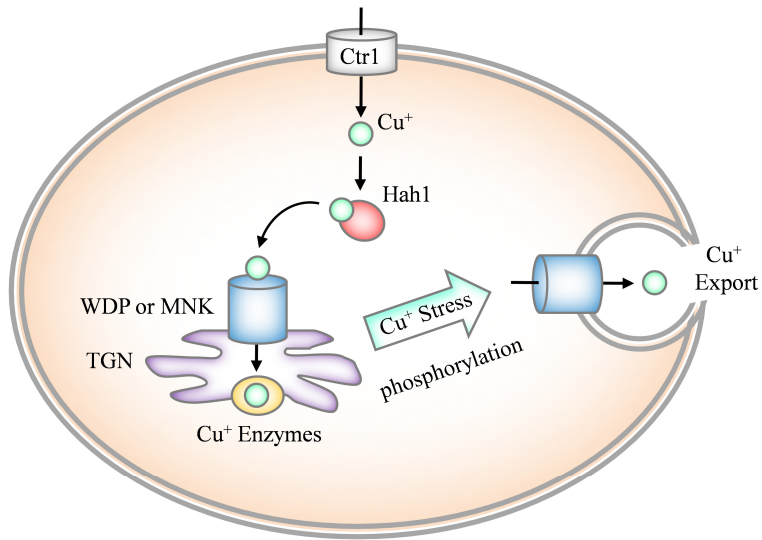
Figure 1. After Cu$^+$ enters the cell via Ctr1, Hah1 safely delivers the metal to either WDP or MNK. Under normal physiological conditions, WDP and MNK are located within the TGN where they deliver Cu$^+$ to various metalloenzymes. Elevated levels of Cu$^+$ induce kinase-mediated phosphorylation and relocalization of WDP/MNK for Cu$^+$ efflux.
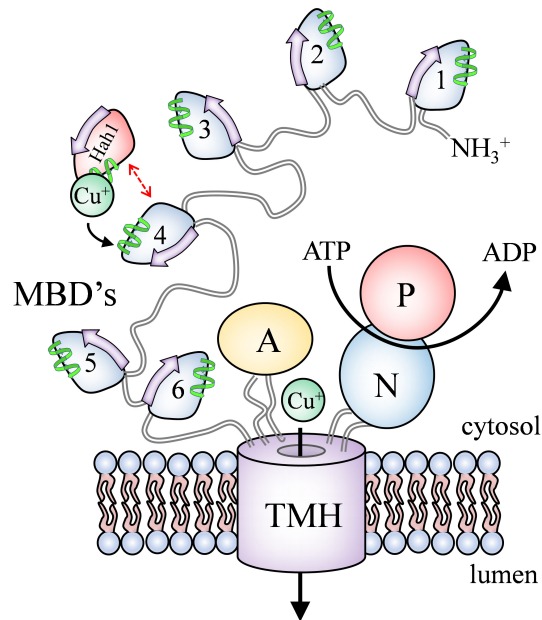
Figure 2. WDP and MNK contain six homologous MBDs connected by flexible linkers of various lengths. Each MBD can acquire Cu$^+$ from Hah1 via weak and dynamic protein interactions. Cu$^+$ is translocated from the cytosol to the lumen through the TMH using the energy from ATP hydrolysis. The N- and P-domains bind and hydrolyze ATP, while the A-domain regulates the translocation of Cu$^+$ through the TMH.

Many similarities exist between Hah1 and the MBDs of MNK and WDP. The individual WDP/MNK MBDs and Hah1 are all homologous, each having a $\beta\alpha\beta\beta\alpha\beta$ protein fold and a CXXC Cu$^+$-binding motif.[19] They can all bind Cu$^+$, with a similar affinity ($\sim 10^{10}$ M^{-1})[20] although recent measurements report binding constants as high as $\sim 10^{18}$ M^{-1}.[21-23] Under a shallow thermodynamic gradient, Hah1 can transfer Cu$^+$ to each MBD with similar efficiency (equilibrium constant for Cu$^+$ transfer is $\sim 10^{0}$).[20,24] The Cu$^+$ transfer is mediated by weak and dynamic protein interactions, $K_D \sim \mu$M, consisting of hydrogen bonds, electrostatic interactions, and metal-bridging of two proteins' CXXC motifs.[4,6,7,25-28] This is illustrated by the NMR structure of Hah1 interacting with the 1st MBD of MNK (Figure 3).
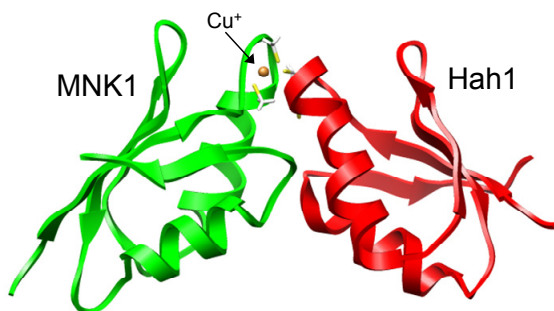
Figure 3. NMR structure (PDB code 2K1R)[29] of the interaction between Hah1 (red) and MNK1 (green). This weak interaction is mediated by hydrogen bonding and electrostatic interactions. The Cu$^+$ transfer occurs via a 3-coordinate Cu$^+$–thiolate intermediate which bridges the two proteins' CXXC motifs.

Despite the many similarities among the MBDs, various differences exist, which perhaps define their differences in functional roles. These differences include electrostatic potentials, locations within the N-terminal tail, ability to re-orient with regard to the adjacent linkers, and complex formation with other MBDs or Hah1.[30-33] A combination of yeast two-hybrid assays[6,34]

and NMR studies[24,31,35,36] have shown that, in general, Hah1 preferentially interacts with MBDs 1–4 over MBDs 5–6. Further, yeast complementation[37,38] and cellular imaging studies[39] showed that only MBDs 5–6 were essential for maintaining WDP function. Taken together, these data suggest that MBDs 1–4 may be more important for Cu^+ sensing or acquisition via protein interactions with Hah1, whereas MBDs 5–6 are more important for Cu^+ translocation at the trans-Golgi network and WDP/MNK relocalization for Cu^+ efflux.

After acquisition of Cu^+ from Hah1, it is possible that MBDs 1–4 subsequently deliver the metal to MBDs 5–6 via intramolecular-interdomain interactions for eventual Cu^+ translocation through the TMH.[31] In this scenario, the MBDs are directly involved in the trafficking and delivery of the metal via Hah1–MBD and MBD–MBD interactions. This direct-trafficking has been proposed as less likely because no observation of Cu^+ transfer from the MBDs to the TMH has been observed, and homologous P_{IB}-type ATPases retain their ability to receive and translocate Cu^+ upon partial or complete removal of their N-terminal MBDs.[40]

Alternatively, the MBDs may behave more like a Cu^+ sensor, in which large-scale tertiary conformational changes in the N-terminal tail regulate WDP/MNK activity or localization.[30,33,35,41-43] For the bacterial homolog of WDP, CopA, recent structural studies have shown MBD interactions with the N- and A-domains, which may serve as a mechanism for the modulation of ATPase activity.[44,45] In addition, Cu^+-dependent phosphorylation sites located in the flexible linkers between MBDs are involved in regulating the cellular localization and function of MNK and WDP.[46-51] Both Hah1–MBD interactions and MBD–MBD interactions would likely change the tertiary structure of the N-terminal tail, and hence modulate the catalytic activity for Cu^+ translocation or change the accessibility of phosphorylation sites signaling the relocalization of WDP/MNK for Cu^+ efflux.

In either role, metal trafficking or regulatory, both intermolecular Hah1–MBD interactions and intramolecular MBD–MBD interactions are vital to proper WDP/MNK function and Cu^+ homeostasis in the cell. Characterizing and understanding these weak and dynamic protein interactions remains an important, yet challenging, task. Surface plasmon resonance studies have been used to study the kinetics of these interactions,[52,53] but non-specific protein–surface interactions may perturb the kinetics and the measurements lack structural information. NMR,[24,29,31,35,36,54-57] X-ray crystallography,[25,58-60] protein docking,[26] and molecular dynamics (MD) simulations[28,32,33,61] have provided detailed structural information on the interaction interfaces, but can only provide estimates on the interaction thermodynamics and kinetics.

1.2. Single-molecule FRET with Nanovesicle Trapping to Probe Hah1–MBD Interactions

To compliment these studies while overcoming some of their limitations, we use single-molecule fluorescence resonance energy transfer (smFRET)[62] in combination with nanovesicle trapping[63-66] to quantify weak and dynamic Hah1–MBD and MBD–MBD interactions (Figure 4). By labeling Hah1 and the MBDs with a smFRET donor-acceptor pair (Cy3-Cy5), we can monitor the interaction dynamics of a single pair in real time. A 532-nm laser excites the donor fluorophore using the evanescent wave resulting from total internal reflection (TIR). The extent of energy transfer from the donor to acceptor, $E_{FRET} = \left(1 + \left(r/r_0\right)^6\right)^{-1}$, is dependent on the distance between the donor and acceptor, r, and a constant, r_0, which accounts for the spectral overlap, quantum yields, and relative orientation of the donor-acceptor pair.[67] Because E_{FRET} is sensitive to distance changes on the ~nm scale, the smFRET technique can probe protein–protein association/dissociation and/or ~nm changes in protein conformation or interaction geometry.

The proteins are trapped within surface-immobilized ~100-nm-diameter lipid vesicles, which maintain an effective protein concentration of ~μM for a single molecule inside a vesicle,

needed for studying weak interactions ($K_D \sim$ μM), while also eliminating non-specific protein–surface interactions. After coating a quartz slide with biotinylated polyethylene glycol (PEG) or bovine serum albumin (BSA), the nanovesicles containing the trapped proteins are immobilized via biotin-avidin chemistry at ~pM concentrations to allow for the imaging of single fluorophores (Figure 4).

This single-molecule approach eliminates the need for synchronization and allows us to observe interaction intermediates that are otherwise masked in ensemble-averaged measurements. Because we only analyze vesicles containing a single donor-acceptor pair, this approach also eliminates homodimeric interactions between two proteins of the same type, which are unavoidable in ensemble measurements.
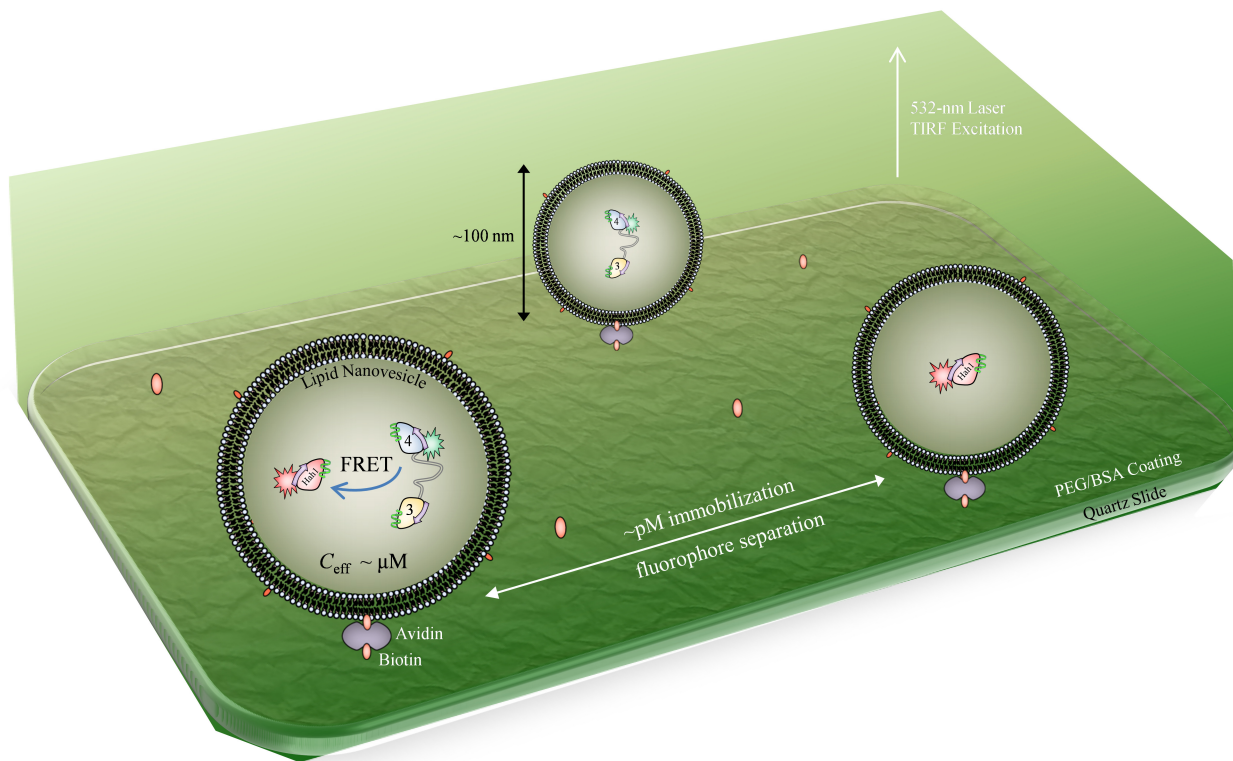
Figure 4. Schematic of nanovesicle trapping with smFRET. The labeled proteins are trapped within ~100-nm-diameter lipid vesicles to maintain a high C_{eff} ~ μM needed to study weak interactions. Subsequently, the nanovesicles are immobilized on a BSA or PEG-coated quartz slide using biotin–avidin chemistry at ~pM concentrations, which allows for spatial separation of the individual fluorophores needed for single-molecule imaging. The donor fluorophore (Cy3, green) is excited by TIRF at 532-nm (green background). Upon protein–protein interaction, the donor transfers its energy via FRET to the acceptor (Cy5, red).

We have previously used this approach to study interactions between Hah1 and the isolated 4th MBD of WDP, denoted MBD4^{SD} (SD: "single domain"), both in the absence (Figure 5) and presence of Cu$^+$ (Figure 6).[68-70]

We labeled Hah1 and MBD4^{SD} at their C-terminal residues with Cy5 and Cy3, respectively (Figure 5A). By measuring the anticorrelated Cy3 and Cy5 intensities of a single Hah1–MBD4^{SD} pair in real time (Figure 5B), we observed three dynamically interconverting E_{FRET} states in the absence of Cu$^+$: $E_{Dissoc} \sim 0.15$, $E_{Mid} \sim 0.50$, and $E_{High} \sim 0.80$. Here, we approximate E_{FRET} as $I_{Cy5}/(I_{Cy5} + I_{Cy3})$, where I_{Cy5} and I_{Cy3} represent the intensities of Cy5 and Cy3, respectively. We assigned the E_{Dissoc} observable to the dissociated state, Hah1 + MBD4^{SD}, due to its relatively low value ($\sim$0.15), which is in agreement with that observed in control measurements where free Cy3 and Cy5 molecules are co-trapped.[66] We assigned the E_{Mid} and E_{High} observables to two Hah1–MBD4^{SD} interaction geometries, denoted ^{1}Hah1–MBD4^{SD} and ^{2}Hah1–MBD4^{SD}, respectively.

As shown in Figure 5B, the dwell times, τ_{Dissoc}, τ_{Mid}, and τ_{High} of E_{Dissoc}, E_{Mid}, and E_{High}, respectively, can be pooled from many Hah1–MBD4^{SD} pairs to examine their respective distributions (Figure 5C). Fitting each distribution with a single-exponential decay function gives the average lifetimes of each E_{FRET} state, $\overline{\tau}_{Dissoc}$, $\overline{\tau}_{Mid}$, and $\overline{\tau}_{High}$. Based on our assignments of the chemical species to each E_{FRET} state, we have presented a kinetic mechanism describing the Hah1–MBD4^{SD} interactions (Figure 5D). From the average lifetimes (Figure 5C) and the relative transition frequency to each state, we can extract all six kinetic rate constants (Figure 5E).[66,69] The rates of association (k_1 and k_2) are $\sim$10^5 M^{-1} s^{-1}, while the rates of dissociation (k_{-1} and k_{-2}) and interconversion (k_3 and k_{-3}) are all $\sim$1 s^{-1}. From the respective ratios of the rate constants, we've also quantified the dissociation constants for each interaction complex (K_{D1} and K_{D2}) which are both $\sim$10^0 μM, and equilibrium of interconversion ($K_3 \sim 1$) as reported in Figure

8

5F. The stabilities of ^{1}Hah1–MBD4^{SD} and ^{2}Hah1–MBD4^{SD} can also be quantified from the Gaussian area ratios of E_{Mid}/E_{Dissoc} and E_{High}/E_{Dissoc} respectively obtained from the compiled E_{FRET} distribution of many interacting pairs (Figure 6A). These area ratios yield dissociation constants $K_{D1} = 7 \pm 6$ μM and $K_{D1} = 9 \pm 7$ μM, which are in agreement with those obtained from the kinetic analysis (Figure 5F).

In the presence of varying stoichiometric equivalents of Cu$^+$, we still observed three dynamically interconverting E_{FRET} states, which have the same E_{FRET} values within experimental error as E_{Dissoc}, E_{Mid}, and E_{High}.[70] The two interaction geometries, ^{1}Hah1–MBD4^{SD} and ^{2}Hah1–MBD4^{SD}, are therefore retained regardless of the Cu$^+$ stoichiometry. However, the stabilities the interaction geometries do change depending on the stochiometric amount of Cu$^+$ (Figure 6). Upon addition of 1 equivalent of Cu$^+$ per protein pair, both ^{1}Hah1–MBD4^{SD} and ^{2}Hah1–MBD4^{SD} interactions are stabilized slightly by a factor of ~1.3 (Figure 6B). We attributed this stabilization to metal-bridging of the two CXXC motifs,[4,6,7,25-28] similar to that observed in the Hah1–MNK1 complex (Figure 3). The distance between the donor and acceptor labeling sites in the Hah1–MNK1 structure, ~5 nm, is indeed in agreement with our measured $E_{Mid} \sim 0.50$ (possible interaction geometries corresponding to E_{High} will be subsequently discussed in Chapter 4, Section 4.1). Under excess equivalents of Cu$^+$ (Figure 6C and D), the CXXC motifs of both Hah1 and MBD4^{SD} become fully metallated, which disrupts the metal-bridging and destabilizes both interaction geometries.
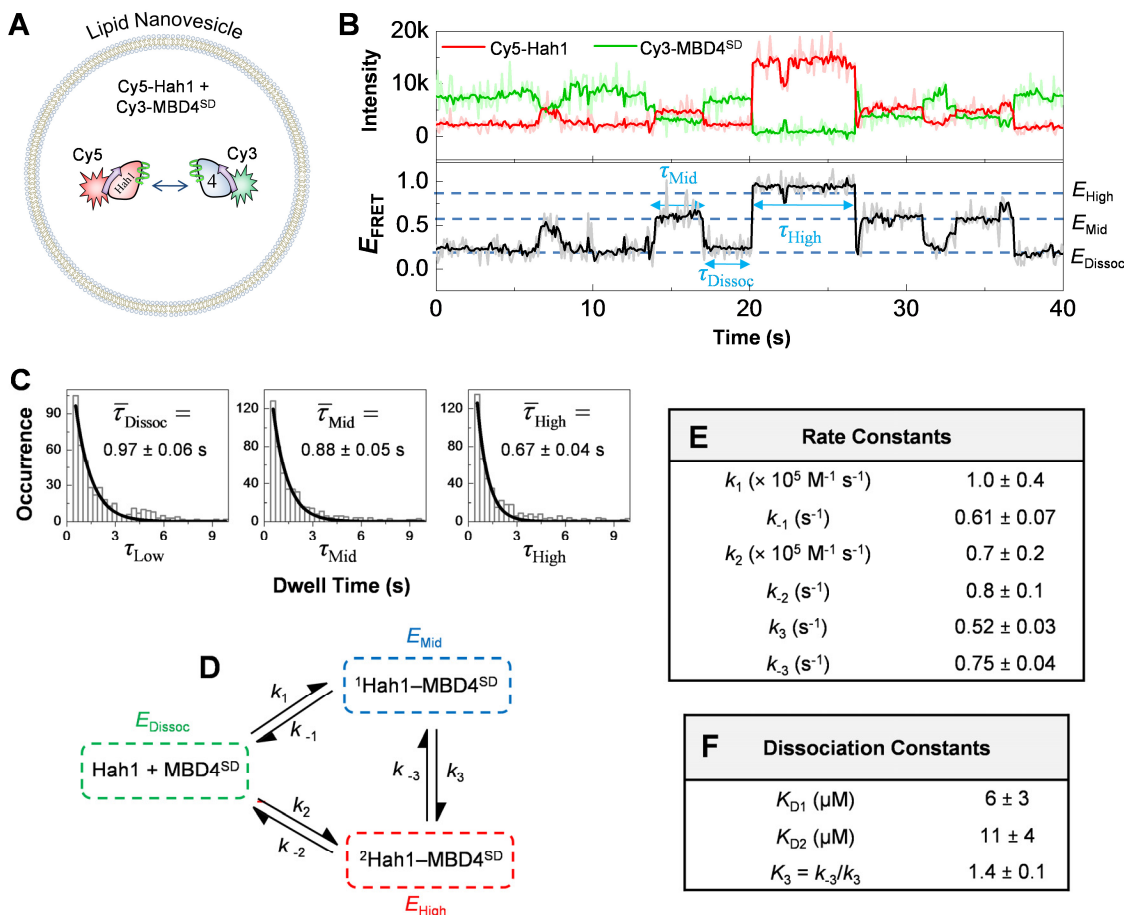
Figure 5. (A) Cy5-Hah1 + Cy3-MBD4^{SD} labeling scheme and corresponding (B) smFRET trajectory. The lighter colors show the original fluorescence intensities (light green for Cy3 and light red for Cy5) and corresponding E_{FRET} (light grey) while the darker colors (dark green for Cy3, dark red for Cy5, and black for E_{FRET}) represent data subjected to non-linear forward-backward filtering (Chapter 2, Section 2.9). Three dynamically interconverting E_{FRET} states are apparent: E_{Dissoc}, E_{Mid}, and E_{High} with dwell times τ_{Dissoc}, τ_{Mid}, and τ_{High}, respectively. (C) The dwell time distributions of τ_{Dissoc}, τ_{Mid}, and τ_{High} for hundreds of interacting pairs can be fit with a single-exponential decay function to obtain the average dwell times, $\overline{\tau}_{Dissoc}$, $\overline{\tau}_{Mid}$, and $\overline{\tau}_{High}$. (D) Kinetic mechanism for Hah1–MBD4^{SD} interactions in the absence of Cu$^+$: E_{Dissoc} contains the dissociated state, Hah1 + MBD4^{SD}. E_{Mid} and E_{High} contain two Hah1–MBD4^{SD} interaction geometries, denoted ^{1}Hah1–MBD4^{SD} and ^{2}Hah1–MBD4^{SD}, respectively. (E) Rate constants computed from $\overline{\tau}_{Dissoc}$, $\overline{\tau}_{Mid}$, and $\overline{\tau}_{High}$ and the ratio of transitions to each state. (F) Dissociation constants, K_{D1} and K_{D2}, for ^{1}Hah1–MBD4^{SD} and ^{2}Hah1–MBD4^{SD}, respectively, and equilibrium of interconversion, K_3, as computed from the respective ratios of the rate constants.
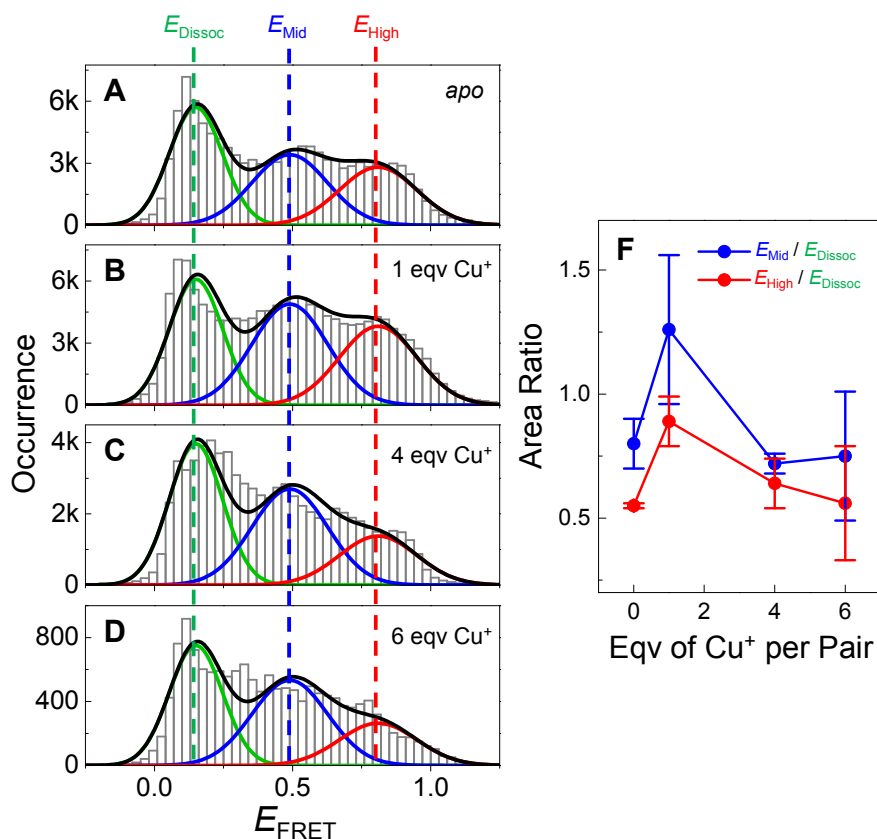
Figure 6. Compiled E_{FRET} distributions for Cy5-Hah1 + Cy3-MBD4^{SD} in (A) the absence of Cu$^+$ (*apo*), and in the presence of (B) 1, (C) 4, and (D) 6 equivalents of Cu$^+$ per protein pair. The distributions were fit globally by sharing the Gaussian peak positions and widths of each E_{FRET} state. (F) The E_{Mid}/E_{Dissoc} and E_{High}/E_{Dissoc} Gaussian area ratios are plotted with increasing equivalents of Cu$^+$ to assess the relative stabilities of the interaction complexes.

We hypothesized that multiple interaction geometries increase the probability of forming a complex for Cu$^+$ transfer. If only one interaction geometry is productive for Cu$^+$ transfer, then the presence of this additional geometry increases the overall rate of association and provides an alternate pathway to the productive complex via interconversion. Additionally, the two interaction geometries may have distinct interfaces which allow for Hah1 to interact with multiple MBDs simultaneously, where concerted motions among the MBDs may facilitate the copper trafficking process. The latter possibility provides a motivation for studying the interaction dynamics of Hah1 with a multi-MBD system of WDP.

11

1.3. Probing Hah1–MBD and MBD–MBD Interactions within a Multi-MBD Construct

Here we seek to understand Hah1–WDP interactions in the context of WDP's multi-MBD structure, using a WDP construct containing its 3rd and 4th MBDs, MBD34. We have systematically probed Hah1–MBD3, Hah1–MBD4, and MBD3–MBD4 interactions using a series of smFRET labeling schemes, which are described in the following chapters. Note that we use the terms "MBD3" and "MBD4" to refer to the respective MBDs within the multidomain MBD34 construct. To probe whether Hah1–MBD interactions are influenced by the presence of additional MBDs, we compare Hah1–MBD4^{SD} vs. Hah1–MBD4 interactions. To probe whether Hah1 interacts preferentially with certain MBDs, we compare Hah1–MBD4 vs. Hah1–MBD3. To understand intramolecular-interdomain interactions between MBDs and how they are coupled to intermolecular interactions with Hah1, we have studied the MBD3–MBD4 interactions both in the absence and presence of Hah1. To probe what role multiple MBDs play during the trafficking of Cu^+, we have studied the Cu^+-dependence of Hah1–MBD4 interactions. Finally, we propose a mechanism for the Hah1–multi-MBD interactions considering our smFRET results in conjunction with proposed protein interaction models based on previous structural data. After quantitative analysis of the interaction thermodynamics, the functional implications of this mechanism for Cu^+ trafficking are subsequently discussed.

EXPERIMENTAL METHODS

2.1. Protein Construct Design and Structure

The pET-21b(+) Hah1 and pET-32 Xa/LIC MBD34 plasmids were obtained from Amy Rosenzweig's Lab at Northwestern University. To ensure site-specific labeling by fluorescent probes, all native cysteines, apart from those located in the CXXC Cu^+-binding motif, were mutated to serines. For Hah1, this included C41S, and for MBD34, these were C71S, C124S, and C197S. Subsequently, a cysteine residue was introduced at a desired location for labeling with a fluorescent probe via maleimide chemistry. This additional labeling cysteine was the C-terminal C69 in Hah1, C206 at the C-terminus in the $MBD34^{L4}$ construct, D95C in the $MBD34^{L3}$ construct, and both D95C and C206 in the $MBD34^{L34}$ construct. The "L4," "L3," and "L34" superscripts indicate the labeling cysteine is on MBD4, MBD3, or both MBD3 and MBD4, respectively. In all MBD34 constructs, the Cu^+-binding cysteines in the CXXC motif of MBD3 were mutated to alanines (i.e., C34A and C37A) to study the Cu^+-dependence of MBD4 exclusively. The locations of the Cu^+-binding motifs, cysteine-to-serine mutations, and labeling sites are annotated within the structure of Hah1 (Figure 7A) and MBD34 (Figure 7B). Site-directed mutagenesis was performed in the Novablue (Novagen) or DH5α (Invitrogen) cloning strain for MBD34 and Hah1 constructs, respectively. All mutations were confirmed by DNA sequencing at the Cornell University Life Sciences CLC Genomics Facility. The distribution of linker distances from the 20 NMR structures of MBD34 is plotted in Figure 7C, giving an average distance of between the MBD's of ~7 nm.
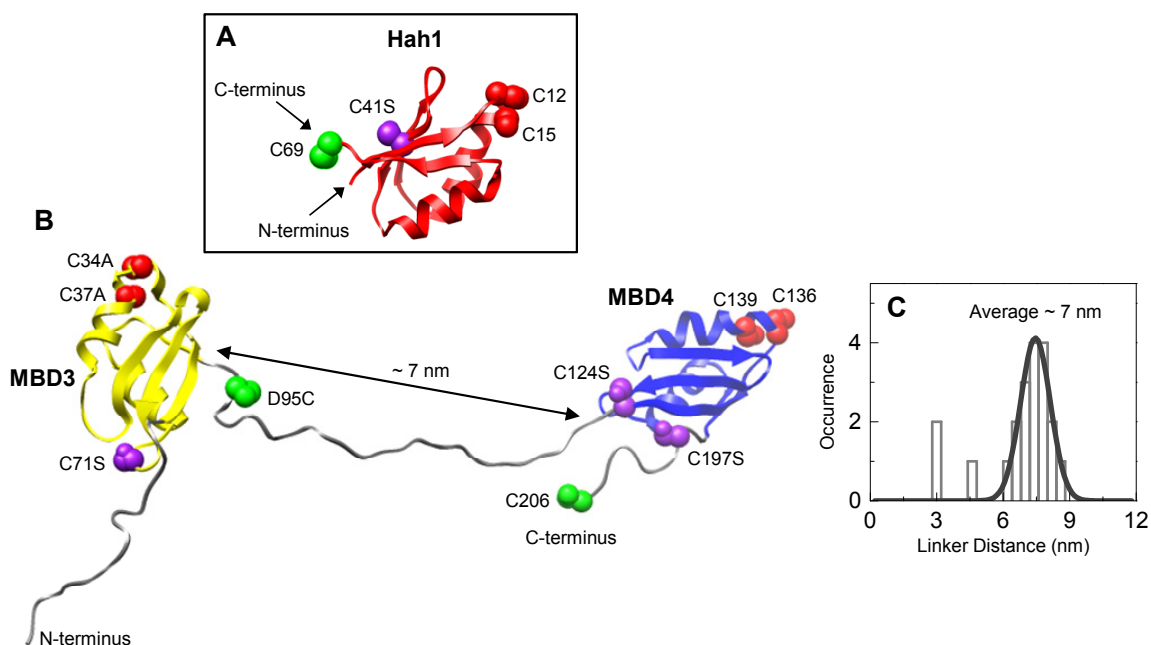
Figure 7. (A) Structure of Hah1 (2K1R)[29] and (B) representative structure of MBD34 (2ROP)[57] showing amino acids of interest as spheres. The CXXC or AXXA motifs are shown in red, cysteine-to-serine mutations in purple, and labeling residues in green. The full structure containing the linker region was provided by Dr. Francesca Cantini (Center for Magnetic Resonance, University of Florence). (C) Distribution of the linker distances in MBD34, as measured between the α-carbons of L93 and C124. The average distance between MBD3 and MBD4 is ~7 nm.

2.2. Protein Expression, Purification, and Characterization

All proteins were expressed and purified as described previously.[68,71] Briefly, the pET-21b(+) Hah1 plasmid was transformed into the BL21(DE3) expression strain (Novagen) and grown to an optical density at 600 nm of 0.8–1.1 before a 3 hr induction with 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG, US Biological). Cells were harvested by centrifugation, re-suspended, and lysed via multiple freeze-thaw cycles in 20 mM MES (US Biological), pH 6.0 buffer containing 1 mM EDTA (Fisher). The filtered lysate was purified through anion exchange (HiTrap Q HP), cation exchange (HiTrap SP HP), and gel filtration (Superdex 200 Prep Grade) columns using FPLC (Amersham Biosciences, GE Healthcare). For the ion exchange columns,

the loading buffer was 20 mM MES, pH 6.0 and the elution buffer was the same with the addition of 1 M NaCl. The gel filtration column was run on the protein pre-reduced with 1 mM *tris*(2-carboxyethyl)phosphine (TCEP, Sigma) using 20 mM MES, 150 mM NaCl, pH 6.0 buffer.

The pET-32 Xa/LIC MBD34 plasmids were transformed into the BL21(DE3)pLysS expression strain (Novagen) and grown to an optical density at 600 nm of 0.8–1.1, followed by a 3 hr induction with 0.9 mM IPTG. The cells were harvested via centrifugation, re-suspended, and lysed via stirring at room temperature for 30 min in 50 mM HEPES (US Biological), 500 mM NaCl, 30 mM imidazole, 5% glycerol, pH 7.5 buffer containing an EDTA-free protease inhibitor mini-tablet (Roche). The MBD34 constructs contained multiple tags at the N-terminus, including a HisTag which allowed for the purification of the protein with a Ni^{2+} column (HisTrap HP), using the lysis buffer for loading and a linear gradient to 300 mM imidazole, pH 7.5 to elute the sample (Buffer B). After imidazole removal via buffer-exchange to 50 mM HEPES, 500 mM NaCl, 5% glycerol, pH 7.5 (Buffer A), 1 unit of Factor Xa (EMD Biosciences) per 100 μg protein was used to cleave the N-terminal purification tags. After 16 hrs at room temperature, the Factor Xa cleavage reaction was stopped by the addition of 1 mM phenylmethylsulfonyl fluoride (PMSF, Sigma). The cleaved protein was then further purified through a Ni^{2+} column with Buffer A as the loading buffer and Buffer B as the elution buffer, followed by a gel filtration column with Buffer A on the protein pre-reduced with 1 mM TCEP.

The purities of all protein samples were confirmed using SDS-PAGE (>95%). Their identities were further confirmed using mass spectrometry (MALDI) at the University of Michigan Protein Structure Facility: $21,453 \pm 21$ Da (expected 21,453 Da) for MBD34^{L4}, $21,335 \pm 21$ Da (expected 21,315 Da) for MBD34^{L3}, and $21,408 \pm 21$ Da (expected 21,425 Da) for MBD34^{L34}. The purity and identity of Hah1 was previously confirmed.[68]

2.3. Fluorescent Labeling, Purification, and Characterization

The labeling reactions were typically performed on the scale of 400–800 μL at a concentration of 0.5 mM protein in 100 mM phosphate, pH 7.0 buffer. The reaction vial was flushed with nitrogen after each step to remove oxygen. After reducing the protein with 10 equivalents of TCEP for 1 hr at room temperature, the cysteines in the CXXC motifs were protected by the addition of 3–6 equivalents of $[Cu(CH_3CN)_4]PF_6$ (Sigma), pre-dissolved in a 50% aqueous solution of acetonitrile. After 1 hr, the targeted cysteine was fluorescently labeled by adding 4–6 equivalents of Cy3- or Cy5-maleimide (GE Healthcare), pre-dissolved in anhydrous dimethylsulfoxide. For the labeling of the MBD34[L34] construct, Cy3- and Cy5-maleimide were added simultaneously as a 1:1 mixture. The labeling reaction was incubated at room temperature for 4 hrs, and overnight at 4 °C.

After quenching the excess maleimide dye with ∼7 equivalents β-mercaptoethanol (βME, Sigma), the unreacted dye was removed using gel filtration (Superdex Peptide HR 10/30, GE Healthcare) with 60 mM MES, 20 mM βME, pH 6.0 buffer for Hah1, or with 50 mM HEPES, 20 mM βME, pH 7.5 buffer for the MBD34 labeling reactions. The negative charge on the fluorescent dyes allowed for the separation of labeled from unlabeled species using anion exchange (Tricorn MonoQ 5/50 GL, GE Healthcare) to a purity greater than 95%. The anion exchange column was run with the same buffer used for the dye removal, and eluted with a gradient to this buffer containing 1 M NaCl.

The labeling stoichiometry and specificity were determined by mass spectrometry. The Hah1 construct is mono-labeled specifically at the C-terminus by Cy5 as reported previously.[68] The MBD34[L4] and MBD34[L3] proteins were confirmed to be mono-labeled with Cy3 by MALDI: 22,181 ± 22 Da (expected 22,205 Da) for Cy3-MBD34[L4], and 22,083 ± 22 Da (expected 22,087 Da) for Cy3-MBD34[L3]. The MBD34[L34] construct was found to be double-labeled with a Cy3 and

a Cy5 using LC-MS: 22,969 ± 2 Da (expected 22,968 Da). Although likely present, the MBD34^{L34} protein containing two Cy3 or two Cy5 labels was not detected by LC-MS.

The Cy3-MBD34^{L4}, Cy3-MBD34^{L3}, and Cy3Cy5-MBD34^{L34} proteins were found to be labeled specifically at their targeted cysteines by analyzing the masses of the trypsin (Roche) digested fragments via LC-MS. For Cy3-MBD34^{L4}, we observed the labeling fragment containing C206-Cy3 (3,403 ± 1 Da, expected 3,405 Da). For Cy3-MBD34^{L3}, the labeling fragment containing C95-Cy3 is 2,238 ± 1 Da (expected 2,238 Da). For Cy3Cy5-MBD34^{L34}, the labeling fragments are those containing C95-Cy3 (2,237 ± 1 Da, expected 2,238 Da), C95-Cy5 (2,263 ± 1 Da, expected 2,264 Da), C206-Cy3 (3,403 ± 1 Da, expected 3,405 Da), and C206-Cy5 (3,429 ± 1 Da, expected 3,431 Da). In all MBD34 constructs, the digested fragments containing the CXXC Cu$^+$-binding motif were found to be unlabeled (6,503 ± 1 Da, expected 6,506 Da). The complete protection of the CXXC motif from maleimide labeling by Cu$^+$ also confirmed that each mutant protein retained its Cu$^+$ binding capability.

2.4. Cu$^+$ Removal

The Cu$^+$ was subsequently removed from the proteins by repeated overnight incubations with 5000 equivalents of bicinchoninic acid (BCA, Sigma), followed by removal of excess BCA and BCA-chelated-Cu$^+$ via ultracentrifugal buffer exchange (using Amicon ultracentrifugation filters with molecular weight cut-offs of 5 kDa for Hah1 and 10 kDa for MBD34). The incubation-removal process was repeated 5–8 times, and the samples were checked for copper content using the BCA copper quantification assay.[72] In all cases, the *apo* samples were found to have a Cu$^+$-to-protein ratio of less than 0.04 (the detection limit of BCA copper assay is ~2 μM). The large excess of BCA needed to completely remove Cu$^+$ from the labeled proteins (25,000–

40,000 equivalents) again confirmed the retention of each proteins' strong Cu^+-binding capabilities.

2.5. Nanovesicle Trapping

A detailed description of our experimental protocol using nanovesicle trapping in combination with smFRET has been reported previously.[66,68] To trap the proteins inside lipid nanovesicles, a sample solution containing 3 μM of each labeled protein, 15 μM TCEP, 2 mM Trolox, and the appropriate amount of $[Cu(CH_3CN)_4]PF_6$ in 60 mM MES, 110 mM NaCl, pH 6.0 buffer (Buffer W) was added to a lipid film containing L-α-phosphatidylcholine (Egg-PC, Avanti Polar Lipids) mixed with 1% 1,2-dipalmitoyl-*sn*-glycero3-phosphoethanolamine-N-(cap biotinyl) (16:0 biotinyl-cap-PE, Avanti Polar Lipids). For the experiment involving Cy3Cy5-MBD34^{L34} in the presence of excess Hah1, 60 μM of unlabeled Hah1 (20 equivalents over Cy3Cy5-MBD34^{L34}) was also present within the hydrating sample along with 1 mM TCEP to ensure sufficient reduction. The final concentration of Egg-PC after hydration was 5 mg/mL, and that of biotinyl-cap-PE was 0.05 mg/mL. After hydrating the lipids for 1 hr at room temperature, the sample was passed through a polycarbonate membrane with a 100-nm pore size ~53 times using a mini-extruder (Avanti Polar Lipids) to make mono-dispersed 100-nm-diameter lipid vesicles.

2.6. Single-molecule FRET Experiments

A homemade quartz flow cell used for single-molecule total internal reflection fluorescence (TIRF) imaging was washed with Buffer W and incubated with 1 mg/mL biotinylated-albumin (Sigma) for 45 min, followed by incubation with 0.2 mg/mL Neutravidin (Invitrogen) for 15 min. After washing the flow cell with excess Buffer W, the sample

containing the proteins trapped in the biotinylated-nanovesicles was incubated at ~30 pM for 5–15 min. The sample incubation time was adjusted depending on the observed density of fluorophores to ensure spatial separation of the nanovesicles. The non-immobilized nanovesicles were washed out with excess Buffer W, and an oxygen scavenging system in Buffer W, supplemented with 2 mM Trolox, was used for the TIRF imaging. We found that partially oxidized Trolox (~6%) prepared before the experiment by brief UV-illumination was critical for maintaining the photostability of Cy3 and Cy5.[73] The oxygen scavenging system consisted of either 0.2 mg/mL glucose oxidase (Sigma), 0.05 mg/mL catalase (Sigma or Roche), and 4% (w/v) (+)-D-glucose (Sigma) (aka the "GODCAT" system) or 50 nM protocatechuate-3,4-dioxygenase (PCD, Sigma) with 2.5 mM protocatechuic acid (PCA, Sigma). The GODCAT system was refreshed in 30-min intervals to maintain pH stability,[74] while the PCD-PCA system was refreshed every 3 hrs.[75]

The flow cell was mounted on a modified IX71 Olympus microscope[76] where Cy3 fluorophores were excited using prism-type TIRF across a $68 \times 137 \ \mu m^2$ area, with a depth of penetration of ~200 nm using a circularly-polarized 532-nm laser (CrystaLaser) at a power of a few mW. The fluorescence emission of both Cy3 and Cy5 were collected through a 60×, NA = 1.2, water-immersion objective (Olympus) and passed through a Dual-View system (Optical Insights) to separate (635DCXR Dichroic, Chroma) and project the Cy3 fluorescence (filtered by HQ580-60m, Choma) and Cy5 fluorescence (filtered by HQ660LP, Chroma) onto each half of an EMCCD camera (Andor iXon). Before entering the Dual-View, the laser light was rejected using a HQ550LP filter (Chroma). Images of Cy3 and Cy5 fluorescence intensities were collected every 50–100 ms for ~2.5 min using Andor iQ software. After 2 min, Cy5 fluorophores were excited directly using a 637-nm laser (CrystaLaser) to check if additional Cy5 fluorophores were present within the nanovesicle. The background-subtracted integrated

19

intensities (over a $\sim 2 \times 2$ μm^2 area) of single Cy3 and Cy5 fluorophores were extracted for each image using a custom-written IDL program to obtain single-molecule fluorescence intensity trajectories.

2.7. Protein–Dye and Protein–Lipid Interaction Controls

To ensure that the labeled proteins are not interacting with the dyes, we performed a control experiment in which Cy3-MBD34 was cotrapped with free Cy5. A single anticorrelated photobleaching event was observed within an imaging area of 46,720 μm^2, whereas typically several hundred such events were observed over this area when Cy5-Hah1 is cotrapped with Cy3-MBD34. A similar experiment cotrapping Cy5-Hah1 with free Cy3 was performed previously, also showing negligible protein–dye interactions.[66,68]

To test whether the proteins were significantly interacting with the lipids within the nanovesicle, we coated the quartz slide in the flow cell with Egg-PC and subsequently flowed in $\sim$nM concentrations of Cy3-MBD34. We only observed a few tens of molecules immobilized as compared with over 1000 immobilized molecules observed during biotin-avidin specific immobilization. Although this number is slightly greater than that observed in the case of Cy5-Hah1,[68] the non-specific MBD34–EggPC interactions are still relatively infrequent ($\sim 6\%$) and are therefore likely to have a minimal contribution to the protein interaction dynamics.

2.8. Nanovesicle Trapping Probabilities

Under direct laser excitation, the number of fluorophores present within each vesicle can be determined by counting the number of photobleaching steps. In several control experiments, we imaged both Cy3 and Cy5 with simultaneous 532-nm and 637-nm laser excitation to determine how many molecules were trapped in each vesicle. The vesicle occupancy and

colocalization probabilities of Cy5-Hah1 and Cy3-MBD34 were found to be similar to that observed in our previous study.[66,68] The majority of the vesicles (86%) only contained one molecule, either Cy3- or Cy5-labeled protein, while only 5% contained one Cy3 and one Cy5 colocalized. The probability of cotrapping two Cy5-Hah1 molecules with one Cy3-MBD34 was small (0.2%), only ~4% relative to the one-to-one Cy3-Cy5 population and thus negligible. During our normal smFRET experiments, as mentioned previously, we further minimized this probability by rejecting trajectories showing additional Cy5 photobleaching events upon 637-nm laser excitation.

2.9. Single-molecule FRET Data Analysis

Only trajectories that showed Cy3-Cy5 anticorrelated intensity fluctuations followed by single-step photobleaching events in both channels were analyzed. The extent of energy transfer from the donor to acceptor, or FRET efficiency, E_{FRET}, was approximated as $I_{Cy5}/(I_{Cy3} + I_{Cy5})$, where I_{Cy3} and I_{Cy5} represent the intensities of Cy3 and Cy5 respectively.

To remove noise and aid in the resolution of E_{FRET} states, I_{Cy3} and I_{Cy5} were filtered using a nonlinear forwards-backwards algorithm[77-79] with filtering parameters $M = 3$, $p = 20$, $K = 5$, where M is the window size, p is an exponential weighting factor, and K is the number of forward-backward predictors. Each intensity (either Cy3 or Cy5), I, for frame number i was filtered according to

$$\hat{I}(i) = \sum_{k=1}^{K} \left[f_k(i)\hat{I}_k^f(i) + b_k(i)\hat{I}_k^b(i) \right]$$

where $\hat{I}_k^f(i)$ is the forward predictor (found by averaging $M - 1$ data points before i) and $\hat{I}_k^b(i)$ is the backward predictor (found by averaging $M - 1$ data points after i). $f_k(i)$ and $b_k(i)$ represent the forward and backward weighting factors, respectively, given by

$$f_k(i) = \left\{ \sum_{j=0}^{M-1} \left[\left(I_{\text{Cy5}}(i-j) - \hat{I}^f_{\text{Cy5},k}(i-j) \right)^2 + \left(I_{\text{Cy3}}(i-j) - \hat{I}^f_{\text{Cy3},k}(i-j) \right)^2 \right] \right\}^{-p}$$

$$b_k(i) = \left\{ \sum_{j=0}^{M-1} \left[\left(I_{\text{Cy5}}(i-j) - \hat{I}^b_{\text{Cy5},k}(i-j) \right)^2 + \left(I_{\text{Cy3}}(i-j) - \hat{I}^b_{\text{Cy3},k}(i-j) \right)^2 \right] \right\}^{-p}$$

The value of M was chosen to be approximately half of the shortest average waiting time, to avoid "smoothing" transitions between E_{FRET} states. Dwell time analysis of each E_{FRET} state (described in Chapter 3, Section 3.4) before and after filtering resulted in identical average lifetimes; thus, this value of M did not "smooth" the majority of the transitions. The values of K and p were typical values recommended by the developers of the algorithm.[77] This nonlinear filter was shown to be effective in reducing noise in single-molecule trajectories.[78,79]

CHAPTER 3

PROBING HAH1–MBD34 INTERACTIONS WITH SINGLE-MOLECULE FRET

3.1. FRET Labeling Schemes for Dissecting Protein Interactions

We used three donor-acceptor (Cy3-Cy5) FRET labeling schemes to dissect the inter- and intra-molecular interactions among Hah1 and the two MBDs of MBD34. To directly observe Hah1–MBD4 interactions, we labeled Hah1 at its C-terminus (i.e., C69) with Cy5, and MBD4 at its C-terminus (i.e., C206) with Cy3 (Figure 8A); we refer to this MBD34 construct as "Cy3-MBD34^{L4}," where L4 denotes that the label is on MBD4. To directly observe Hah1–MBD3 interactions, we labeled MBD3 at its C-terminus (i.e., C95) with Cy3 (Figure 8D); we refer to this MBD34 construct as "Cy3-MBD34^{L3}." To observe intramolecular-interdomain MBD3–MBD4 interactions within MBD34, we labeled these two domains at their respective C-terminals (i.e., C95 and C206) with the Cy3-Cy5 pair (Figure 8G); the Cy3 or Cy5 can be attached to either of the two cysteines; we refer to this double-labeled MBD34 construct as "Cy3Cy5-MBD34^{L34}." This double labeling also generated MBD34 molecules that contain two Cy3 or two Cy5 probes; these could be easily differentiated in our smFRET experiments and were excluded in our data analysis (Chapter 2, Section 2.9). In all MBD34 constructs, the cysteines in the CXXC motif of MBD3 were mutated to alanines to remove MBD3's Cu^{+}-binding capability, so Cu^{+}-transfer could only occur between Hah1 and MBD4.
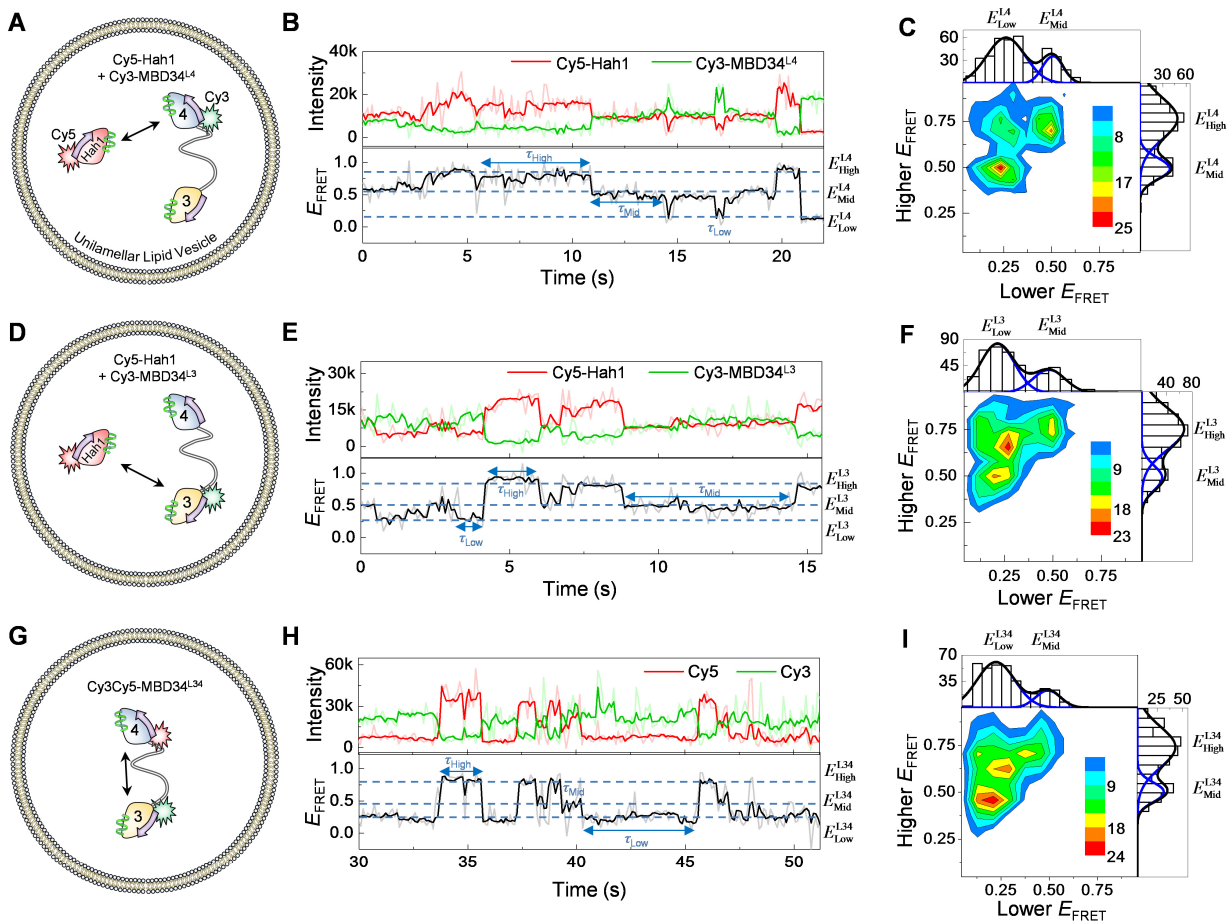
Figure 8. (A) Cy5-Hah1 + Cy3-MBD34^{L4} labeling scheme to probe Hah1–MBD4 interactions, (B) corresponding smFRET trajectory, and (C) two-dimensional (2-D) histogram of the average lower vs. average higher E_{FRET} state for 226 pairs. (D) Cy5-Hah1 + Cy3-MBD34^{L3} labeling scheme to probe Hah1–MBD3 interactions, (E) corresponding smFRET trajectory, and (F) 2-D E_{FRET} histogram for 285 pairs. (G) Cy3Cy5-MBD34^{L34} labeling scheme to probe MBD3–MBD4 interactions, (H) corresponding smFRET trajectory, and (I) 2-D E_{FRET} histogram for 248 molecules. In all trajectories, the light colors show the original fluorescence intensities and corresponding E_{FRET} while the darker colors represent data subjected to non-linear forward-backward filtering (Chapter 2, Section 2.9). For each 2-D histogram, three populations are observed, corresponding to E_{Low}-E_{Mid}, E_{Low}-E_{High}, and E_{Mid}-E_{High} combinations. The 1-D projections of the histograms and their Gaussian resolution allow for the determination of the center E_{FRET} values.

3.2. Observation of Dynamic Protein–Protein Interactions in the Absence of Cu⁺.

By trapping two protein molecules labeled with the Cy3-Cy5 pair or a single protein labeled with this pair within an immobilized nanovesicle and measuring their smFRET, we first studied the dynamic interactions between Hah1 and MBD34 and between the two domains of MBD34 in the absence of Cu⁺.

For all labeling schemes we observed anticorrelated Cy3-Cy5 donor-acceptor fluorescence intensity fluctuations, reporting the dynamic intermolecular Hah1–MBD (Figure 8B, E) and intramolecular MBD3–MBD4 interactions (Figure 8H). In each case, three interconverting E_{FRET} states are apparent, at E_{FRET} ~0.2, ~0.5, and ~0.8. We refer to these respective states as E_{Low}^{L4}, E_{Mid}^{L4}, and E_{High}^{L4} for the Cy5-Hah1 + Cy3-MBD34^{L4} labeling scheme; E_{Low}^{L3}, E_{Mid}^{L3}, and E_{High}^{L3} for the Cy5-Hah1 + Cy3-MBD34^{L3} labeling scheme; and E_{Low}^{L34}, E_{Mid}^{L34}, and E_{High}^{L34} for the Cy3Cy5-MBD34^{L34} labeling scheme.

Because many smFRET trajectories showed only two E_{FRET} states before Cy3 or Cy5 photobleached, we pooled data from a few hundred interacting pairs for each labeling scheme and examined the 2-dimensional (2-D) histogram of the average lower (either E_{Low} or E_{Mid}) vs. higher (either E_{Mid} or E_{High}) E_{FRET} values Figure 8C, F, I). Three distinct populations are clear in each 2-D histogram, whose Gaussian-resolved peak values are $E_{Low}^{L4} = 0.26 \pm 0.01$, $E_{Mid}^{L4} = 0.50 \pm 0.01$, and $E_{High}^{L4} = 0.77 \pm 0.01$ for Cy5-Hah1 + Cy3-MBD34^{L4}; $E_{Low}^{L3} = 0.21 \pm 0.01$, $E_{Mid}^{L3} = 0.49 \pm 0.02$, and $E_{High}^{L3} = 0.76 \pm 0.01$ for Cy5-Hah1 + Cy3-MBD34^{L3}; and $E_{Low}^{L34} = 0.22 \pm 0.01$, $E_{Mid}^{L34} = 0.50 \pm 0.02$, and $E_{High}^{L34} = 0.75 \pm 0.01$ for Cy3Cy5-MBD34^{L34}.

The large values of E_{Mid} and E_{High} indicate the formation of intermolecular Hah1–MBD or intramolecular MBD3–MBD4 interaction complexes, as the Cy3 and Cy5 labels should be

within a few nm from each other. The significant difference between E_{Mid} and E_{High} indicates that these two E_{FRET} states correspond to two protein interaction geometries, similar to that observed in our previous study of Hah1 interacting with the single-domain construct, MBD4^{SD}.[68] The similarity in the E_{Mid} and E_{High} values across the three labeling schemes indicates that the two interaction geometries are conserved between Hah1–MBD4, Hah1–MBD3, and MBD3–MBD4 interactions, consistent with Hah1 and all WDP MBDs being homologous in sequence and structure.

For the two labeling schemes that probe intermolecular Hah1–MBD interactions, the E_{Low}^{L4} and E_{Low}^{L3} states represent the case where the proteins are far apart, yet their E_{FRET} values are higher than that of a completely dissociated state, $E_{Dissoc} = 0.15 \pm 0.01$, which was independently determined in control experiments with vesicles containing free Cy3 and Cy5 and in our previous study of Hah1–MBD4^{SD} interactions.[66,68] Therefore, besides the dissociated species, the E_{Low}^{L4} state must contain Cy5-Hah1 interactions with the unlabeled MBD3 in Cy3-MBD34^{L4}, and the E_{Low}^{L3} state must contain Cy5-Hah1 interactions with the unlabeled MBD4 in Cy3-MBD34^{L3}; these interactions would give an $E_{FRET} \sim 0.22$ based on our E_{FRET} vs. distance calibration curve (Figure S1A), considering the MBD3-to-MBD4 linker distance of $\sim$7 nm (Figure 7C).

For Cy3Cy5-MBD34^{L34}, E_{Low}^{L34} corresponds to the state where MBD3 and MBD4 are separate with their linker in an extended conformation (denoted as MBD34$_{ext}$); this extended conformation was observed in MBD34's NMR structure (Figure 7B).[57] As noted above, the average distance between the two domains is $\sim$7 nm in MBD34$_{ext}$, corresponding to an expected $E_{FRET} \sim 0.22$ (Figure 7C and Figure S1A), consistent with the value of E_{Low}^{L34}. It is worth noting

that $E_{\text{Low}}^{\text{L34}}$ also approximates the E_{FRET} value when Cy5-Hah1 interacts with the respective unlabeled domain of Cy3-MBD34$^{\text{L4}}$ or Cy3-MBD34$^{\text{L3}}$ (i.e., $E_{\text{Low}}^{\text{L34}} \approx E_{\text{Low}}^{\text{L3}} \approx E_{\text{Low}}^{\text{L4}}$).

Combining all above results, we have studied how Hah1 interacts with each MBD within the double-domain WDP construct, MBD34, and how the two domains in MBD34 interact with each other in the absence of Cu$^+$. Hah1 can interact with each MBD forming two different complexes. Correlating with our previous work on Hah1 interaction with MBD4$^{\text{SD}}$,[68-70] the results show that Hah1 can interact with MBD4 in two geometries in both the single-domain MBD4$^{\text{SD}}$ and the double-domain MBD34 constructs. For MBD3, our results represent the first direct observation of any complex formation with Hah1. Moreover, we observed two MBD3–MBD4 interactions with geometries similar to Hah1–MBD interactions; this is the first direct observation of intramolecular-interdomain complexes between MBD3 and MBD4. It appears that the two interaction geometries are conserved for any pair between Hah1, MBD3, and MBD4, which are all homologous to each other. This conservation in interaction geometries indicates that Hah1, MBD3, and MBD4 use similar protein surface patches for their interactions.

3.3. Stabilities of Hah1–MBD34 Intermolecular Interactions in the Absence of Cu$^+$.

Figure 9B and C show the E_{FRET} distributions from hundreds of smFRET trajectories of Cy5-Hah1 interacting with Cy3-MBD34$^{\text{L4}}$ or Cy3-MBD34$^{\text{L3}}$, along with that of Cy5-Hah1 interacting with Cy3-MBD4$^{\text{SD}}$ that we reported previously (Figure 9A).[66,68] These E_{FRET} distributions can be Gaussian-resolved to individual E_{FRET} states; the relative areas of the resolved peaks reflect the relative stabilities of chemical species associated with the E_{FRET} states. For Cy5-Hah1 + Cy3-MBD4$^{\text{SD}}$, as determined previously,[66,68] the E_{FRET} distribution contains three peaks (Figure 9A): $E_{\text{Dissoc}} \sim 0.15$, $E_{\text{Mid}} \sim 0.50$, and $E_{\text{High}} \sim 0.81$, corresponding to the dissociated state and the two Hah1–MBD4$^{\text{SD}}$ interaction complexes, respectively.

27

For Cy5-Hah1 + Cy3-MBD34$^{\text{L4}}$, we used four Gaussian peaks to resolve the E_{FRET} distribution (Figure 9B). Two of them are centered at $E_{\text{Mid}}^{\text{L4}} \sim 0.50$ and $E_{\text{High}}^{\text{L4}} \sim 0.75$, as resolved in Figure 8C, which correspond to the complexes between Hah1 and MBD4. For the other two, one accounts for the dissociated state (E_{Dissoc}), which is centered at ~ 0.15 as resolved from the Cy5-Hah1 + Cy3-MBD4$^{\text{SD}}$ E_{FRET} distribution (Figure 9A) and appears as a shoulder in Figure 9B; the other accounts for the state where Cy5-Hah1 forms complexes with the unlabeled MBD3 in Cy3-MBD34$^{\text{L4}}$ ($E_{\text{Low}}^{\text{L4}}{}' \sim 0.27$) as discussed in Section 3.2, and its center was floated in Gaussian-resolving the E_{FRET} distribution. Note, E_{Dissoc} and $E_{\text{Low}}^{\text{L4}}{}'$ are unresolved within the $E_{\text{Low}}^{\text{L4}}$ state in Figure 8C.

Similarly, for Cy5-Hah1 + Cy3-MBD34$^{\text{L3}}$, we used four Gaussian peaks to resolve the E_{FRET} distribution, centered at $E_{\text{Dissoc}} \sim 0.15$, $E_{\text{Low}}^{\text{L3}}{}' \sim 0.27$, $E_{\text{Mid}}^{\text{L3}} \sim 0.50$, and $E_{\text{High}}^{\text{L3}} \sim 0.75$, corresponding to the dissociated state, the state where Hah1 interacts with the unlabeled MBD4, and the two states where Hah1 complexes with MBD3, respectively (Figure 9C). We shared the center positions and widths of these peaks with those for resolving the E_{FRET} distribution of Cy5-Hah1 + Cy3-MBD34$^{\text{L4}}$ in Figure 9B; this is a valid approximation as the center positions determined through the 2-D E_{FRET} analyses are the same within experimental error for Cy5-Hah1 + Cy3-MBD34$^{\text{L4}}$ and Cy5-Hah1 + Cy3-MBD34$^{\text{L3}}$ (Figure 8C, F). Furthermore, because E_{Dissoc} represents the same dissociated state for both Cy5-Hah1 + Cy3-MBD34$^{\text{L4}}$ and Cy5-Hah1 + Cy3-MBD34$^{\text{L3}}$ labeling schemes, the relative peak area of the E_{Dissoc} state is shared between Figure 9B and C.
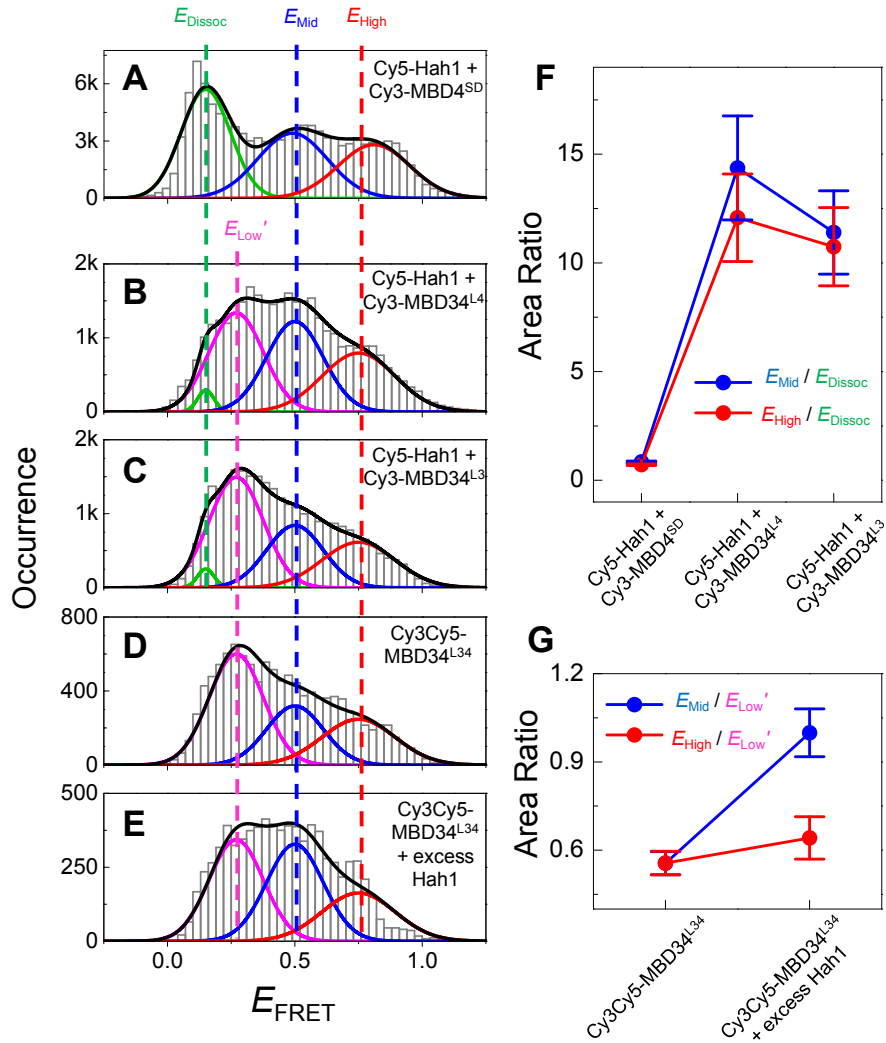
Figure 9. Compiled E_{FRET} distributions (bin size = 0.03) and corresponding Gaussian fits for (A) Cy5-Hah1 + Cy3-MBD4^{SD}, (B) Cy5-Hah1 + Cy3-MBD34^{L4}, (C) Cy5-Hah1 + Cy3-MBD34^{L3}, (D) Cy3Cy5-MBD34^{L34}, and (E) Cy3Cy5-MBD34^{L34} with excess Hah1. The E_{FRET} distributions were fitted globally by sharing the center values and widths of each E_{FRET} state. (F) The area ratios of E_{Mid} and E_{High} states with respect to E_{Dissoc} for Cy5-Hah1 + Cy3-MBD4^{SD}, Cy5-Hah1 + Cy3-MBD34^{L4}, and Cy5-Hah1 + Cy3-MBD34^{L3} determined from A−C. (G) The area ratios of E_{Mid} and E_{High} with respect to E_{Low}' for Cy3Cy5-MBD34^{L34} and Cy3Cy5-MBD34^{L34} with excess Hah1 determined from D−E.

The deconvolution of $E_{Low}{}'$ from E_{Dissoc} in these E_{FRET} distributions allows us to account for Cy5-Hah1's interactions with the unlabeled domain within Cy3-MBD34^{L4} or Cy3-MBD34^{L3}, so that we can quantify the stability of the Hah1–MBD interactions relative to the dissociated state E_{Dissoc}, given by the peak area ratios in the E_{FRET} distributions. In both Cy5-Hah1 + Cy3-MBD34^{L4} and Cy5-Hah1 + Cy3-MBD34^{L3} labeling schemes, the E_{Mid}/E_{Dissoc} and E_{High}/E_{Dissoc} area ratios are comparable (Figure 9F), indicating that Hah1–MBD3 and Hah1–MBD4 interactions have similar stabilities. However, compared with Hah1–MBD4^{SD} interactions, Hah1 interactions with MBD4 and MBD3 within MBD34 are both more stable by an order of magnitude, reflected by the increase in E_{Mid}/E_{Dissoc} and E_{High}/E_{Dissoc} area ratios (Figure 9F). This increased stability indicates that there are concerted actions between the two MBDs (and perhaps the linker region) in MBD34 that facilitate interactions with Hah1.

The similar stability of Hah1–MBD3 and Hah1–MBD4 interactions suggests that Hah1 does not have a significant preference for interacting with one MBD over the other within the double-domain MBD34 construct. This is contrary to previous NMR experiments, which only detected Hah1 complex formation with MBD4 and not MBD3 (in the presence of Cu$^+$),[24,57] but is in agreement with the yeast two-hybrid assay, which also detected Hah1–MBD3 interactions.[6] The comparable complex formation of Hah1 with these MBDs observed here may explain why Cu$^+$-loaded Hah1 can fully metallate both MBD3 and MBD4.[20,24,57]

The area percentages (χ) of the E_{FRET} states in the E_{FRET} distributions can be used to analyze the population percentages of all complexes. From the Cy5-Hah1 + Cy3-MBD34^{L4} E_{FRET} distribution (Figure 9B), the total population percentage of Hah1 in complexes with MBD4 is $\chi_{Mid}^{L4} + \chi_{High}^{L4} = 62 \pm 4\%$. From the Cy5-Hah1 + Cy3-MBD34^{L3} E_{FRET} distribution (Figure 9C), the total population percentage of Hah1 in complexes with MBD3 is $\chi_{Mid}^{L3} + \chi_{High}^{L3} = 52 \pm 9\%$. Interestingly, the total population percentage of Hah1 in complexes with either MBD3

or MBD4 would then sum to greater than unity ($114 \pm 10\%$). This suggests there must be overlap ($>14 \pm 10\%$) between the population of Hah1 in complexes with MBD3 and that of Hah1 in complexes with MBD4. In other words, there must be a population in which Hah1 is in close proximity with MBD4 and MBD3 *simultaneously*, e.g., forming 3-body interactions.

In the Cy5-Hah1 + Cy3-MBD34$^{\text{L4}}$ E_{FRET} distribution (Figure 9B), the $E_{\text{Low}}^{\text{L4}}{}'$ state represents Hah1 interactions with the unlabeled MBD3 in the MBD34$_{\text{ext}}$ conformation (as any interactions involving MBD4 are contained in $E_{\text{Mid}}^{\text{L4}}$ or $E_{\text{High}}^{\text{L4}}$). Its population percentage, $\chi_{\text{Low}}^{\text{L4}}{}' = 35 \pm 4\%$, is significantly less than the population percentage of all Hah1 complexes with MBD3 determined from the Cy5-Hah1 + Cy3-MBD34$^{\text{L3}}$ E_{FRET} distribution ($\chi_{\text{Mid}}^{\text{L3}} + \chi_{\text{High}}^{\text{L3}} = 52 \pm 9\%$, Figure 9C); the difference between the above percentages again indicates that there is overlap ($17 \pm 10\%$) between the apparent Hah1–MBD3 and Hah1–MBD4 complexes.

Similarly, in the Cy5-Hah1 + Cy3-MBD34$^{\text{L3}}$ E_{FRET} distribution (Figure 9C), the $E_{\text{Low}}^{\text{L3}}{}'$ state represents Hah1 interactions with the unlabeled MBD4 in the MBD34$_{\text{ext}}$ conformation. Its population percentage, $\chi_{\text{Low}}^{\text{L3}}{}' = 46 \pm 9\%$, is less than the population percentage of all Hah1 complexes with MBD4 determined from the Cy5-Hah1 + Cy3-MBD34$^{\text{L4}}$ E_{FRET} distribution ($\chi_{\text{Mid}}^{\text{L4}} + \chi_{\text{High}}^{\text{L4}} = 62 \pm 4\%$, Figure 9B). In agreement with the analysis above, the difference ($16 \pm 8\%$) again reflects a population overlap in the apparent Hah1–MBD3 and Hah1–MBD4 complexes.

To summarize, the population analysis of Cy5-Hah1 interacting with Cy3-MBD4$^{\text{SD}}$, Cy3-MBD34$^{\text{L4}}$, and Cy3-MBD34$^{\text{L3}}$ demonstrates that Hah1 interactions with MBDs within the double-domain MBD34 construct are significantly more stable than with the isolated MBD4$^{\text{SD}}$.

Moreover, Hah1–MBD3 and Hah1–MBD4 interactions have similar stability; and Hah1 can be in close proximity to MBD3 and MBD4 *simultaneously,* with a percentage occurrence of ~16%.

3.4. Timescales of Intermolecular Hah1–MBD34 Interactions in the Absence of Cu$^+$.

From the smFRET trajectories (Figure 8B, E, H), we can quantify the stochastic dwell times (τ_{Low}, τ_{Mid}, and τ_{High}) of the three E_{FRET} states (E_{Low}, E_{Mid}, and E_{High}). The distributions of these dwell times all follow single-exponential decay approximately: $f(\tau) = N\exp(-\tau/\overline{\tau})$, where $\overline{\tau}$ is a time constant and N is a scaling factor (Figure 10 and Figure S2). $\overline{\tau}$ is also equivalent to the average of the respective dwell time and represents the apparent lifetime of that E_{FRET} state. Table 1 summarizes the apparent lifetimes ($\overline{\tau}_{\text{Low}}$, $\overline{\tau}_{\text{Mid}}$, and $\overline{\tau}_{\text{High}}$) of each E_{FRET} state for each labeling scheme.

For the Cy5-Hah1 + Cy3-MBD34$^{\text{L4}}$ and Cy5-Hah1 + Cy3-MBD34$^{\text{L3}}$ labeling schemes, $\overline{\tau}_{\text{Mid}}$ and $\overline{\tau}_{\text{High}}$ represent the apparent lifetimes of the interaction complexes between Hah1 and MBD4 or between Hah1 and MBD3. Unfortunately, $\overline{\tau}_{\text{Low}}$ is not the lifetime of the dissociated state, as the E_{Low} state in the smFRET trajectories contains contributions from both the dissociated state (E_{Dissoc}) and the complexes in which Cy5-Hah1 interacts with the unlabeled domain in Cy3-MBD34$^{\text{L4}}$ or Cy3-MBD34$^{\text{L3}}$ (i.e., E_{Low}'), as discussed in Sections 3.2 and 3.3. In both labeling schemes, the population of E_{Low}' state is much higher than that of E_{Dissoc}, shown by the resolved E_{FRET} distributions (Figure 9B, C). Therefore, $\overline{\tau}_{\text{Low}}$ predominantly represents the apparent lifetime of E_{Low}' for these two labeling schemes.

Comparing the apparent lifetimes of Hah1–MBD4 and Hah1–MBD4$^{\text{SD}}$ complexes (from Cy5-Hah1 interacting with Cy3-MBD34$^{\text{L4}}$ or Cy3-MBD4$^{\text{SD}}$), $\overline{\tau}_{\text{Mid}}$ is similar in both cases (~0.9 s), whereas $\overline{\tau}_{\text{High}}$ is slightly longer for Hah1–MBD4 (~1.1 s) vs. Hah1–MBD4$^{\text{SD}}$ (~0.7 s) (Table

32

1). These comparable or slightly longer lifetimes of the Hah1–MBD4 interaction complexes cannot account fully for the order-of-magnitude increase in complex stability observed in the E_{FRET} distributions (Figure 9F). Therefore, there must be a significant decrease in the lifetime of the dissociated state for Hah1 interacting with the double-domain construct MBD34. This decrease in the lifetime of the dissociated state has also been proposed by van Dongen *et al.* to rationalize the increased affinity of Hah1 interacting with a four-domain WDP construct (MBDs 1–4) as compared with single-isolated MBDs, arising from the presence of multiple binding sites for Hah1.[34] Unfortunately, we could not directly determine the lifetime of the dissociated state due to its minor contribution to the experimental $\overline{\tau}_{Low}$.
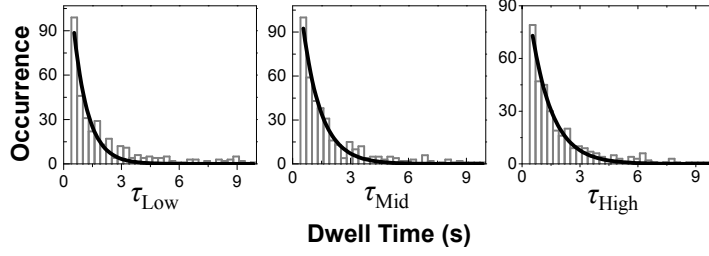
Figure 10. Distributions of the dwell time τ_{Low}, τ_{Mid}, and τ_{High} from E_{FRET} trajectories of Cy5-Hah1 + Cy3-MBD34^{L4}. Bin size = 0.3 s. Solid lines are fits with a single-exponential decay function, $f(\tau) = Ne^{-\tau/\bar{\tau}}$. Here $\bar{\tau}$ is the decay time constant, which also represents the average dwell time. N is a scaling factor. This analysis was performed for the dwell times from E_{FRET} trajectories for all labeling schemes (Appendix, Section S2). $\bar{\tau}_{Low}$, $\bar{\tau}_{Mid}$, and $\bar{\tau}_{High}$ represent the apparent lifetime of respective states (Table 1).

Table 1. The average dwell time (seconds) of each E_{FRET} state

Experiment	$\bar{\tau}_{Low}$	$\bar{\tau}_{Mid}$	$\bar{\tau}_{High}$
apo Cy5-Hah1 + Cy3-MBD4^{SD}	0.97 ± 0.07	0.88 ± 0.05	0.67 ± 0.04
apo Cy5-Hah1 + Cy3-MBD34^{L4}	0.77 ± 0.06	0.95 ± 0.05	1.09 ± 0.06
+ 1 eqv Cu$^+$	0.83 ± 0.07	0.85 ± 0.05	0.95 ± 0.05
+ 2 eqv Cu$^+$	0.62 ± 0.04	0.44 ± 0.02	0.53 ± 0.03
+ 4 eqv Cu$^+$	0.67 ± 0.04	0.43 ± 0.02	0.48 ± 0.02
apo Cy5-Hah1 + Cy3-MBD34^{L3}	0.89 ± 0.06	0.80 ± 0.04	0.69 ± 0.03
Cy3Cy5-MBD34^{L34}	0.59 ± 0.03	0.64 ± 0.03	0.61 ± 0.03
+ Excess Hah1	0.48 ± 0.02	0.38 ± 0.01	0.34 ± 0.01

Comparing the apparent lifetimes of Hah1–MBD4 and Hah1–MBD3 complexes (from Cy5-Hah1 interacting with Cy3-MBD34^{L4} or Cy3-MBD4^{L3}), $\bar{\tau}_{Mid}$ and $\bar{\tau}_{High}$ are slightly longer for Hah1–MBD4 complexes, but are still in similar magnitude (~1 s, Table 1). This is consistent with the similar stabilities of Hah1–MBD4 and Hah1–MBD3 interactions, given by the area ratios of their respective E_{FRET} states (Figure 9F). Their $\bar{\tau}_{Low}$'s are also similar, and mainly reflect the lifetimes for Cy5-Hah1 interactions with the respective unlabeled domain within Cy3-MBD34^{L4} or Cy3-MBD34^{L3}.

In short, here we have examined the apparent lifetimes of the E_{FRET} states for Cy5-Hah1 interacting with Cy3-MBD4^{SD}, Cy3-MBD34^{L4}, or Cy3-MBD34^{L3} to assess the kinetic aspect of the trends in complex stability for Hah1–MBD4^{SD}, Hah1–MBD4, and Hah1–MBD3 interactions. Comparing Hah1–MBD4 vs. Hah1–MBD3 interactions, the apparent lifetimes of the protein complexes are similar, consistent with their similar apparent stabilities. However, the apparent lifetimes of their complexes are also similar to those of Hah1–MBD4^{SD} interactions; therefore, the increased complex stabilities in Hah1's interactions with MBD34 likely come from an increase in the rate of protein association, rather than decrease in protein dissociation.

3.5. Stabilities and Kinetics of Intramolecular-Interdomain Interactions within MBD34.

Using Cy3Cy5-MBD34^{L34}, we observed intramolecular interactions between MBD3 and MBD4 (Section 3.2). Besides the extended conformation, MBD34$_{ext}$ (i.e, E_{Low}^{L34}), two intramolecular-interdomain complexes are clear (i.e., E_{Mid}^{L34} and E_{High}^{L34}), which interconvert dynamically (Figure 8H and I). Scheme 1 presents the simplest kinetic scheme describing these intramolecular-interdomain interactions within MBD34.

$$K_1^{34} = \frac{\left[{}^1MBD34\right]}{\left[MBD34_{ext}\right]} = \frac{k_1}{k_{-1}}$$

$$K_{int}^{34} = \frac{\left[{}^2MBD34\right]}{\left[{}^1MBD34\right]} = \frac{k_3}{k_{-3}}$$

$$K_2^{34} = \frac{\left[{}^2MBD34\right]}{\left[MBD34_{ext}\right]} = \frac{k_2}{k_{-2}}$$
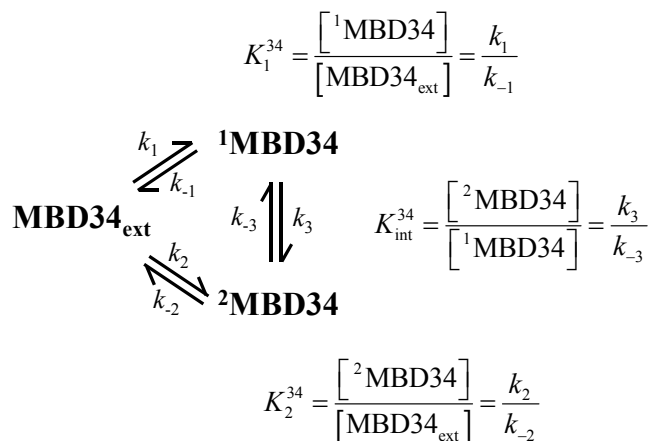
Scheme 1. Kinetic mechanism of the intramolecular-interdomain interactions between MBD3 and MBD4. Rate constants: $k_1 = 0.9 \pm 0.1$ s^{-1}, $k_{-1} = 0.88 \pm 0.09$ s^{-1}, $k_2 = 0.79 \pm 0.05$ s^{-1}, $k_{-2} = 0.76 \pm 0.08$ s^{-1}, $k_3 = 0.69 \pm 0.03$ s^{-1}, and $k_{-3} = 0.87 \pm 0.04$ s^{-1}. Equilibrium constants: $K_1^{34} = 1.0 \pm 0.2$, $K_2^{34} = 1.0 \pm 0.1$, and $K_{int}^{34} = 0.79 \pm 0.05$.

Accordingly, we used three Gaussian peaks to resolve the E_{FRET} distribution of Cy3Cy5-MBD34$^{\text{L34}}$ (Figure 9D). They are centered at $E_{\text{Low}}^{\text{L34}}{}' \sim 0.27$, $E_{\text{Mid}}^{\text{L34}} \sim 0.50$, and $E_{\text{High}}^{\text{L34}} \sim 0.75$, corresponding to the extended conformation, MBD34$_{\text{ext}}$, and the two intramolecular-interdomain complexes, denoted as ^{1}MBD34 and ^{2}MBD34, respectively. The center value and width of the $E_{\text{Low}}^{\text{L34}}{}'$ peak were shared with $E_{\text{Low}}^{\text{L3}}{}'$ and $E_{\text{Low}}^{\text{L4}}{}'$ in Figure 9B and C because the observed E_{FRET} for MBD34$_{\text{ext}}$ in the Cy3Cy5-MBD34$^{\text{L34}}$ labeling scheme should be similar to Cy5-Hah1 interacting with the unlabeled domain in Cy3-MBD34$^{\text{L4}}$ or Cy3-MBD34$^{\text{L3}}$. The center values and widths of $E_{\text{Mid}}^{\text{L34}}$ and $E_{\text{High}}^{\text{L34}}$ were shared with the other distributions as well because the two interaction geometries between any pair of Hah1, MBD3, and MBD4 appear conserved as noted previously (Section 3.2). Relative to the extended conformation, MBD34$_{\text{ext}}$, the equilibrium stability constants of ^{1}MBD34 ($K_1^{34} = 0.6 \pm 0.2$) and ^{2}MBD34 ($K_2^{34} = 0.6 \pm 0.2$) can be obtained from the peak area ratios in the Cy3Cy5-MBD34$^{\text{L34}}$ E_{FRET} distribution (Figure 9D and Scheme 1). These two intramolecular complexes are approximately equal in stability.

To determine the six kinetic rate constants for the intramolecular-interdomain MBD3–MBD4 interactions (Scheme 1), we used the single-exponential decay constants obtained from the τ_{Low}, τ_{Mid}, and τ_{High} dwell-time distributions, $f_{\text{Low}}(\tau)$, $f_{\text{Mid}}(\tau)$, and $f_{\text{High}}(\tau)$, along with the relative transition frequencies to each state obtained from the Cy3Cy5-MBD34$^{\text{L34}}$ smFRET experiment (Figure S2E). Below, we present the previously derived dwell-time distribution functions, $f_{\text{Low}}(\tau)$, $f_{\text{Mid}}(\tau)$, and $f_{\text{High}}(\tau)$, for a 3-state equilibrium process (equations 1.1–1.3).[66,68,69]

$$f_{\text{Low}}(\tau) = (k_1 + k_2) e^{-(k_1 + k_2)\tau} \tag{1.1}$$

$$f_{\text{Mid}}(\tau) = (k_{-1} + k_3) e^{-(k_{-1} + k_3)\tau} \tag{1.2}$$

$$f_{\text{High}}(\tau) = (k_2 + k_{-3}) e^{-(k_2 + k_{-3})\tau} \tag{1.3}$$

37

Fitting the distributions of τ_{Low}, τ_{Mid}, and τ_{High} gives the values of the exponential terms.

The relationship between ratio of transition frequencies and the rate constants are as follows (equations 2.1–2.3):[66,68,69]

$$\frac{N_{\text{Low}\rightarrow\text{Mid}}}{N_{\text{Low}\rightarrow\text{High}}} = \frac{k_1}{k_2} \tag{2.1}$$

$$\frac{N_{\text{Mid}\rightarrow\text{Low}}}{N_{\text{Mid}\rightarrow\text{High}}} = \frac{k_{-1}}{k_3} \tag{2.2}$$

$$\frac{N_{\text{High}\rightarrow\text{Low}}}{N_{\text{High}\rightarrow\text{Mid}}} = \frac{k_1}{k_2} \tag{2.3}$$

where $N_{\text{x}\rightarrow\text{y}}$ represents the number of transitions between E_{FRET} states E_{x} and E_{y}. With six equations (1.1 – 2.3) and six unknowns (k_1, k_{-1}, k_2, k_{-2}, k_3, and k_{-3}) we can readily solve for all rate constants as reported in Scheme 1.

The intramolecular-interdomain association rate constants (k_1 and k_2), the dissociation rate constants (k_{-1} and k_{-2}), and the interconversion rate constants (k_3 and k_{-3}) all occur in similar timescales, ~ 1 s^{-1}. The rate constants also give the equilibrium constants, $K_1^{34} = 1.0 \pm 0.2$, K_2^{34} = 1.0 ± 0.1, and $K_{\text{int}}^{34} = 1.3 \pm 0.3$, consistent with those obtained from analyzing the Gaussian-resolved E_{FRET} distribution (Figure 9D).

In the presence of excess unlabeled Hah1, the three E_{FRET} states were still observed for Cy3Cy5-MBD34$^{\text{L34}}$, shown by the 2-D E_{FRET} analysis (Figure S3B) and the E_{FRET} distribution (Figure 9E). Therefore, the two intramolecular-interdomain complexes of MBD34 still occur in the presence of Hah1. However, the apparent stabilities of these three states have changed, reflected by the changes in the peak area ratios in the E_{FRET} distribution (Figure 9G), and so have the average dwell times of these three states (Table 1). These changes indicate that Hah1 interacts with MBD34 regardless of its two domains being in the extended conformation or

forming intramolecular-interdomain complexes. Particularly, the ability of Hah1 to interact with the intramolecular-interdomain complexes of MBD34 indicates that Hah1, MBD3 and MBD4 can come together to form 3-body interactions, consistent with the population overlap between the complexes observed in the Cy5-Hah1 + Cy3-MBD34^{L4} and Cy5-Hah1 + Cy3-MBD34^{L3} experiments (Section 3.3).

3.6. Cu^{+}-Dependence of Hah1–MBD34 Interactions.

We further investigated the Cu^{+}-dependence of Hah1–MBD34 interactions using the Cy5-Hah1 + Cy3-MBD34^{L4} labeling scheme, which directly probes Hah1–MBD4 interactions. Cu^{+}-transfer can only occur between Hah1 and MBD4 here because the CXXC Cu^{+}-binding motif of MBD3 was mutated to AXXA. The Cu^{+}-dependence of the isolated Hah1–MBD3 interactions should be minimal, as Hah1 has merely small conformational changes upon Cu^{+}-binding.[80]

Three apparent E_{FRET} states ($E_{\mathrm{Low}}^{\mathrm{L4}}$, $E_{\mathrm{Mid}}^{\mathrm{L4}}$, and $E_{\mathrm{High}}^{\mathrm{L4}}$) are still observed in the presence of 1, 2, and 4 eqv Cu^{+} (Figure S3A). Again, the apparent $E_{\mathrm{Low}}^{\mathrm{L4}}$ state contains contributions from the dissociated state, E_{Dissoc}, and Hah1 interactions with the unlabeled MBD3, $E_{\mathrm{Low}}^{\mathrm{L4}}{}'$. All three E_{FRET} values are similar to those in the absence of Cu^{+} (Figure 8C), indicating that within our experimental limit the Hah1–MBD4 interaction geometries remain largely unchanged by Cu^{+}.

Yet, the stabilities and dynamics of the Hah1–MBD4 interactions are dependent on the presence of Cu^{+}, as shown by the Gaussian-resolved E_{FRET} distributions and average lifetimes (Figure 11). In the presence of 1 eqv Cu^{+}, the $E_{\mathrm{Mid}}^{\mathrm{L4}}$/$E_{\mathrm{Dissoc}}$ and $E_{\mathrm{High}}^{\mathrm{L4}}$/$E_{\mathrm{Dissoc}}$ area ratios in the E_{FRET} distribution and the lifetimes of $E_{\mathrm{Mid}}^{\mathrm{L4}}$ and $E_{\mathrm{High}}^{\mathrm{L4}}$ do not change much compared with those of the *apo* protein interactions (Figure 11D, E). Therefore, Hah1 interactions with MBD4 are

39

unperturbed by 1 eqv Cu^+. This is in contrast to the Cu^+-dependence of Hah1–MBD4^{SD} interactions, which showed a factor of ~1.3 stabilization at 1 eqv Cu^+; this stabilization was attributable to Cu^+-bridging at the protein interaction interface.[4,6,7,25-28,70] This contrast indicates that the Cu^+-bridging-induced stabilization is insignificant for Hah1 interacting with the multi-domain construct MBD34, possibly because the *apo* Hah1–MBD4 interactions are already ~16 times more stable than the *apo* Hah1–MBD4^{SD} (Figure 9F). An NMR characterization of Hah1's interactions with a WDP multi-domain construct containing MBDs 4–6 also showed that Cu^+ does not greatly perturb Hah1–MBD interactions.[36]

In the presence of 2 eqv Cu^+, the E_{Mid}^{L4}/E_{Dissoc} and E_{High}^{L4}/E_{Dissoc} area ratios both decrease by a factor of ~2 compared with the *apo* and 1 eqv Cu^+ conditions (Figure 11D). This decrease in stability can be attributed to a decrease in the average lifetimes of the E_{Mid}^{L4} and E_{High}^{L4} states, $\overline{\tau}_{Mid}$ and $\overline{\tau}_{High}$, which also decrease by a factor of ~2 (Figure 11E). Although the change in the area ratios between 1 and 2 eqv Cu^+ is subtle, a *t*-test confirms that the values are different beyond error at the 95% confidence level. Moreover, the accompanying decrease in $\overline{\tau}_{Mid}$ and $\overline{\tau}_{High}$ by the same factor (~2) further validates our observation. No further change was observed in the E_{FRET} distribution (Figure 11C, D) or average lifetimes (Figure 11E) with 4 eqv Cu^+, indicating that both proteins are fully metallated at 2 eqv Cu^+.

The observed destabilization of the Hah1–MBD4 interactions at excess Cu^+ relative to the *apo* condition was not observed in the Hah1–MBD4^{SD} interactions.[70] Therefore, the destabilization must be associated with the multi-domain nature of MBD34 compared with the isolated MBD4^{SD}. It is possible that the full metallation disrupts the concerted actions between the two domains of MBD34 for interacting with Hah1.
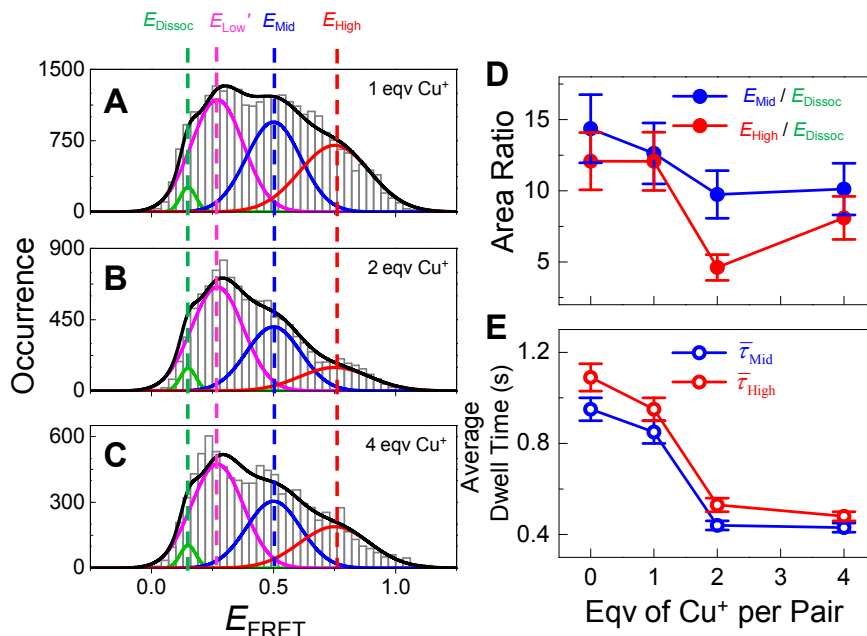
Figure 11. Compiled E_{FRET} distributions (bin size = 0.03) and Gaussian fits for Cy5-Hah1 + Cy3-MBD34^{L4} in the presence of (A) 1 eqv, (B) 2 eqv, and (C) 4 eqv of Cu$^+$ per protein pair. The E_{FRET} distributions were fitted globally with those in Figure 9 by sharing the center values and widths of each E_{FRET} state. The area ratio between $E_{Low}^{L4}{'}$ and E_{Dissoc}^{L4} was kept constant using the approximation that the isolated Hah1−MBD3 interactions (represented by $E_{Low}^{L4}{'}$) should be largely independent of Cu$^+$. (D) The area ratios of E_{Mid}^{L4} and E_{High}^{L4} with respect to E_{Dissoc}^{L4} for Cy5-Hah1 + Cy3-MBD34^{L4} with varying equivalents of Cu$^+$. (E) The average dwell times of E_{Mid}^{L4} and E_{High}^{L4}, $\overline{\tau}_{Mid}$ and $\overline{\tau}_{High}$ respectively, for Cy5-Hah1 + Cy3-MBD34^{L4} with varying equivalents of Cu$^+$.

CHAPTER 4

STRUCTURAL MODELS AND HAH1-MBD34 INTERACTION MECHANISM

We have used smFRET measurements combined with nanovesicle trapping to probe the weak and dynamic interactions between Hah1 and the double-domain MBD34 construct of WDP. By placing the FRET donor or acceptor on Hah1, MBD3, or MBD4 (Section 3.1), we have examined how Hah1 interacts with MBD3 and MBD4 and how MBD3 and MBD4 interact with each other at the single-molecule level. For all cases, we observed two major interaction complexes that interconvert dynamically (Section 3.2). The similarity in E_{FRET} values across all labeling schemes indicates that Hah1–MBD4, Hah1–MBD3, and MBD3–MBD4 interaction geometries are conserved, attributable to the structural and sequence homology among Hah1, MBD3, and MBD4.

In the absence of Cu^+, Hah1 interactions with MBD3 and MBD4 of MBD34 have similar stabilities (Section 3.3). The Hah1–MBD4 interactions are significantly more stable than the Hah1–MBD4^{SD} interactions; this enhanced stability is associated with an increase in the protein association rate, not a decrease in the dissociation rates (Section 3.4). An overlap population exists between Hah1–MBD3 and Hah1–MBD4 complexes, attributable to Hah1 interacting with MBD3 and MBD4 *simultaneously*, forming 3-body interactions. These 3-body interactions were further supported by intramolecular-interdomain MBD3–MBD4 complexes interacting with Hah1 (Section 3.5).

In the presence of Cu^+ and regardless of Cu^+ stoichiometry, the Hah1–MBD4 interaction geometries appear unchanged, although the stabilities and lifetimes of the interaction complexes decreased under excess Cu^+ (Section 3.6). This decrease in stability may be due to a disruption

of concerted interactions within the double-domain MBD34, as this trend was not observed for Hah1–MBD4^{SD} interactions.

Based on the above results, below we propose structural models for Hah1–MBD34 interactions. These models are then used to formulate a Hah1–MBD34 interaction mechanism that includes both 2-body and 3-body interactions between Hah1, MBD3, and MBD4.

4.1. Structural Models of 2-body Protein Interaction Complexes.

Two major interaction geometries, E_{Mid} and E_{High}, were observed for Hah1–MBD4, Hah1–MBD3, and MBD3–MBD4 interactions. To better understand how Hah1 and MBD34 interact, here we propose *possible* structural models of interaction complexes based on our smFRET results and past structural studies of these or homologous proteins. The experimental constraints for the structural models include: (1) The distance r between the donor and acceptor labeling sites in the complexes should be consistent with the observed E_{FRET} values of E_{Mid} and E_{High}. From the calibration curve measuring E_{FRET} as a function of r (Figure S1A), for $E_{Mid} = 0.50$, $r \sim 5 \pm 1$ nm, and for $E_{High} = 0.75$, $r \sim 3 \pm 1$ nm. (2) The stabilities of the interaction complexes should be comparable, as observed in our smFRET thermodynamic analysis (Chapter 3, Section 3.3). (3) The protein interaction interfaces for the complexes giving rise to E_{Mid} and E_{High} should be spatially distinct (i.e., no significant overlap) because our data indicate simultaneous 3-body interactions between Hah1, MBD3, and MBD4.

Hah1 and all WDP/MNK MBDs share the same βαββαβ protein fold and all contain the CXXC motif. The two α-helices are on one side of the protein (i.e., the "face" side) and the four β-strands form a β-sheet on the other side (i.e., the "back" side). In our cartoon representations of the interaction models we denote the "face" side of the protein with a green helix and the "back" side with a purple arrow (Figure 12).

43

In our previous study of Hah1−MBD4[SD] interactions,[70] we have proposed a structural model giving rise to the E_{Mid} state based on an interaction geometry between Hah1 and the 1st N-terminal MBD of MNK, MNK1, observed by NMR (Figure 3).[29] This interaction geometry was also observed between yeast homologs Atx1 and Ccc2a by NMR,[55] and in Hah1 dimers by X-ray crystallography.[25] In this geometry, Hah1 and MBD interact in a face-to-face manner: their CXXC motifs face each other, where a metal ion can coordinate to cysteines from both proteins, thus offering a facile pathway for metal transfer via ligand exchange.[4,6,7,25-27] We generated these face-to-face structural models between Hah1, MBD4, and MBD3 by superimposing the structures of MBD4 and MBD3[57] onto the Hah1−MNK1 structure (Figure 12A). We abbreviate these complexes as $H_{ff}4$, $H_{ff}3$, and $3_{ff}4$, where H stands for Hah1, 4 (or 3) stands for MBD4 (or MBD3), and the "ff" subscripts designate the face-to-face interaction mode of the two proteins/domains. In all three models, the donor-acceptor distances, r, range from ∼3.9–4.4 nm, giving expected E_{FRET} ∼0.54–0.63, within error of E_{Mid} (Table 2).

To further assess these face-to-face models, we examined their interaction stability using PISA (Protein Interfaces, Surfaces and Assemblies service at the European Bioinformatics Institute, http://www.ebi.ac.uk/pdbe/prot_int/pistart.html),[81] in comparison with the same analysis on the experimentally observed face-to-face Hah1−MNK1 interaction (Table 2). All the face-to-face models ($H_{ff}4$, $H_{ff}3$, and $3_{ff}4$) have comparable stability to the Hah1−MNK1 complex based on the buried surface area (BSA, ∼ 510–620 Å^2), solvation free energy change (−ΔG_{solv}, ∼ 5−11 kcal/mol), and favorable interactions between residues (H-bonding and salt bridges). Appendix, Section S4.1 describes the residues involved in these face-to-face models.
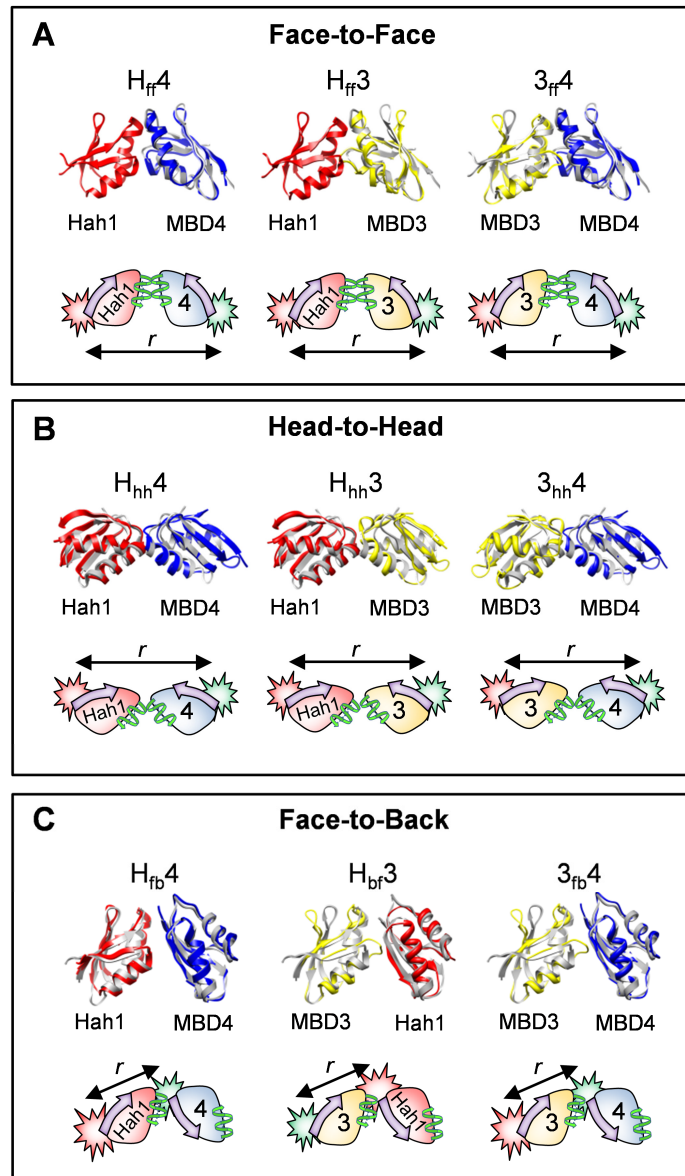
Figure 12. Structural models of (A) face-to-face, (B) head-to-head, and (C) face-to-back interaction models and corresponding cartoon representations showing the distance between the donor and acceptor, *r*. The models were generated by overlaying Hah1 (1FEE, red),[25] MBD3 (2ROP, yellow),[57] and MBD4 (2ROP, blue)[57] with the Hah1–MNK1 NMR structure (2K1R, grey)[29] for face-to-face (A), the Atx1 dimer crystal structure (2XMT, grey)[58] for head-to-head (B), and the Hma7 MBD dimer crystal structure (3DXS, grey)[59] for the face-to-back (C). In overlaying, the distance between the backbone α-carbons were minimized using the iterative magic fit function in SwissPDB Viewer,[82] followed by an energy minimization of the side-chains using the GROMOS 43B1 force field.[83]

Table 2. Analysis of Face-to-Face, Head-to-Head, and Face-to-Back Complexes

Structure	BSA (Å^2)[a]	ΔG_{solv} (kcal/mol)[b]	Favorable Interactions	r (nm)[c]	Expected E_{FRET}[d]	Observed E_{FRET}
Hah1– Mnk1[e]	539	−5.5	6 H-bonds, 5 Salt Bridges	4.77	0.49 ± 0.13	
H$_{ff}$4	510	−8.0	5 H-bonds, 1 Salt Bridge	4.44	0.54 ± 0.13	E_{Mid} = 0.50 ± 0.02
H$_{ff}$3	579	−6.2	6 H-bonds, 1 Salt Bridge	4.22	0.57 ± 0.13	
3$_{ff}$4	616	−11.2	1 H-Bond	3.86	0.63 ± 0.12	
Atx1 Dimer[f]	235	−7.8	1 H-bond	2.66	0.80 ± 0.09	
H$_{hh}$4	375	−7.8	1 H-bond	4.31	0.56 ± 0.13	E_{High} = 0.75 ± 0.02
H$_{hh}$3	166	−2.6	1 Salt Bridge	4.11	0.59 ± 0.13	
3$_{hh}$4	468	−10.5	2 H-Bonds	3.65	0.66 ± 0.12	
Hma7 Dimer[g]	370	−2.8	10 H-bonds, 1 Salt Bridge	3.28	0.73 ± 0.11	
H$_{fb}$4[#]	283	−2.9	3 H-bonds	3.91	0.62 ± 0.13	
H$_{bf}$4	240	−0.7	1 H-bond	3.01	0.75 ± 0.10	
H$_{fb}$3	179	+0.9	1 H-bond	4.20	0.58 ± 0.13	E_{High} = 0.75 ± 0.02
H$_{bf}$3[#]	334	−2.2	2 H-bonds	2.85	0.77 ± 0.10	
3$_{fb}$4[#]	426	−6.2	5 H-bonds	3.28	0.71 ± 0.11	
3$_{bf}$4	243	−0.4	1 H-bond	3.76	0.64 ± 0.12	

[a]Buried surface area. [b]Solvation free energy change. [c]Distance between Cy3 and Cy5 labeling sites. [d]Calculated using the E_{FRET} calibration curve (Figure S1A). [e]PDB code 2K1R. [f]PDB code 2XMT. [g]PDB code 3DXS. [#]Indicates the more favorable interaction complexes based on interface analysis.

For the structural model of protein interactions giving rise to E_{High}, there are two candidates from past studies on homologous proteins. One candidate is the crystal structure of a Atx1 dimer,[58] in which the two monomers interact "head-on" symmetrically in the α-helical and loop region next to the CXXC motif (Figure 12B and Appendix, Section S4.2); we refer to this interaction geometry as head-to-head. The other is the crystal structure of an asymmetric dimer of the MBD of Hma7, a Cu^+-transporting ATPase in *Arabidopsis thaliana*; this MBD is homologous to Hah1 and WDP/MNK MBDs.[59] In this Hma7 MBD dimer, the face side of one monomer interacts with the back side of the other monomer (Figure 12C and Appendix, Section S4.3); we refer to this interaction geometry as face-to-back.

Using the head-to-head structure of the Atx1 dimer, we generated the corresponding structural models of Hah1−MBD4, Hah1−MBD3, and MBD3−MBD4 interactions, abbreviated as $H_{hh}4$, $H_{hh}3$, and $3_{hh}4$, respectively (Figure 12B). The three models all have comparable stability to that of the Atx1 dimer (Table 2). However, the expected E_{FRET} is ~0.56–0.66 (Table 2), which is smaller than the experimentally observed E_{High} (~0.75). Furthermore, in all three cases, the interaction interface has a considerable overlap (~70%) with that of the face-to-face model (Appendix, Section S4.1 vs. S4.2); this overlap would forbid a coexistence of the head-to-head and face-to-face interactions in a 3-body interaction among Hah1, MBD4, and MBD3. Thus, the head-to-head interaction geometry is a less likely model for the E_{High} state, although we cannot definitively rule it out. Appendix, Section S4.2 describes the residues involved in the head-to-head interaction models.

To generate the face-to-back interaction models using the Hma7 MBD dimer structure, there are two possible combinations for each pair between Hah1, MBD4, and MBD3. In one combination, one protein presents its face side, and the other protein presents its back side. In the other combination the two proteins switch in position. This leads to a total of six different face-

47

to-back interaction models (three of which are presented in Figure 12C): $H_{fb}4$ or $H_{bf}4$, $H_{fb}3$ or $H_{bf}3$, and $3_{fb}4$ or $3_{bf}4$. The interaction stability analysis indicates that for each interaction pair, one combination is clearly more favorable than the other (Table 2). For example, the interface of the $H_{fb}4$ model has a larger buried surface area, more negative $\Delta G_{solv,}$, and more H-bonds than that of $H_{bf}4$. Similarly, $H_{bf}3$ is more favorable than $H_{fb}3$, and $3_{fb}4$ more favorable than $3_{bf}4$. The three more favorable combinations ($H_{fb}4$, $H_{bf}3$, and $3_{fb}4$) also have stabilities comparable to the experimentally observed Hma7 MBD dimer and the face-to-face models (Table 2). We thus chose these three as the likely face-to-back interaction models. Their expected E_{FRET} values (~0.62–0.77) are also consistent with the observed E_{High} (~0.75). Furthermore, the face and back sides do not overlap, allowing for possible coexistence of face-to-back and face-to-face interactions in a 3-body interaction complex (see Section 4.3). Appendix, Section S4.3 describes the residues involved in the face-to-back interactions.

4.2. Molecular Dynamics Simulations of Hah1–MBD4 Interaction Models

To further assess the face-to-face, face-to-back, and head-to-head interaction geometries, we used Hah1–MBD4 interactions as representatives and performed molecular dynamics simulations on the four possible models: $H_{ff}4$, $H_{fb}4$, $H_{bf}4$, and $H_{hh}4$. All simulations were performed by Linghao Zhong (Pennsylvania State University, Mont Alto). Computational details for the simulations are found in Appendix, Section S5. Figure 13A presents the center-of-mass (COM) distance between the two proteins as a function of simulation time to assess the inter-protein distances and relative stability/association of the Hah1–MBD4 interaction models. The distance between the donor and acceptor labeling sites (r, Figure 13B) vs. simulation time is also presented as this distance gives rise to our E_{FRET} observables. The $H_{ff}4$ inter-protein distance remains stable, ~25 Å, throughout the simulation, with an $r \sim 5$ nm corresponding to the E_{Mid}

state. This further supports the $H_{ff}4$ model as the geometry giving rise to the E_{Mid} state. At 6 ns, the $H_{fb}4$ complex undergoes a slight shift in inter-protein distance, from ~25 to ~28 Å, but remained stable through the rest of the simulation. Additionally, the $H_{fb}4$ complex retains an r of ~3 nm, which is consistent with the E_{High} state. The $H_{bf}4$ complex quickly dissociates as indicated by the sharp increase in inter-protein distance at 1.5 ns (>40 Å), which is consistent with our PISA interface analysis that indicated this complex was unfavorable (Section 4.1). For the $H_{hh}4$ model, its geometry is unstable, and its protein distance fluctuates in and out of association/dissociation before finally reaching a completely different geometry (denoted as "$H_{hh}4\_new$"). While the donor-acceptor distance, $r \sim 3$ nm, suggests this $H_{hh}4\_new$ geometry could be within the E_{High} state, no experimental evidence exists for this new geometry and it could correspond to a local minimum within the MD simulation. Regardless, the initial $H_{hh}4$ geometry did not remain intact and therefore is less likely a model giving rise to the E_{high} state, consistent with our previous assessment in Section 4.1.

The final geometries of $H_{ff}4$, $H_{fb}4$, and $H_{hh}4\_new$ and their interaction energies, E_{int}, are also presented in Figure 13C–D. For $H_{hh}4\_new$, $E_{int} = -52$ kcal/mol, which is much less than that of $H_{ff}4$ (−290 kcal/mol) or $H_{fb}4$ (−208 kcal/mol). In conclusion, our molecular dynamics simulations are in agreement with our PISA interface analysis (Section 4.1) and further support our proposal of the $H_{ff}4$ and $H_{fb}4$ models to the E_{Mid} and E_{High} observables.
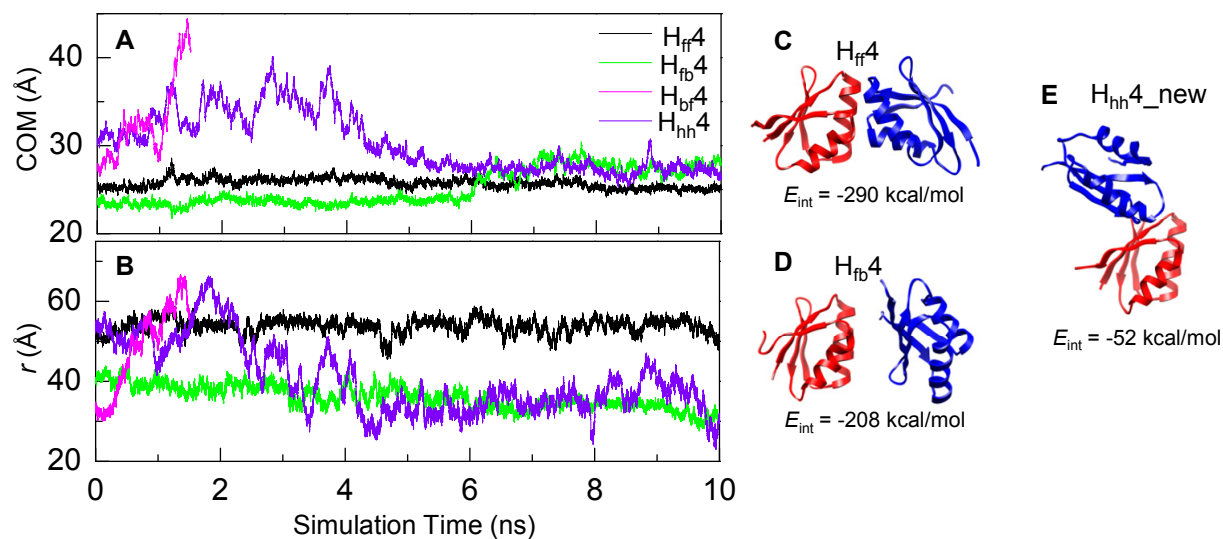
Figure 13. (A) The inter-protein distance (between the center-of-mass, COM, of each protein) and donor-acceptor distance, r, are plotted vs. molecular dynamics simulation time for $H_{ff}4$ (black), $H_{fb}4$ (green), $H_{bf}4$ (magenta), and $H_{hh}4$ (violet). The final structures (Hah1 in red and MBD4 in blue) and interaction energies, E_{int}, are shown for $H_{ff}4$ (C), $H_{fb}4$ (D), and "$H_{hh}4\_new$" (E). The $H_{hh}4\_new$ is distinct from the initial $H_{hh}4$ geometry. E_{int} is a rough estimation of the stability and does not directly correspond to the free energies of interaction (typically, ~few kcal/mol, as determined by PISA and previous MD simulations).[32]

4.3. Structural Models of 3-body Hah1–MBD34 Interactions.

Using the models of the 2-body interactions ($H_{ff}4$, $H_{ff}3$, $3_{ff}4$, $H_{fb}4$, $H_{bf}3$, and $3_{fb}4$, Figure 12A,C), four different 3-body interactions are possible: $H_{fb}4_{ff}3$, $H_{ff}4_{bf}3$, $3_{fb}H_{ff}4$, and $3_{fb}H_{fb}4$ (Figure 14). The first two involve Hah1 interacting with the two MBD3–MBD4 intramolecular complexes: for example in $H_{fb}4_{ff}3$, Hah1 interacts with MBD4 in a face-to-back geometry, while MBD4 and MBD3 are in a face-to-face geometry (Figure 14A). The latter two have Hah1 sandwiched between MBD3 and MBD4.

To check if the linkage between MBD4 and MBD3 would forbid these 3-body interactions, especially the two where Hah1 is sandwiched between the MBDs, we measured the distance between the C-terminal of MBD3 and the N-terminal of MBD4 in all 3-body interaction models (Table 3); the distances range from 2.1 to 5.8 nm, significantly shorter than the average linker length of 7.4 ± 0.7 nm from the NMR structure of MBD34 (Figure 7). In addition, Hah1 does not obstruct the space between the C-terminal of MBD3 and the N-terminal of MBD4 in these 3-body models. Therefore, the linker region of MBD34 should be long and flexible enough to accommodate the proposed 3-body interactions.

For the two proteins located on the outer-most side of the 3-body interaction, the distance between the donor and acceptor labeling sites is short enough (4.4–5.7 nm) to result in an apparent E_{Mid} state ($E_{FRET} \sim 0.4$–0.5) if these sites are labeled (Table 3). For the Cy5-Hah1 + Cy3-MBD34^{L3} labeling scheme, the $H_{fb}4_{ff}3$ (**d**) and $H_{ff}4_{bf}3$ (**e**) species would be within the E_{Mid}^{L3} state. For the Cy3Cy5-MBD34^{L34} experiment in the presence of unlabeled Hah1, the $3_{fb}H_{ff}4$ (**i**) and $3_{fb}H_{fb}4$ (**h**) species would be observed within the E_{Mid}^{L34} state.
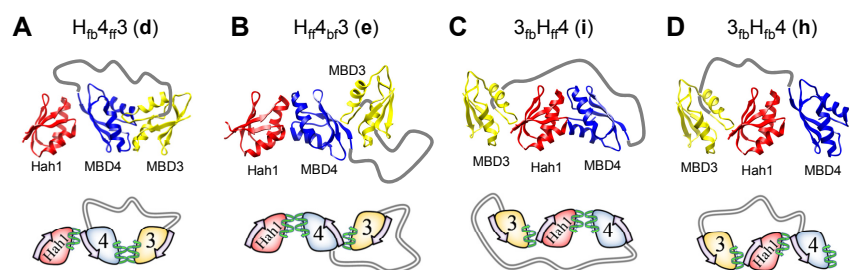
Figure 14. Illustration of 3-body Hah1 (red), MBD3 (yellow), MBD4 (blue) models derived from overlaying combinations of "Face-to-Face" and "Face-to-Back" interactions. The linker region is hand-drawn schematically in grey. (A) $H_{fb}4_{ff}3$ (**d**) and (B) $H_{ff}4_{bf}3$ (**e**) involve Hah1 interacting with MBD34 intramolecular-interdomain complexes, while (C) $3_{fb}H_{ff}4$ (**i**) and (D) $3_{fb}H_{fb}4$ (**h**) involve Hah1 sandwiched between MBD3 and MBD4. The cartoon representations are shown below each model.

Table 3. Hah1–MBD34 3-Body Model Dimensions

3-Body Model	MBD3(C-term)–MBD4(N-term) Distance (nm)[a]	Outer-most r (nm)[b]	Outer-most Expected E_{FRET}
$H_{fb}4_{ff}3$ (**d**)	4.12	$r_{H3} = 5.64$	0.37 ± 0.11
$H_{ff}4_{bf}3$ (**e**)	2.14	$r_{H3} = 5.68$	0.36 ± 0.11
$3_{fb}H_{ff}4$ (**i**)	5.81	$r_{34} = 4.44$	0.54 ± 0.13
$3_{fb}H_{fb}4$ (**h**)	3.80	$r_{34} = 4.59$	0.52 ± 0.13

[a]Distance between α-carbons of L93 and C124. [b]Distance between donor-acceptor labeling sites for outer-side proteins in 3-body interaction: r_{H3} represents the donor-acceptor distance between Hah1 and MBD3 in both $H_{fb}4_{ff}3$ and $H_{ff}4_{bf}3$. r_{34} represents the donor-acceptor distance between MBD3 and MBD4 in both $3_{fb}H_{ff}4$ and $3_{fb}H_{fb}4$. [c]Based on E_{FRET} calibration curve (Figure S1A).

4.4. Kinetic mechanism of Hah1–MBD34 Interactions.

Taking into account the dissociated state, the six possible 2-body interactions, and the four possible 3-body interactions, we can formulate a kinetic mechanism for Hah1−MBD34 interactions (Scheme 2), where each chemical species is indexed alphabetically (**a**)–(**k**). This mechanism includes the following major features that were observed experimentally: (1) intermolecular binding and unbinding of Hah1 with MBD4 or MBD3, each forming two interaction geometries, (2) intramolecular binding and unbinding between MBD4 and MBD3, forming two interaction geometries, (3) interconversion between interaction geometries, and (4) formation of 3-body interactions between Hah1, MBD4, and MBD3.
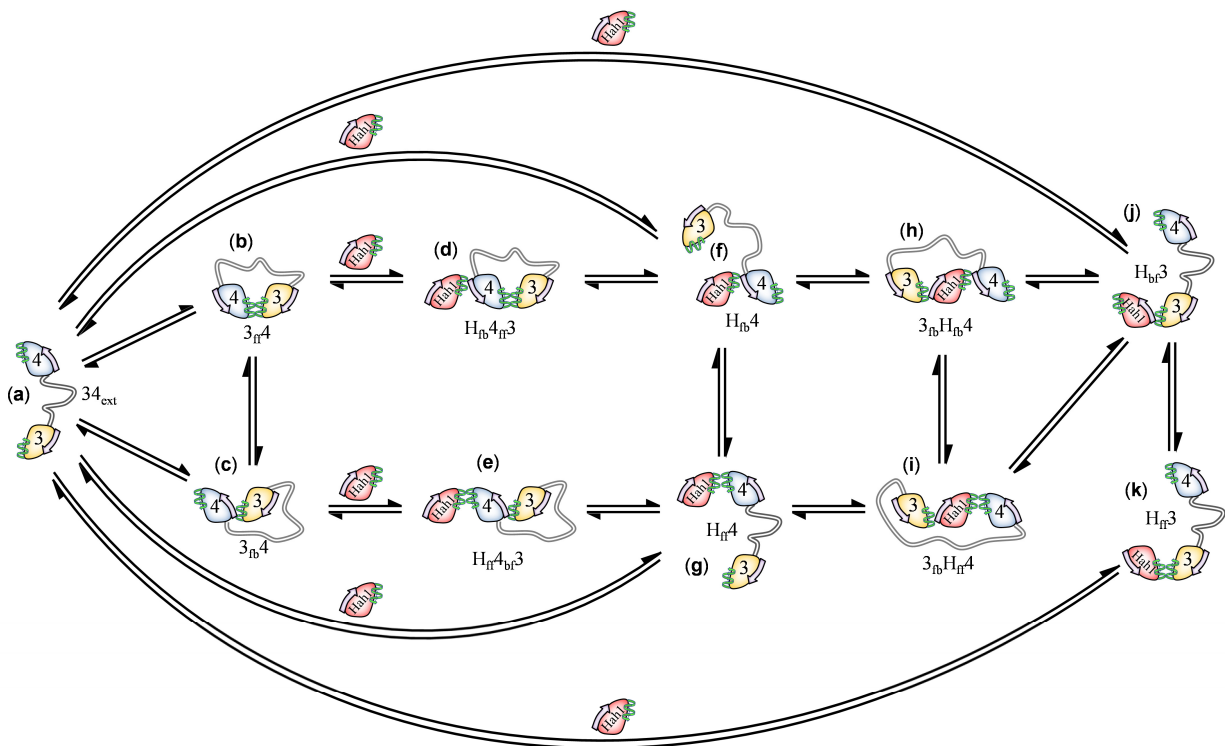
Scheme 2. Proposed Hah1−MBD34 interaction mechanism based on the structural models including 2-body Hah1−MBD3, Hah1−MBD4, and MBD3−MBD4 interactions in face-to-face or face-to-back orientation ($H_{ff}3$, $H_{bf}3$, $H_{ff}4$, $H_{fb}4$, $3_{ff}4$, and $3_{fb}4$) and 3-body interactions ($H_{fb}4_{ff}3$, $H_{ff}4_{bf}3$, $3_{fb}H_{fb}4$, and $3_{fb}H_{ff}4$).

The mechanism in Scheme 2 is certainly dependent on the structural models for the observed E_{Mid} and E_{High} states. The interaction geometries here are only models that are supported by data and deduced from known structures of protein complexes. Within either E_{Mid} or E_{High} states, additional subpopulations could exist that are unresolved in our measurements. The dynamic linker between MBD4 and MBD3 may also play a role in the complex formation.[33] Although the details of the mechanism of Hah1−MBD34 interactions may vary, the proposed mechanism here intends to convey the major features of their interactions within the limit of our measurements. Below we use this mechanism to provide a more quantitative description of, as well as functional insights into, Hah1−MBD34 interactions.

4.5. Relation of Chemical Species to E_{FRET} States.

Depending on where Cy3 and Cy5 are labeled on Hah1 and MBD34 (Figure 8A, D, and G), the species in Scheme 2 have different E_{FRET} values. Here we describe the assignment of these chemical species to each E_{FRET} state considering the labeling positions within our models.

The Hah1–MBD34 interaction mechanism along with the expected E_{FRET} observables for the Cy5-Hah1 + Cy3-MBD34^{L4} (Scheme 3), Cy5-Hah1 + Cy3-MBD34^{L3} (Scheme 4), and Cy3Cy5-MBD34^{L34} (Scheme 5) labeling schemes are presented for further clarification. Each chemical species is indexed alphabetically (**a**)–(**k**), as in Scheme 2, for convenient referencing in the subsequent discussion. Table 4 summarizes the assignments, along with the population percentages of each species calculated using the relative Gaussian areas and thermodynamic decomposition (see Section 4.6). The dissociated state, E_{Dissoc} (green), is the same regardless of the labeling scheme, as it always contains 34_{ext} (**a**), $3_{\text{ff}}4$ (**b**), and $3_{\text{fb}}4$ (**c**).
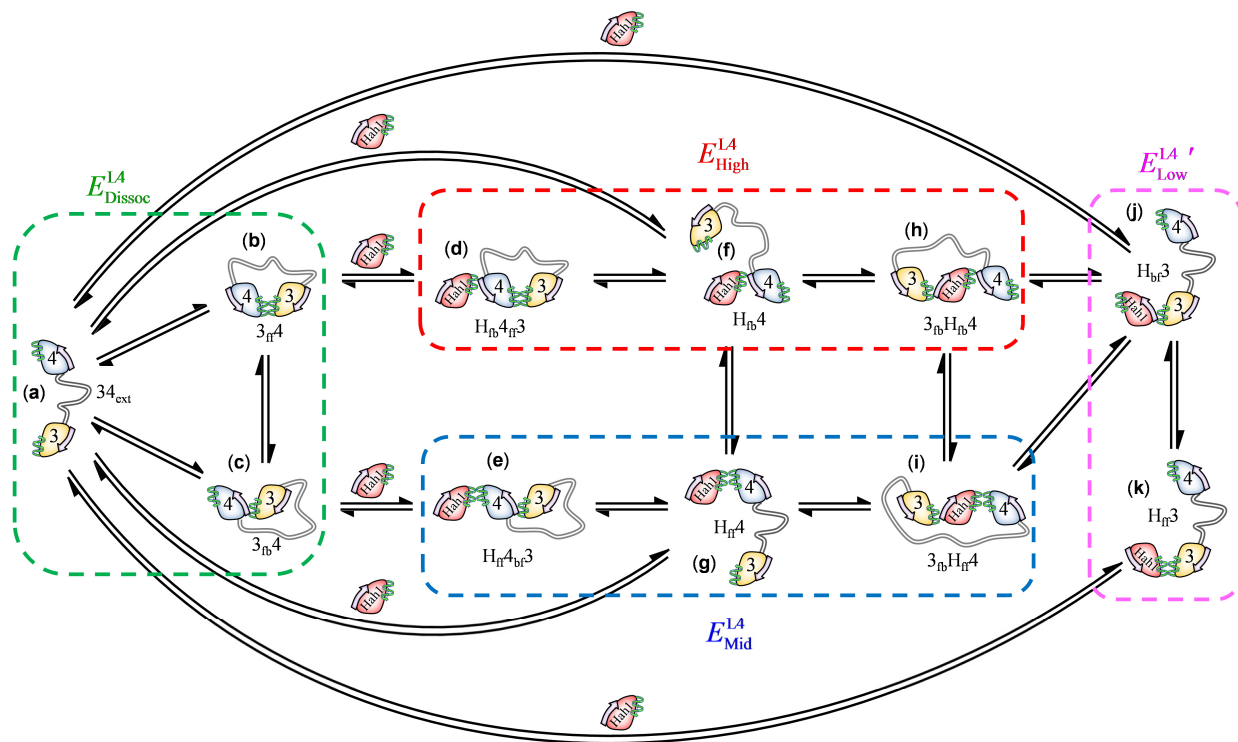
Scheme 3. Expected E_{FRET} observables for the Cy5-Hah1 + Cy3-MBD34^{L4} labeling scheme. The dashed boxes outline the species contained in each E_{FRET} state.

For Cy5-Hah1 + Cy3-MBD34^{L4} (Scheme 3), all Hah1–MBD4 face-to-face interaction geometries (H$_{ff}$4$_{bf}$3 (**e**), H$_{ff}$4 (**g**), and 3$_{fb}$H$_{ff}$4 (**i**)) occur within E_{Mid}^{L4} (blue), while all Hah1–MBD4 face-to-back interaction geometries (H$_{fb}$4$_{ff}$3 (**d**), H$_{fb}$4 (**f**), and 3$_{fb}$H$_{fb}$4 (**h**)) occur within E_{High}^{L4} (red). When Hah1 is interacting with the unlabeled MBD3 independently (H$_{bf}$3 (**j**) and H$_{ff}$3 (**k**)), this is observed within the $E_{Low}^{L4}{}'$ state (magenta).

56
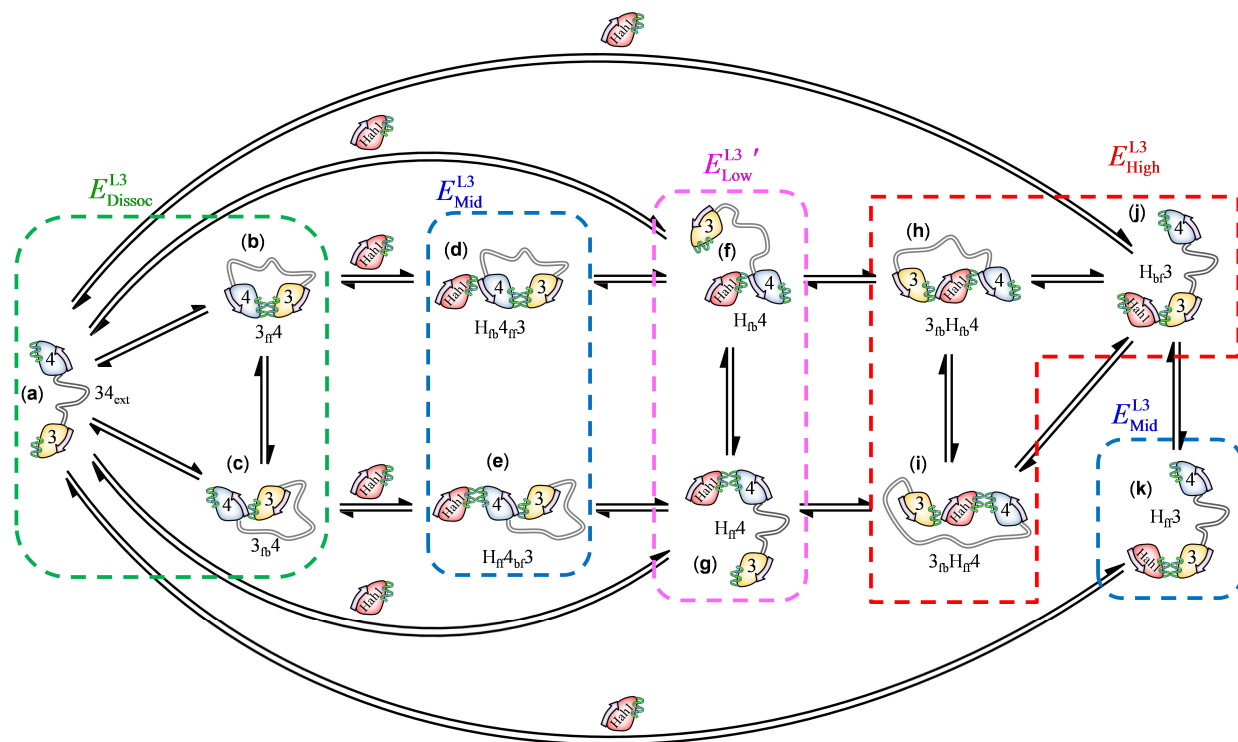
Scheme 4. Expected E_{FRET} observables for the Cy5-Hah1 + Cy3-MBD34^{L3} labeling scheme. The dashed boxes outline the species contained in each E_{FRET} state.

For Cy5-Hah1 + Cy3-MBD34^{L3} (Scheme 4), the Hah1–MBD3 face-to-face interaction geometry, $H_{ff}3$ (**k**), occurs within E_{Mid}^{L3} (blue), while all Hah1–MBD3 face-to-back geometries ($3_{fb}H_{fb}4$ (**h**), $3_{fb}H_{ff}4$ (**i**), and $H_{bf}3$ (**j**)) occur within E_{High}^{L3} (red). When Hah1 is interacting with the unlabeled MBD4 independently ($H_{fb}4$ (**f**) and $H_{ff}4$ (**g**)), this is observed within the $E_{Low}^{L3}{}'$ state (magenta). As described previously in Section 4.3, the $H_{fb}4_{ff}3$ (**d**) and $H_{ff}4_{bf}3$ (**e**) species should occur within the E_{Mid}^{L3} state (blue) due to the close proximity of Hah1 and MBD3.
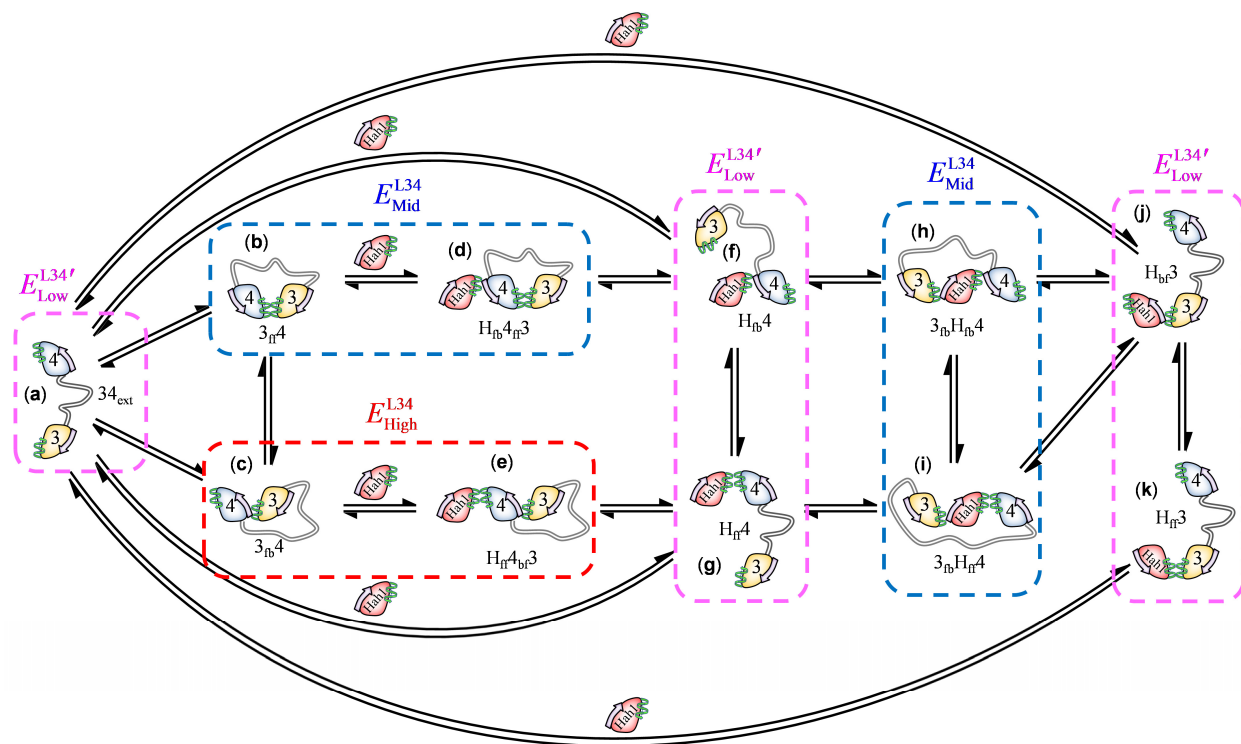
Scheme 5. Expected E_{FRET} observables for the Cy3Cy5-MBD34$^{\text{L34}}$ labeling scheme in the presence of one unlabeled Hah1 molecule. The dashed boxes outline the species contained in each E_{FRET} state.

For Cy3Cy5-MBD34$^{\text{L34}}$ in the presence of one unlabeled Hah1 molecule (Scheme 5), all MBD34$_{\text{ext}}$ conformations (34_{ext} (**a**), H$_{\text{fb}}$4 (**f**), H$_{\text{ff}}$4 (**g**), H$_{\text{bf}}$3 (**j**), and H$_{\text{ff}}$3 (**k**)) occur within $E_{\text{Low}}^{\text{L34}\prime}$ (magenta). The face-to-face MBD3–MBD4 interactions (3_{ff}4 (**b**) and H$_{\text{fb}}$4$_{\text{ff}}$3 (**d**)) give rise to the $E_{\text{Mid}}^{\text{L34}}$ state (blue), while the face-to-back MBD3–MBD4 interactions (3_{fb}4 (**c**) and H$_{\text{ff}}$4$_{\text{bf}}$3 (**e**)) give rise to the $E_{\text{High}}^{\text{L34}}$ state (red). The species where Hah1 is sandwiched in between MBD3 and MBD4 (3_{fb}H$_{\text{fb}}$4 (**h**) and 3_{fb}H$_{\text{ff}}$4 (**i**)) will also give rise to the $E_{\text{Mid}}^{\text{L34}}$ state (blue) due to the close proximity of MBD3 and MBD4, as indicated in Section 4.3.

58

Table 4. E_{FRET} Assignments of Chemical Species for Each Labeling Scheme and Their Population Percentages[a]

Chemical Species	Cy5-Hah1 + Cy3-MBD34^{L4}	Cy5-Hah1 + Cy3-MBD34^{L3}	Cy3Cy5-MBD34^{L34}	P (%)[a]
34_{ext} (**a**)	E_{Dissoc}	E_{Dissoc}	E_{Low}'	1.1 ± 0.2
$3_{ff}4$ (**b**)	E_{Dissoc}	E_{Dissoc}	E_{Mid}	0.6 ± 0.1
$3_{fb}4$ (**c**)	E_{Dissoc}	E_{Dissoc}	E_{High}	0.6 ± 0.1
$H_{ff}4$ (**g**)	E_{Mid}	E_{Low}'	E_{Low}'	25 ± 6
$H_{fb}4$ (**f**)	E_{High}	E_{Low}'	E_{Low}'	20 ± 5
$H_{ff}3$ (**k**)	E_{Low}'	E_{Mid}	E_{Low}'	18 ± 8
$H_{bf}3$ (**j**)	E_{Low}'	E_{High}	E_{Low}'	17 ± 8
$H_{fb}4_{ff}3$ (**d**)	E_{High}	E_{Mid}	E_{Mid}	4 ± 2
$H_{ff}4_{bf}3$ (**e**)	E_{Mid}	E_{Mid}	E_{High}	4 ± 2
$3_{fb}H_{fb}4$ (**h**)	E_{High}	E_{High}	E_{Mid}	4 ± 2
$3_{fb}H_{ff}4$ (**i**)	E_{Mid}	E_{High}	E_{Mid}	4 ± 2

[a]Population percentages of chemical species obtained as described in Section 4.6.

4.6. *Thermodynamic Decomposition and Quantification of Hah1–MBD34 Interactions.*

With the E_{FRET} assignments of the chemical species (**a**)–(**k**) for each labeling scheme (Scheme 3, Scheme 4, Scheme 5, and Table 4) and the relative stabilities of the E_{FRET} states from the E_{FRET} distributions (Figure 9B-D), we can now decompose the population percentages of these species (i.e., the relative occurrence of each species at any time point among many interacting pairs) and examine their thermodynamic properties.

From the Gaussian-resolved E_{FRET} distributions (Figure 9B-D), the area percentage (χ) of each E_{FRET} state represents the sum of the population percentages (P) of all the chemical species contributing to that E_{FRET} state. For example, in the Cy5-Hah1 + Cy3–MBD34^{L4} labeling scheme

(Scheme 3), the area percentage ($\chi_{\text{Dissoc}}^{\text{L4}}$) of the $E_{\text{Dissoc}}^{\text{L4}}$ state equals the sum of population percentages of 34_{ext} (**a**), $3_{\text{ff}}4$ (**b**), and $3_{\text{fb}}4$ (**c**) that constitute the dissociated state:

$$\chi_{\text{Dissoc}}^{\text{L4}} = P_a + P_b + P_c = P_{\text{Dissoc}} \tag{3.1}$$

Here, P_x denotes the population percentage of chemical species **x**. For the area percentages of the $E_{\text{Low}}^{\text{L4}}{}'$, $E_{\text{Mid}}^{\text{L4}}$, and $E_{\text{High}}^{\text{L4}}$ states,

$$\chi_{\text{Low}}^{\text{L4}}{}' = P_k + P_j \tag{3.2}$$

$$\chi_{\text{Mid}}^{\text{L4}} = P_g + P_e + P_i \approx P_g + 2P_S \tag{3.3}$$

$$\chi_{\text{High}}^{\text{L4}} = P_f + P_d + P_h \approx P_f + 2P_S \tag{3.4}$$

and $\chi_{\text{Dissoc}}^{\text{L4}} + \chi_{\text{Low}}^{\text{L4}}{}' + \chi_{\text{Mid}}^{\text{L4}} + \chi_{\text{High}}^{\text{L4}} = 1$.

Similarly, for the Cy5-Hah1 + Cy3-MBD34$^{\text{L3}}$ labeling scheme (Scheme 4):

$$\chi_{\text{Dissoc}}^{\text{L3}} = P_a + P_b + P_c = P_{\text{Dissoc}} \tag{4.1}$$

$$\chi_{\text{Low}}^{\text{L3}}{}' = P_g + P_f \tag{4.2}$$

$$\chi_{\text{Mid}}^{\text{L3}} = P_k + P_d + P_e \approx P_g + 2P_S \tag{4.3}$$

$$\chi_{\text{High}}^{\text{L3}} = P_j + P_h + P_i \approx P_j + 2P_S \tag{4.4}$$

and $\chi_{\text{Dissoc}}^{\text{L3}} + \chi_{\text{Low}}^{\text{L3}}{}' + \chi_{\text{Mid}}^{\text{L3}} + \chi_{\text{High}}^{\text{L3}} = 1$. Note, $\chi_{\text{Dissoc}}^{\text{L3}} = \chi_{\text{Dissoc}}^{\text{L4}}$, as the dissociated state is the same for these two labeling schemes; this relationship was also used in constraining the global Gaussian-resolution of the E_{FRET} distributions in Figure 9B and C.

The chemical species 34_{ext} (**a**), $3_{\text{ff}}4$ (**b**), and $3_{\text{fb}}4$ (**c**) are resolved in the labeling scheme Cy3Cy5-MBD34$^{\text{L34}}$ and correspond to the $E_{\text{Low}}^{\text{L34}}{}'$, $E_{\text{Mid}}^{\text{L34}}$, and $E_{\text{High}}^{\text{L34}}$ states, respectively (Figure 9D). The area percentages in Figure 9D are related to the population percentages by:

$$\chi_{\text{Low}}^{\text{L34}'} = P_\text{a}/P_{\text{Dissoc}} \qquad\qquad (5.1)$$

$$\chi_{\text{Mid}}^{\text{L34}} = P_\text{b}/P_{\text{Dissoc}} \qquad\qquad (5.2)$$

$$\chi_{\text{High}}^{\text{L34}} = P_\text{c}/P_{\text{Dissoc}} \qquad\qquad (5.3)$$

where $P_{\text{Dissoc}} = P_\text{a} + P_\text{b} + P_\text{c}$ as in equations 3.1 and 4.1.

Unfortunately the above equations are underdetermined for solving for the population percentage of every species in Scheme 2. We thus approximated that all 3-body interaction species have similar stabilities; i.e., their population percentages are approximately the same: $P_\text{d} \approx P_\text{e} \approx P_\text{h} \approx P_\text{i} \equiv P_\text{S}$. This should be a valid approximation because the face-to-face and face-to-back interactions (and thus any combination thereof) have comparable stabilities (Chapter 3, Section 3.3). This approximation leads to the simplification of equations 3.3, 3.4, 4.3, and 4.4 and allows for solving for the population percentages of the chemical species (Table 4). Also, because of this approximation, P_S effectively represents the average population percentage of all 3-body interaction species.

Using the population percentage of each species and the effective concentration (C_{eff}) of each protein inside the nanovesicle ($\sim 6 \pm 5$ μM),[70] we can compute the equilibrium constants of the kinetic processes in Scheme 2. These equilibrium constants can be classified into four categories (Table 5): (1) Dissociation constants, K_D, of Hah1 binding to either MBD4 or MBD3 (e.g., K_D^f, K_D^g, K_D^j, K_D^k, K_D^d, and K_D^e), (2) stability constants, K_A, of intramolecular-interdomain interactions (e.g., K_A^b and K_A^c), (3) interconversion equilibrium constants, K_{int}, between different interaction geometries (e.g., $K_{\text{int}}^{\text{H4}}$, $K_{\text{int}}^{\text{H3}}$, and K_{int}^{34}), and (4) stability constants, K_A, of the 3-body interactions relative to the 2-body interactions (e.g., $K_\text{A}^{\text{S/f}}$, $K_\text{A}^{\text{S/g}}$, and $K_\text{A}^{\text{S/j}}$).

The dissociation constants (K_D) range from 10^{-1} to 10^0 μM, in agreement with the SPR studies,[52,53] and an order of magnitude smaller than the Hah1−MBD4^{SD} interactions. This likely comes from an increased rate of association (see Chapter 3, Section 3.4), perhaps because of additional electrostatic/orientational steering present in the double-domain MBD34 for interacting with Hah1.[84] The interconversion equilibrium constants (K_{int}) are around 1, consistent with the two Hah1−MBD interaction geometries having comparable stabilities. From the stability constants ($K_A \sim 0.2$), the 3-body interactions are less stable than the 2-body interactions, perhaps due to entropy loss in forming 3-body interactions.

Table 5. Equilibrium Constants for the Hah1–MBD34 Interaction Mechanism in Scheme 2[a]

Dissociation Constant and Definition[b]	Value (μM)	Equilibrium Constant and Definition	Value
$K_D^{\,f} = \left(P_a/P_f\right) C_{eff}$	0.34 ± 0.28	$K_{int}^{H4} = P_g \big/ P_f$	1.3 ± 0.4
$K_D^{\,g} = \left(P_a/P_g\right) C_{eff}$	0.27 ± 0.22	$K_{int}^{H3} = P_k \big/ P_j$	1.1 ± 0.7
		$K_{int}^{34} = P_b \big/ P_c$	1.0 ± 0.3
$K_D^{\,j} = \left(P_a/P_j\right) C_{eff}$	0.41 ± 0.37	$K_A^{\,b} = P_b \big/ P_a$	0.56 ± 0.15
$K_D^{\,k} = \left(P_a/P_k\right) C_{eff}$	0.37 ± 0.33	$K_A^{\,c} = P_c \big/ P_a$	0.56 ± 0.15
		$K_A^{S/f} = P_S \big/ P_f$	0.21 ± 0.13
$K_D^{\,d} = \left(P_b/P_d\right) C_{eff}$	0.92 ± 0.89	$K_A^{S/g} = P_S \big/ P_g$	0.16 ± 0.10
$K_D^{\,e} = \left(P_c/P_e\right) C_{eff}$	0.92 ± 0.89	$K_A^{S/j} = P_S \big/ P_j$	0.25 ± 0.18

[a] P_x represents the relative population of chemical species x as annotated in Scheme 2. [b] C_{eff} is the effective protein concentration inside the nanovesicle, 6 ± 5 μM, calculated using the vesicle inner-diameter, 80 ± 20 nm.[70]

4.7. Functional Implications of Hah1–MBD34 Interactions.

Inside cells Hah1 delivers Cu^+ to WDP (or MNK), which translocates Cu^+ across the Golgi membrane for either incorporation to downstream copper proteins or efflux. Many features of the interactions between Hah1 and the double-domain WDP construct MBD34 are advantageous for fulfilling this function. These advantages are better illustrated by an analogy to the modern cargo transportation industry. In this analogy, Hah1 behaves like a delivery truck with Cu^+ as its cargo. The N-terminal region of WDP, with multiple MBDs, behaves like a warehouse distribution center, where the MBDs are the loading docks. The distribution center must operate with both efficiency and versatility to receive, re-route, and export shipments from many trucks.

The operation versatility of the WDP distribution center is accomplished by providing multiple MBD docking sites for the Hah1 truck to deliver its cargo. The truck can park frontward or backward at the dock (i.e., with two Hah1–MBD interaction geometries). Even better, the truck can interconvert between its docking geometries dynamically, thus allowing either of the two interfaces to be exposed for interaction with an additional MBD. The 3-body interactions where Hah1 is sandwiched between MBDs allow for the re-routing of the delivery truck, i.e., a Hah1 molecule can be handed over directly from one MBD to another. This re-routing of Hah1 would especially be useful when the initially targeted MBD is already loaded with Cu^+. Consistent with this scenario, a decrease in complex stability and lifetime was observed when Hah1 and MBD34 are fully metallated (Chapter 3, Section 3.6), assisting the departure of Hah1.

WDP's intramolecular MBD–MBD interactions provide a way for internal redistribution of the Cu^+ cargo, either to vacate space for next Hah1 delivery or to traffic Cu^+ downstream. This redistribution also occurs in a versatile manner, as multiple binding geometries were observed between MBD3 and MBD4. This internal cargo redistribution among MBDs can be directly

coupled to the cargo delivery or export through Hah1 interactions with the intramolecular MBD–MBD complexes.

All of these processes occur on similar timescales (~1 s), including the protein associations at the effective μM concentration inside nanovesicles. (Note the intracellular concentration of the yeast Hah1 homologue Atx1 is also about μM.[85]) Their similarity in timescale suggests that all these processes should occur comparably inside cells for function.

Our proposed Hah1–MBD34 interaction mechanism may also help understand the regulatory function of the MBDs, where Hah1–MBD or MBD–MBD interactions modulate the ATPase activity associated with Cu^+-translocation[13,18,40,44,45] or the kinase-mediated phosphorylation associated with the relocalization of WDP/MNK for Cu^+-efflux.[13,39,46-50] It was proposed that large-scale conformational changes within the N-terminal tail of WDP/MNK act as the regulatory switch,[30,33,35,41-43] which disrupts MBD interactions with the catalytic core affecting Cu^+-translocation or exposes/hides phosphorylation sites in the linker regions. The 3-body interactions where Hah1 is sandwiched between MBDs could induce large-scale conformational changes in the cytoplasmic tail of WDP, and hence may play a role in this regulatory switching mechanism.

CHAPTER 5

SUMMARY AND FUTURE DIRECTIONS

5.1. What We Learned

We have systematically characterized the interaction dynamics of Hah1 with a multi-domain construct of WDP, MBD34 using smFRET. Two conserved, dynamically interconverting, protein interaction geometries were observed between Hah1–MBD3, Hah1–MBD4, and MBD3–MBD4. Hah1 interacts with MBD3 and MBD4 with a similar stability within MBD34, and may even interact with both MBDs simultaneously by forming 3-body interactions with a relative occurrence of ~16%. Compared with an isolated MBD4 (MBD4^{SD}), Hah1−MBD4 interactions within the multi-domain construct are more stable by an order of magnitude, suggesting that concerted interactions facilitate complex formation. The kinetics and thermodynamics of the intramolecular-interdomain MBD3–MBD4 interactions were quantified. We further studied the Cu$^+$-dependence of Hah1–MBD4 interactions and found they become destabilized at excess Cu$^+$ concentrations.

We further proposed models for Hah1−MBD34 interactions, including 2-body and 3-body interactions, and formulated a kinetic mechanism that includes all our experimentally observed interaction processes. We were able to quantify the equilibria between different chemical species in our proposed interaction mechanism by using the relative area percentages obtained from the Gaussian-resolved E_{FRET} distributions. The Hah1-MBD34 interaction dynamics suggest an efficient and versatile means for WDP (or MNK) to receive, re-route, and export Cu$^+$ that is delivered by Hah1, analogous to a modern-day warehouse distribution center. The simultaneous 3-body interactions between Hah1, MBD3, and MBD4 may also induce large-scale conformational changes in the N-terminal tail which could serve as a regulatory switch for

66

WDP/MNK function.

5.2. Future Directions

5.2.1. Further Interrogation of Hah1–MBD and MBD–MBD Interactions.

Due to the complexity of the Hah1–MBD and MBD–MBD interaction dynamics within the full WDP system, there are many variations of future smFRET experiments that could further interrogate this multi-bodied interaction network. Ideally, these experiments should build on our previous and current studies while providing further functional insight into the mechanisms of Cu^+ trafficking in the cell.

To further interrogate how intramolecular-interdomain MBD3–MBD4 interactions are coupled to intermolecular interactions with Hah1, we can immobilize the Cy3Cy5-MBD34^{L34} construct, modified with an Avitag (a sequence of amino acids that binds to biotin) at the N-terminus, on a biotinylated PEG-coated surface. Unlabeled Hah1 can then be flowed at varying concentrations to monitor the perturbation in the MBD3–MBD4 dynamics via smFRET (Figure 15A). We could also probe all three interactions simultaneously using 3-color smFRET.[86-88] This would involve cotrapping Cy3-Hah1 and MBD34^{L34} labeled with two acceptors, Cy5 and Cy5.5 (or Cy7), within the nanovesicle (Figure 15B). Such an experiment, although technically challenging due to substantial overlap in the fluorophore emission spectra, would directly confirm the presence of our proposed 3-body interactions between Hah1, MBD3, and MBD4.

We could also investigate other MBD–MBD interaction dynamics within the WDP system using smFRET. Because we have already probed intramolecular interactions between MBD4 and the adjacent MBD3, it would be interesting to look at MBD4's interactions with a more distant MBD; e.g., MBD2 or MBD1. We already have protein constructs containing MBDs 1–4 and MBDs 1–6 which could be easily modified to introduce the necessary labeling

cysteines (Figure 15C–E). A systematic comparison of MBD4 interactions with MBD3, MBD2, and MBD1 would answer whether the proximity (in primary sequence) between two MBD's affects their ability to interact with one another. Combining the information from these various smFRET experiments may also yield information on the overall tertiary structure of the N-terminal tail, and how this might change in the presence of Cu^+ or Hah1. This would give direct evidence for the proposed large-scale conformational changes within the N-terminal tail of WDP/MNK which could serve as a regulatory switch for cellular relocalization or Cu^+ translocation. [30,33,35,41-43] Various combinations of the previously mentioned 3-color smFRET labeling scheme could also be implemented to study the interaction of Hah1 with these larger multi-MBD systems.

Ultimately, to test whether Hah1 interacting with each MBD in multiple geometries and/or forming 3-body interactions is directly related to WDP/MNK function (i.e., Cu^+ translocation or relocalization for Cu^+ efflux), we would need to study the system *in vivo*. Multiple Hah1–MBD interaction geometries and 3-body interactions would be eliminated or disrupted by destabilizing the face-to-back interactions via mutagenesis. For example, S124 of MBD4, located on the back side, could be mutated to alanine to disrupt hydrogen bonding in the $H_{fb}4$ interaction (Figure S10), while not effecting the $H_{ff}4$ interaction. Immunofluorescence microscopy can then be used to determine whether WDP retains its ability to relocalize for Cu^+ export compared to the wild-type.[39] Cellular vitality under conditions of Cu^+ stress could also be determined by yeast complimentation assays.[37,38] Additionally, the catalytic phosphorylation of WDP could be monitored readily using standard biochemical procedures.[30,51]
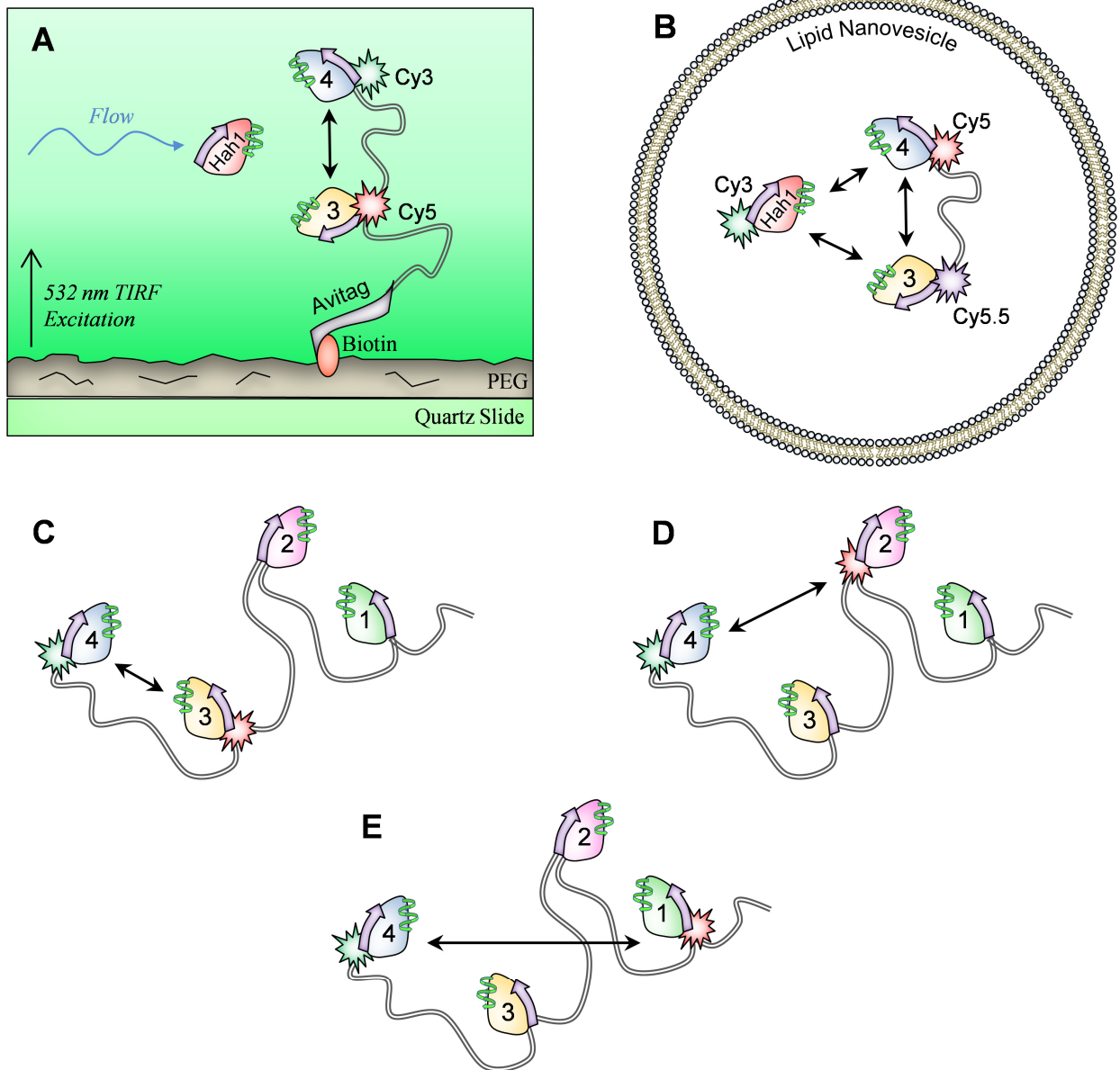
Figure 15. (A) Cy3Cy5-MBD34^{L34} immobilized on a PEG-coated slide using Avitag-biotin interactions. Unlabeled Hah1 can be flowed at various concentrations to systematically examine the perturbation in MBD3–MBD4 interactions. (B) Three-color smFRET labeling scheme used to simultaneously probe Hah1–MBD4, Hah1–MBD3, and MBD3–MBD4 interactions. Labeling schemes to probe MBD4 interactions with (C) MBD3, (D) MBD2, and (E) MBD1 within a WDP construct containing MBDs 1–4.

69

5.2.2. Beyond Cu$^+$: Exploring Vitamin B$_{12}$ Trafficking.

Besides Cu$^+$, nature must also deliver many other potentially toxic, yet essential, cofactors to their cellular destinations.[89] For example, the chaperone adenosyltransferase (ATR) binds and delivers adenosylcobalamin (AdoCbl), a Vitamin B$_{12}$ derivative, to methylmalonyl-CoA mutase (MCM).[90,91] In addition to AdoCbl, other cobalamin derivatives exist depending on the upper-axial ligand coordinated to cobalt (Figure 16A); these include cyanocobalamin (CNCbl), hydroxylcobalamin (HOCbl), and methylcobalamin (MeCbl).[92] Only AdoCbl and MeCbl are the biologically active forms of the cofactor. For instance, MCM utilizes the labile organometallic bond in AdoCbl to isomerize methylmalonyl-CoA to succinyl-CoA, for use in the citric acid cycle.[92]

In collaboration with Prof. Ruma Banerjee and Michael Lofgren at the University of Michigan, we seek to first quantify the binding/unbinding dynamics of various cobalamin cofactors with ATR and, ultimately, understand how ATR's interactions with MCM are coupled to cofactor transfer at the single-molecule level. Because the various cobalamin derivatives exhibit visible absorption bands around ~460 nm, we can monitor their binding/unbinding kinetics to ATR in real time using single-molecule fluorescence quenching (smFQ).

ATR is a homotrimer, with each monomer consisting of a five-helix bundle, and three cobalamin binding sites located in between each monomer (Figure 16B).[93-95] For our smFQ experiments, we took advantage of ATR's trimeric nature. While the introduction of a labeling cysteine for maleimide conjugation produces three labeling sites, one or two of these sites can be conjugated to a maleimide-PEG-biotin linker for immobilization, and another can be labeled with the maleimide fluorescent reporter (Figure 16C). Because ATR contains three binding sites (with different distances from the dye) and can potentially bind two or even three substrates simultaneously, we expect to see multiple smFQ states.
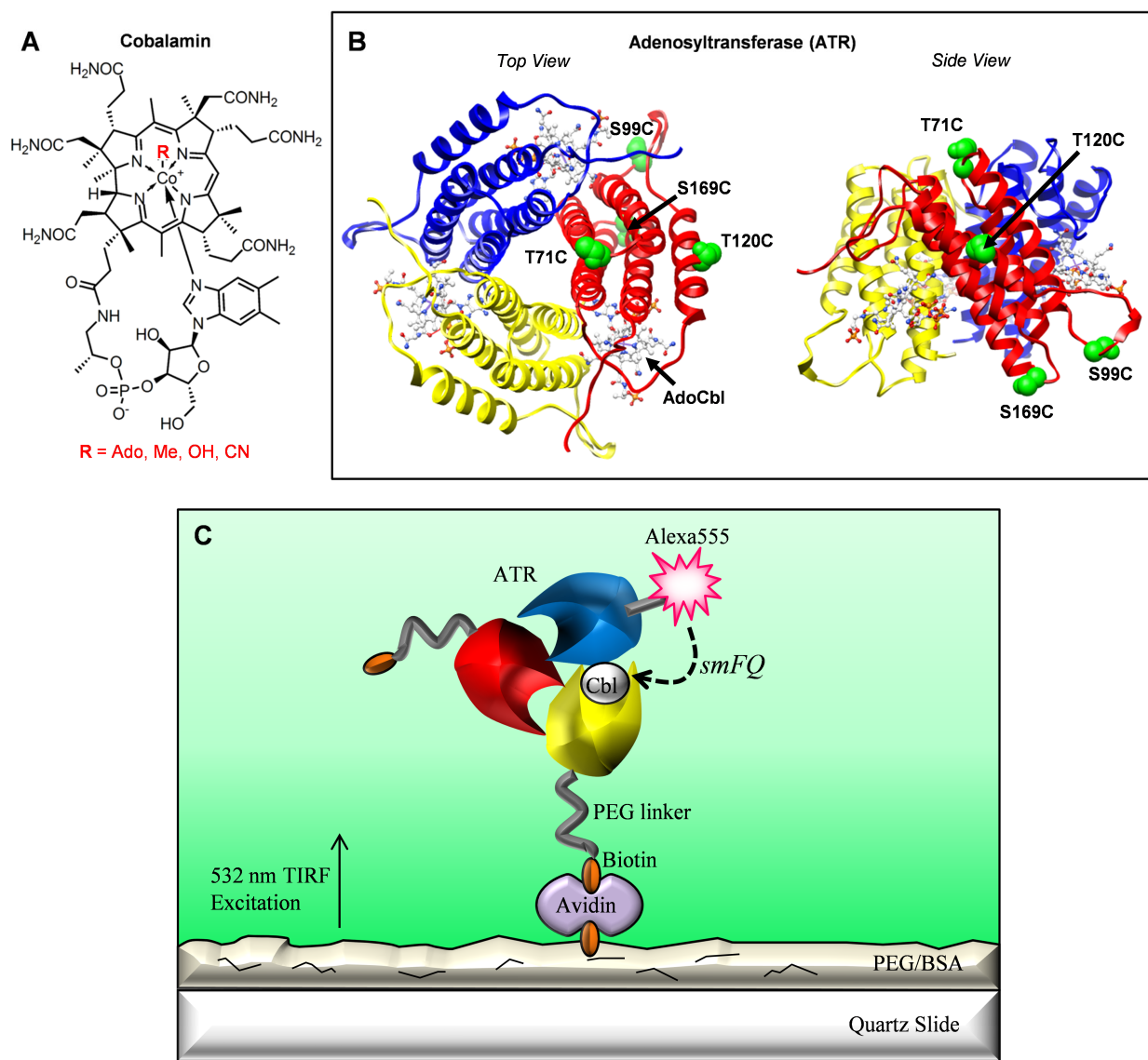
Figure 16. (A) Structure of cobalamin (Cbl). The upper axial ligand, R, can be an adenosyl (Ado), methyl (Me), hydroxo (OH), or cyano (CN) group. (B) Top and side views of the *M. extorquens* ATR homology model, based on the structure of *M. tuberculosis* ATR (PDB code 2G2D).[93] The homology model was overlayed with the *L. reuteri* ATR structure (PDB code 3CI1)[94] to model the AdoCbl binding sites. The positions of the potential labeling sites, T71C, S99C, T120C, and S169C, are shown in green. (C) smFQ experimental set-up. One labeling site is conjugated via maleimide chemistry to Alexa555, while the other two sites are conjugated to a PEG linker containing a biotin group. The labeled ATR is immobilized on a PEG- or BSA-coated quartz slide, and single Alexa555 fluorophores are excited using 532-nm TIRF. Flowing various concentrations of Cbl results in dynamic binding/unbinding events, reported by smFQ.

To determine the appropriate labeling sites (i.e., residues to be mutated to cysteine for maleimide conjugation to the dye or PEG-biotin), we considered a variety of factors: (1) the residue was not conserved throughout various ATR homologues, (2) the residue was not involved in stabilizing secondary or tertiary structure, (3) the residue was similar to cysteine, (4) the residue was solvent accessible for maleimide labeling, and (5) the residue was not located too close to the cobalamin binding site (which we also assume constitutes the ATR–MCM interface). We also mutated two of the native cysteines to serines (C64S and C154S) to avoid potential additional labeling at these sites (site-directed mutagenesis was performed by Michael Lofgren). After these considerations, we chose four potential labeling sites: T71C, T120C, S99C, and S169C (Figure 16B). While all constructs retained their ability to bind and release AdoCbl (activity assay performed by Michael Lofgren), the T120C and S169C mutants were susceptible to aggregation at room temperature. We therefore chose the T71C and S99C ATR constructs for our smFQ experiments.

After labeling the ATR constructs with Alexa555 and PEG-biotin, we performed smFQ experiments (Figure 16C) with various concentrations of AdoCbl, MeCbl, HOCbl, and CNCbl. Figure 17A illustrates a typical smFQ trace for the ATR T71C construct in the presence of AdoCbl. Highly dynamic binding and unbinding events are observed with multiple smFQ states that are clearly distinguishable from dye blinking (such blinking events would result in intensity values similar to the photobleached state, ~0). To ensure these additional states were not due to collision-induced quenching of the dye by the substrate, control experiments performed by Tai-Yen Chen showed that these intensity fluctuations do not occur upon addition of AdoCbl to Alexa555-labeled DNA at the single-molecule level. The highest intensity state represents the *apo* protein, while the lower intensity states represent quenching due to cobalamin binding to one or more of the three sites. The normalized intensity distributions compiled for many ATR

molecules (Figure 17B) show the population of the lower intensity states increases with increasing cobalamin concentration, as expected. These distributions also indicate the presence of multiple smFQ states, which can be attributed to the presence of multiple binding sites with different distances from the dye reporter and various cofactor binding stoichiometries (single, double, and triple-bound species).

We are currently deconvoluting the number of smFQ states present for each condition, possibly by using 2-dimensional smFQ state analysis (similar to that used in our smFRET study) or perhaps hidden Markov modeling.[96,97] Once the number of states is determined, we can determine the thermodynamics and kinetics of cobalamin binding/unbinding through Gaussian-deconvolution of the normalized intensity distributions and waiting time analysis of the smFQ states, respectively. It is likely that smFQ information from both labeling sites (T71C and S99C) will be utilized as these two perspectives will yield different, yet complimentary, observables. These experiments will produce quantitative kinetic information on substrate specificity, which is important for understanding the mechanism by which ATR binds and delivers cobalamin. We also intend to explore the ATR–MCM interaction dynamics, perhaps via smFRET with nanovesicle trapping, and explore how these are coupled to cobalamin binding/unbinding and transfer. Overall, this study broadens our scope to interrogating the trafficking of other metal-containing cofactors, where the presence of cofactor can be directly probed, thus offering a unique contribution to the field of bioinorganic chemistry.
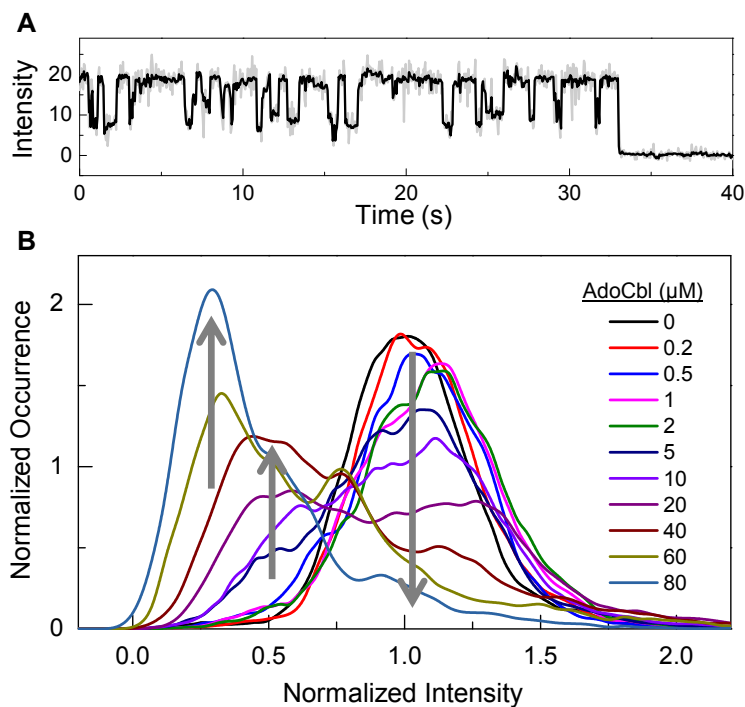
Figure 17. (A) smFQ trajectory of the ATR T71C construct labeled with Alexa555 in presence of 5 μM AdoCbl. The light grey represents the raw intensity corrected by the TIRF laser profile and the black represents the data subjected to nonlinear forwards-backwards filtering (using the same parameters described in Chapter 2, Section 2.9). (B) Compiled normalized intensity histograms for hundreds of ATR molecules showing an increase in the population of quenched states with increasing AdoCbl concentration.

APPENDIX

S1. Calibration of E_{FRET} vs. Donor-Acceptor Distance

An empirical E_{FRET} vs. distance calibration curve (Figure S1A) was prepared by Debashis Panda by compiling smFRET data from various Holliday Junction (HJ) constructs labeled with Cy3 and Cy5. The E_{FRET} values are taken as $I_A/(I_A + I_D)$, as described in the Chapter 2, Section 2.9. The distances are the anchor-to-anchor distances of the FRET donor-acceptor pair Cy3−Cy5. These HJs are four-way junctions of DNA. In the presence of salt and at room temperature, they form two dynamically interconverting X-shaped conformations, each like two DNA helices stacked on top of each other.[98-100] Using structural models of these HJ conformations, the distances between Cy3 and Cy5 labeling sites (i.e., dye anchoring points) can be measured readily. The corresponding E_{FRET} values were obtained from smFRET measurements, as we reported previously.[76,101]

In the calibration curve, data were included from the following HJs: HJ1 and HJ1a (described in Sarkar et al),[76] HJC2 and HJC2a (described in Andoy et al),[101] HJC3 (Figure S1B), HJC4 (Figure S1C), and HJC4a (Figure S1D). For the previously unreported HJ constructs (HJC3, HJC4, and HJC4a), the Cy3 and Cy5 labeling positions are shown along with the names of each strand, "a, b, c, and d," and the names of each arm, "M, N, P, and Q" (Figure S1B-D). The BiotinTEG moiety at the 5′ end of strand c was used for biotin-avidin immobilization. Because the HJ is in dynamic equilibrium between two conformational states, constructs with arms M and N/Q labeled give rise to a low E_{FRET} state of ~0.2 and a high E_{FRET} state of ~0.7 (see cartoon representations of dynamic equilibrium in Figure S1B-C). The HJC4a construct, however, gives rise to just one apparent E_{FRET} state ~0.4, because Cy3 and Cy5 are within arms'

M and P, which do not move significantly with respect to one another during the HJC4a's structural dynamics.

To obtain the corresponding distances between labeling sites (i.e., anchor-to-anchor distances), r, for each E_{FRET} state, a HJ model was generated by overlaying and extending B-form DNA onto the four arms of the crystal structure of a HJ junction (PDB code 1DCW).[76,98] The distance measurement was made between two phosphate backbone atoms corresponding to the anchoring position of the dyes. Within each conformation, the dye-labeled strands can be positioned in two possible, but slightly different ways. In all constructs except HJC4a, this results in two similar donor-acceptor distances for each E_{FRET} state: r_{12} and r_{34} for the high E_{FRET} and r_{12}' and r_{34}' for the low E_{FRET} (Figure S1B-C). In HJC4a, this results in four similar donor-acceptor distances: r_{12}, r_{34}, r_{12}', and r_{34}' (Figure S1D). The distance we used for each E_{FRET} state, r, was therefore taken as an average of these possibilities.

The data were fitted empirically using $E_{FRET} = \left(1 + \left((r+\alpha)/(r_0\beta)\right)^6\right)^{-1}$, with $\alpha = 5.3 \pm 0.6$ nm and $\beta = 1.9 \pm 0.1$, and r_0 (the Cy3-Cy5 Förster radius) fixed at the previously reported value of 5.4 nm.[102] Here α and β are correction parameters: α corrects for the additional distance due to the dye linker length (since we only measured the distance between the dye anchoring points for r) and β corrects for deviation in r_0 due to the relative orientation and linker flexibility of the dyes.[103] Note the E_{FRET} value used here is a simple experimental approximation (= $I_A/(I_D+I_A)$), and it is not corrected by the relative fluorescence detection efficiencies and quantum yields of the donor and acceptor.[102,103] Therefore, the numerical values and α and β should not be interpreted literally, but should be treated as empirical fitting parameters to generate a best fit to the data, despite their physical connection. This empirical calibration curve provides a direct correlation between an experimental observable and a distance quantity that can be independently determined using structural modeling. The fitted curve and the HJ smFRET data

demonstrate that the apparent E_{FRET} varies with distance over a range relevant to our measured E_{FRET} values (~0.27–0.75).
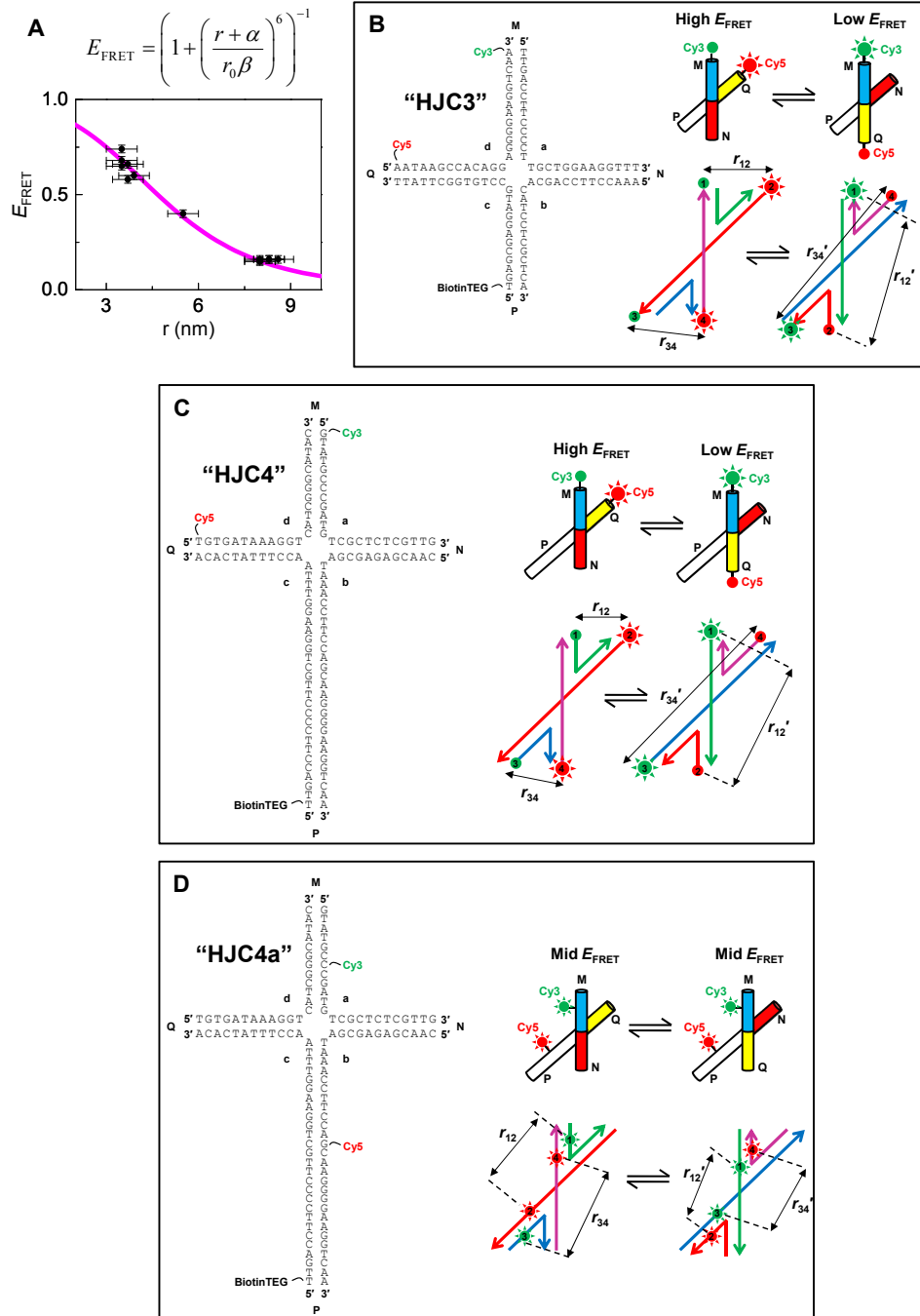
Figure S1. (A) Empirical E_{FRET} calibration curve based on smFRET data for various Holliday junction (HJ) constructs. The labeling schemes, dynamic equilibrium, and an illustration of determining r are shown for HJC3 (B), HJC4 (C), and HJC4a (D).

S2. Dwell Time Distributions of E_{FRET} States

The dwell times of the E_{Low}, E_{Mid}, and E_{High} states, τ_{Low}, τ_{Mid}, and $\tau_{High,}$ respectively, can

be compiled from many interacting pairs to examine their statistical distributions (Figure S2).

The distributions follow an apparent single-exponential decay, $f(\tau) = N\exp(-\tau/\overline{\tau})$ which can be

used to obtain the average lifetime of the E_{FRET} state, $\overline{\tau}$. Fast transitions consisting of 1–2 image

frames (100–200 ms) were excluded from these distributions as these dwell times and

corresponding E_{FRET} values are unreliable due to our experimental resolution. The distributions

in Figure S2 were obtained from analyzing the raw trajectories (not subjected to nonlinear

filtering described in Chapter 2, Section 2.9). Dwell time analysis performed after nonlinear

filtering resulted in identical average lifetimes.
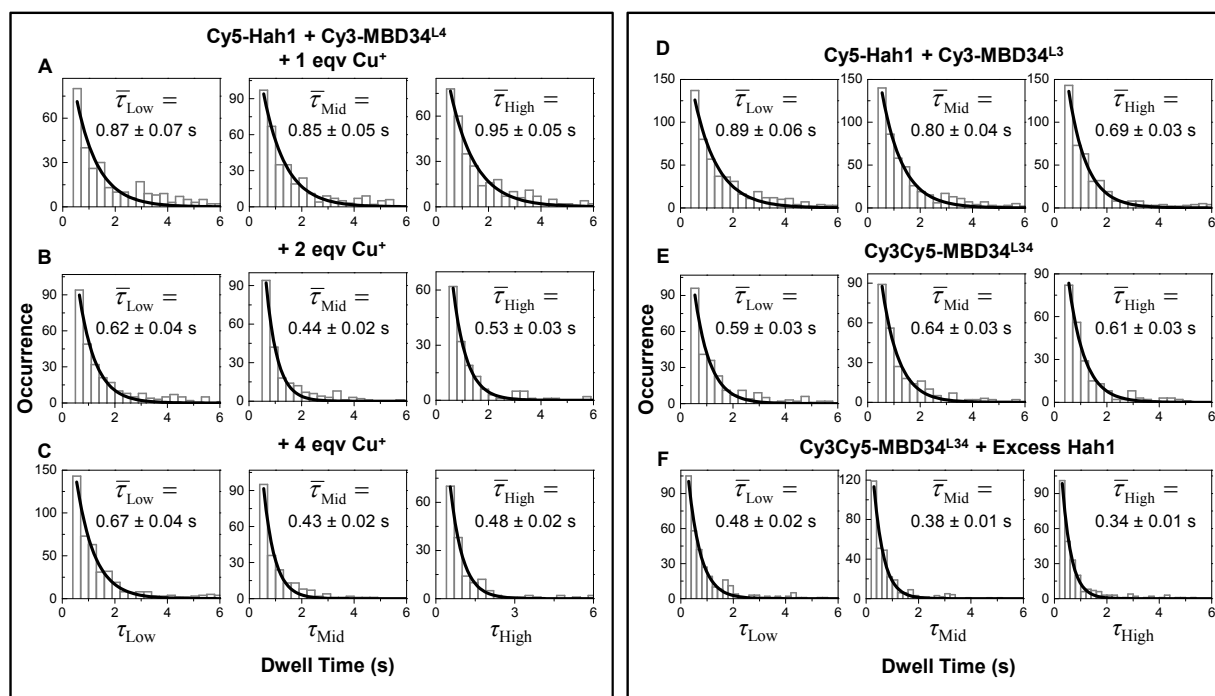
Figure S2. Dwell time distributions for Cy5-Hah1 + Cy3-MBD34^{L4} in the presence of (A) 1 eqv Cu^{+}, (B) 2 eqv Cu^{+}, (C) 4 eqv Cu^{+}, (D) Cy5-Hah1 + Cy3-MBD34^{L3}, (E) Cy3Cy5-MBD34^{L34}, and (F) Cy3Cy5-MBD34^{L34} in the presence of excess unlabeled Hah1. The average lifetimes obtained from the single-exponential fit, $\overline{\tau}_{Low}$, $\overline{\tau}_{Mid}$, and $\overline{\tau}_{High}$, are given for each distribution.

S3. Two-Dimensional E_{FRET} Histogram Analysis

The 2-D E_{FRET} analysis of Cy5-Hah1 + Cy3-MBD34^{L4}, in the presence of varying stoichiometric equivalents of Cu$^+$ (Figure S3A), resolves three apparent E_{FRET} states: E_{Low}^{L4}, E_{Mid}^{L4}, and E_{High}^{L4}. All E_{FRET} values are consistent with those observed in the *apo* condition (Chapter 3, Section 3.2, Figure 8C) suggesting that the Hah1–MBD4 interaction geometries remain largely unchanged in the presence of Cu$^+$. In addition, three E_{FRET} values, E_{Low}^{L34}, E_{Mid}^{L34}, and E_{High}^{L34} were also observed for Cy3Cy5-MBD34^{L34} in the presence of excess Hah1 (Figure S3B). The E_{FRET} values are similar to those observed in the absence of Hah1 (Chapter 3, Section 3.2, Figure 8I), indicating the two intramolecular MBD3–MBD4 interaction geometries are conserved in the presence of Hah1.
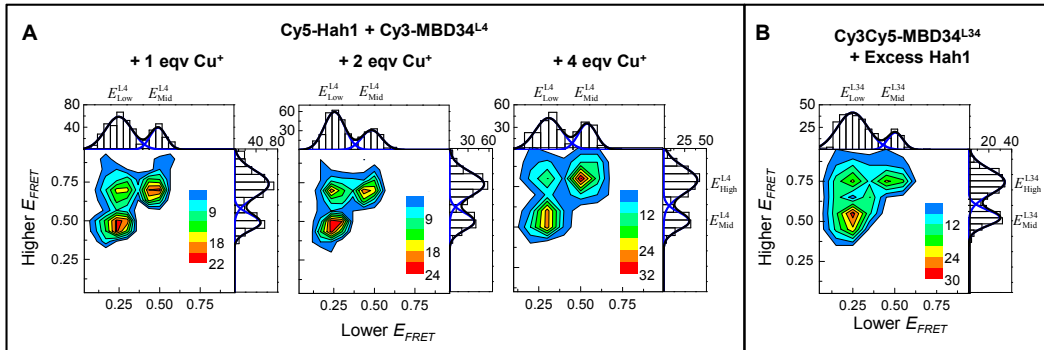
Figure S3. (**A**) Two-dimensional histogram of the average lower vs. higher E_{FRET} state for Cy5-Hah1 + Cy3-MBD34^{L4} with varying equivalents of Cu$^+$. For each condition, three E_{FRET} states are observed. For 1 eqv Cu$^+$ (249 pairs), E_{Low}^{L4} = 0.25 ± 0.01, E_{Mid}^{L4} = 0.49 ± 0.01, and E_{High}^{L4} = 0.74 ± 0.01. For 2 eqv Cu$^+$ (220 pairs), E_{Low}^{L4} = 0.26 ± 0.01, E_{Mid}^{L4} = 0.49 ± 0.01, and E_{High}^{L4} = 0.73 ± 0.01. For 4 eqv Cu$^+$ (183 pairs), E_{Low}^{L4} = 0.25 ± 0.01, E_{Mid}^{L4} = 0.49 ± 0.01, and E_{High}^{L4} = 0.76 ± 0.01. (**B**) The two-dimensional E_{FRET} histogram for Cy3Cy5-MBD34^{L34} (187 molecules) in the presence of excess unlabeled Hah1 also shows three E_{FRET} states: E_{Low}^{L34} = 0.24 ± 0.01, E_{Mid}^{L34} = 0.51 ± 0.01, and E_{High}^{L34} = 0.75 ± 0.01.

S4. Detailed Description of Protein–Protein Interaction Interfaces

Below we present the details about the interaction interfaces for the face-to-face (Section S4.1), head-to-head (Section S4.2), and face-to-back (Section S4.3) models. The interface analyses were done using PISA.[81]

S4.1. Face-to-Face Interaction Models. The face-to-face interaction interfaces are illustrated and described for the $H_{ff}4$ (Figure S4), $H_{ff}3$ (Figure S5), $3_{ff}4$ (Figure S6) models. The interface residues are mainly located within the two α-helices and loop regions next to the CXXC Cu^+-binding motif. The face-to-face interaction models are also summarized in Chapter 4, Section 4.1 and Table 2.

B Hah1 Interface Residues			**C** Favorable Interactions		**D** MBD4 Interface Residues		
Residue	BSA (Å²)	-ΔG$_{solv}$ (kcal/mol)	Hah1	MBD4	Residue	BSA (Å²)	-ΔG$_{solv}$ (kcal/mol)
THR 11	24.83	-0.05	*Hydrogen Bonds*		ALA 132	15.79	-0.15
CYS 12	68.99	2.32	CYS 12 [SG]	CYS 136 [O]	GLY 133	3.76	0.04
GLY 13	23.07	0.14	CYS 12 [SG]	CYS 139 [SG]	THR 135	39.15	0.54
GLY 14	43.78	0.48	CYS 15 [H]	CYS 139 [SG]	CYS 136	20.90	0.70
CYS 15	36.71	1.05	ARG 21 [NH2]	ALA 132 [O]	SER 138	50.73	0.40
GLU 17	6.34	-0.02	CYS 12 [SG]	CYS 136 [H]	CYS 139	26.94	1.03
ALA 18	38.08	0.54	THR 11 [O]	SER 138 [OG]	SER 142	28.98	-0.01
ARG 21	84.83	-0.82	*Salt Bridges*		MET 146	39.00	1.11
VAL 22	19.39	0.31	LYS 25 [NZ]	ASP 185 [OD1]	GLU 184	19.25	-0.12
LYS 25	46.62	-0.98			ASP 185	59.03	-0.40
LYS 57	46.78	0.19			MET 186	64.67	0.24
THR 58	26.90	0.09			GLY 187	54.31	0.29
GLY 59	46.72	0.15			PHE 188	60.26	0.81
LYS 60	8.21	-0.13			GLU 189	16.45	0.26

Figure S4. (A) Illustration of interface residues for $H_{ff}4$. Hah1's interface residues are shown in green while MBD4's interface residues are in orange. In addition, each protein is rotated ~90° to illustrate the interface residues more clearly. (B) Hah1 interface residues with corresponding buried surface area, BSA, and solvation free energy change, ΔG$_{solv}$. (C) Favorable interactions – hydrogen bonds and salt bridges. (D) MBD4 interface residues and corresponding BSA and ΔG$_{solv}$.

Figure S5. (A) Illustration of interface residues for $H_{ff}3$. Hah1's interface residues are shown in green while MBD3's interface residues are in purple. In addition, each protein is rotated ~90° to illustrate the interface residues more clearly. (B) Hah1 interface residues with corresponding BSA and ΔG_{solv}. (C) Favorable interactions – hydrogen bonds and salt bridges. (D) MBD3 interface residues and corresponding BSA and ΔG_{solv}.

B	MBD3 Interface Residues			C	Favorable Interactions		D	MBD4 Interface Residues		
Residue	BSA (Å^2)	$-\Delta G_{solv}$ (kcal/mol)		MBD3	MBD4		Residue	BSA (Å^2)	$-\Delta G_{solv}$ (kcal/mol)
GLY 31	3.63	-0.01		*Hydrogen Bonds*			ALA 132	10.30	-0.12
HIS 33	0.47	0.00		SER 36 [OG]	CYS 136 [H]		GLY 133	19.11	0.29
ALA 34	54.12	2.00					MET 134	2.06	0.01
LYS 35	28.30	0.31					THR 135	67.06	0.85
SER 36	70.01	0.34					CYS 136	39.68	1.19
ALA 37	21.97	0.89					ALA 137	1.02	-0.01
LEU 39	39.20	0.63					SER 138	41.12	0.24
ASN 40	44.40	0.58					CYS 139	28.34	1.16
GLU 43	44.78	-0.44					HIS 141	3.14	0.02
ASN 44	22.23	-0.30					SER 142	32.42	-0.17
ALA 83	16.68	-0.18					MET 146	78.46	1.53
LEU 84	47.52	0.34					ILE 147	7.38	0.00
PRO 85	8.89	0.11					GLN 149	15.16	-0.22
PRO 86	133.30	1.24					LEU 150	1.50	0.02
GLY 87	24.10	-0.07					ALA 182	1.84	-0.02
ASN 88	54.24	-0.43					GLU 184	12.46	-0.14
PHE 89	25.33	0.41					ASP 185	38.85	-0.28
							MET 186	68.17	0.38
							GLY 187	61.05	0.20
							PHE 188	53.28	0.68
							GLU 189	9.66	0.15

Figure S6. (A) Illustration of interface residues for 3$_{ff}$4. MBD3's interface residues are shown in purple while MBD4's interface residues are in orange. In addition, each protein is rotated ~90° to illustrate the interface residues more clearly. (B) MBD3 interface residues with corresponding BSA and ΔG_{solv}. (C) Favorable interactions – hydrogen bonds. (D) MBD4 interface residues and corresponding BSA and ΔG_{solv}.

S4.2. Head-to-Head Interaction Models. The head-to-head interaction interfaces are illustrated and described for the $H_{hh}4$ (Figure S7), $H_{hh}3$ (Figure S8), $3_{hh}4$ (Figure S9) models. Similar to the face-to-face interaction, the head-to-head interface residues are located within the two α-helices, but more confined to the loop regions next to the CXXC Cu^+-binding motif. The head-to-head interaction models are also summarized in Chapter 4, Section 4.1 and Table 2.

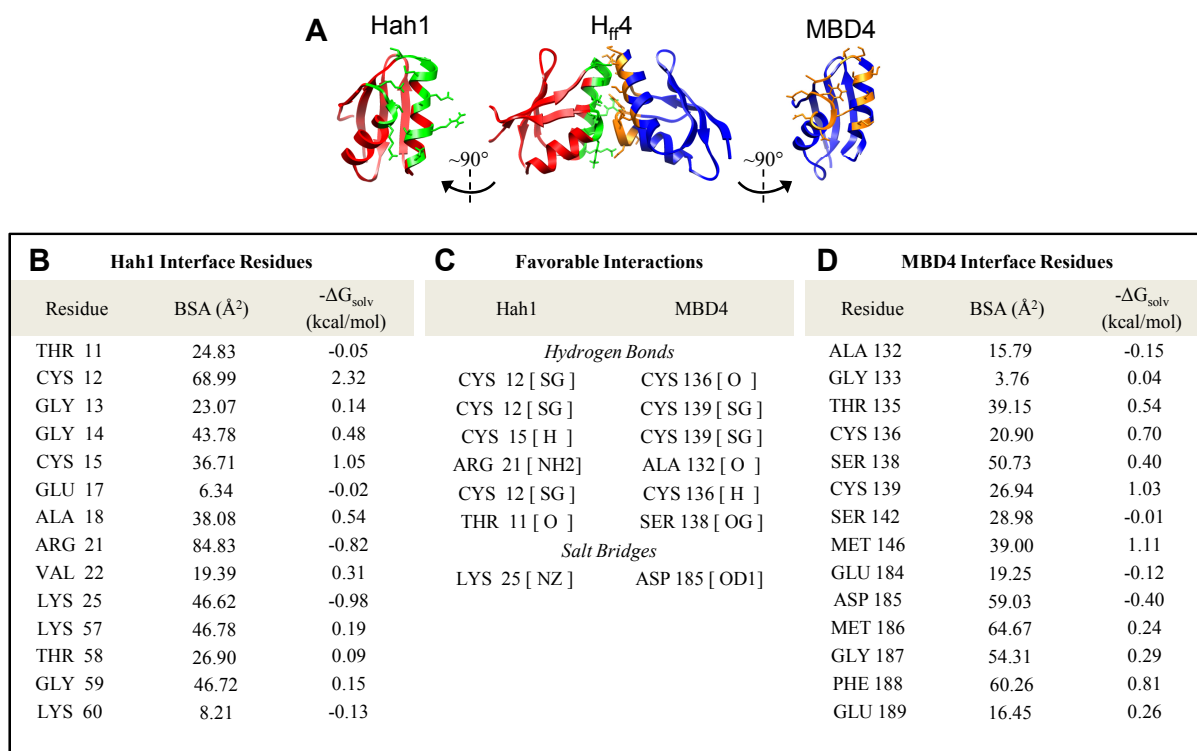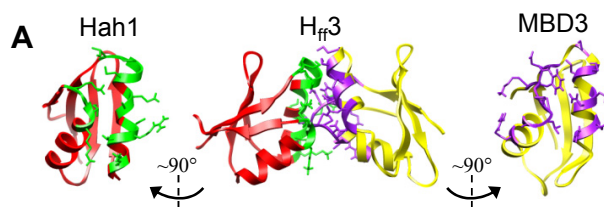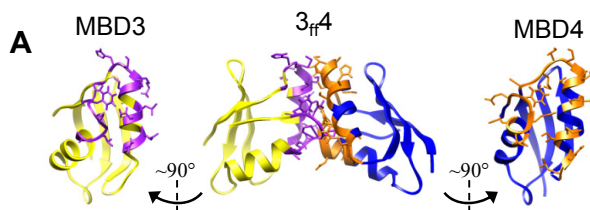| **B**	**Hah1 Interface**		**C**	**Favorable Interactions**		**D**	**MBD4 Interface**	
Residue	BSA (Å^2)	$-\Delta G_{solv}$ (kcal/mol)	Hah1		MBD4	Residue	BSA (Å^2)	$-\Delta G_{solv}$ (kcal/mol)
ASP 9	17.23	0.09		*Hydrogen Bonds*		ALA 122	1.49	0.00
MET 10	0.34	0.01	ASP 9 [O]		THR 135 [HG1]	GLY 133	23.49	0.24
THR 11	65.45	-0.20				MET 134	3.89	-0.03
CYS 12	57.63	1.91				THR 135	107.12	0.45
GLY 13	14.72	-0.07				CYS 136	65.14	2.14
GLY 14	28.90	0.28				ALA 137	10.29	0.12
CYS 15	22.03	0.87				SER 138	41.97	0.13
GLU 17	11.83	-0.14				CYS 139	19.90	0.69
ALA 18	25.84	0.41				HIS 141	29.38	0.91
ARG 21	19.27	-0.22				SER 142	28.97	0.02
LEU 35	2.51	0.04				ILE 143	1.96	0.03
THR 58	20.63	-0.21				MET 186	10.62	0.08
GLY 59	15.04	0.03				GLY 187	1.89	-0.02
LYS 60	38.14	-0.78				PHE 188	62.68	1.00
THR 61	2.14	-0.02						

Figure S7. (A) Illustration of interface residues for $H_{hh}4$. Hah1's interface residues are shown in green while MBD4's interface residues are in orange. In addition, each protein is rotated ~90° to illustrate the interface residues more clearly. (B) Hah1 interface residues with corresponding BSA and ΔG_{solv}. (C) Favorable interactions – hydrogen bonds. (D) MBD4 interface residues and corresponding BSA and ΔG_{solv}.

A

Hah1 $H_{hh}3$ MBD3

~90°   ~90°

B	**Hah1 Interface Residues**		**C**	**Favorable Interactions**		**D**	**MBD3 Interface Residues**	
Residue	BSA (Å^2)	$-\Delta G_{solv}$ (kcal/mol)	Hah1	MBD3		Residue	BSA (Å^2)	$-\Delta G_{solv}$ (kcal/mol)
ASP 9	10.10	-0.09	*Salt Bridges*			ASP 30	3.58	-0.03
MET 10	0.31	-0.00	ASP 9 [OD2]	HIS 30 [NE2]		GLY 31	5.19	0.07
THR 11	28.15	0.04				HIS 33	6.72	0.13
CYS 12	24.82	0.63				ALA 34	22.91	0.58
GLY 13	21.06	0.13				LYS 35	3.05	0.05
GLY 14	12.97	0.16				SER 36	16.85	0.08
CYS 15	7.35	0.29				ALA 37	2.08	0.01
GLU 17	9.13	-0.10				LEU 39	1.88	0.03
ALA 18	3.69	0.06				ASN 40	19.38	0.04
ARG 21	3.78	-0.04				GLU 43	0.86	-0.01
LEU 35	20.70	0.32				PRO 85	7.57	0.12
PRO 36	3.28	0.05				PRO 86	34.00	0.52
THR 58	6.31	-0.06				ASN 88	27.09	-0.34
GLY 59	6.18	-0.00				PHE 89	9.98	0.16
LYS 60	9.63	-0.16						
THR 61	4.09	-0.05						

Figure S8. (A) Illustration of interface residues for $H_{hh}3$. Hah1's interface residues are shown in green while MBD3's interface residues are in purple. In addition, each protein is rotated ~90° to illustrate the interface residues more clearly. (B) Hah1 interface residues with corresponding BSA and ΔG_{solv}. (C) Favorable interactions – salt bridges. (D) MBD3 interface residues and corresponding BSA and ΔG_{solv}.

B	MBD3 Interface Residues		
Residue	BSA (Å²)	-ΔG$_{solv}$ (kcal/mol)	
ASP 30	3.31	-0.01	
GLY 31	14.30	0.20	
HIS 33	13.38	0.41	
ALA 34	88.48	1.82	
LYS 35	21.13	0.22	
SER 36	58.42	0.20	
ALA 37	5.95	-0.05	
LEU 39	4.12	0.07	
ASN 40	49.85	0.18	
ILE 41	0.17	0.00	
GLU 43	1.70	-0.02	
LEU 84	1.39	0.01	
PRO 85	12.53	0.20	
PRO 86	26.54	0.41	
ASN 88	79.24	-0.88	
PHE 89	41.85	0.67	

C	Favorable Interactions	
	MBD3	MBD4
	Hydrogen Bonds	
	PHE 89 [H]	CYS 136 [SG]
	ASP 30 [O]	THR 135 [HG1]

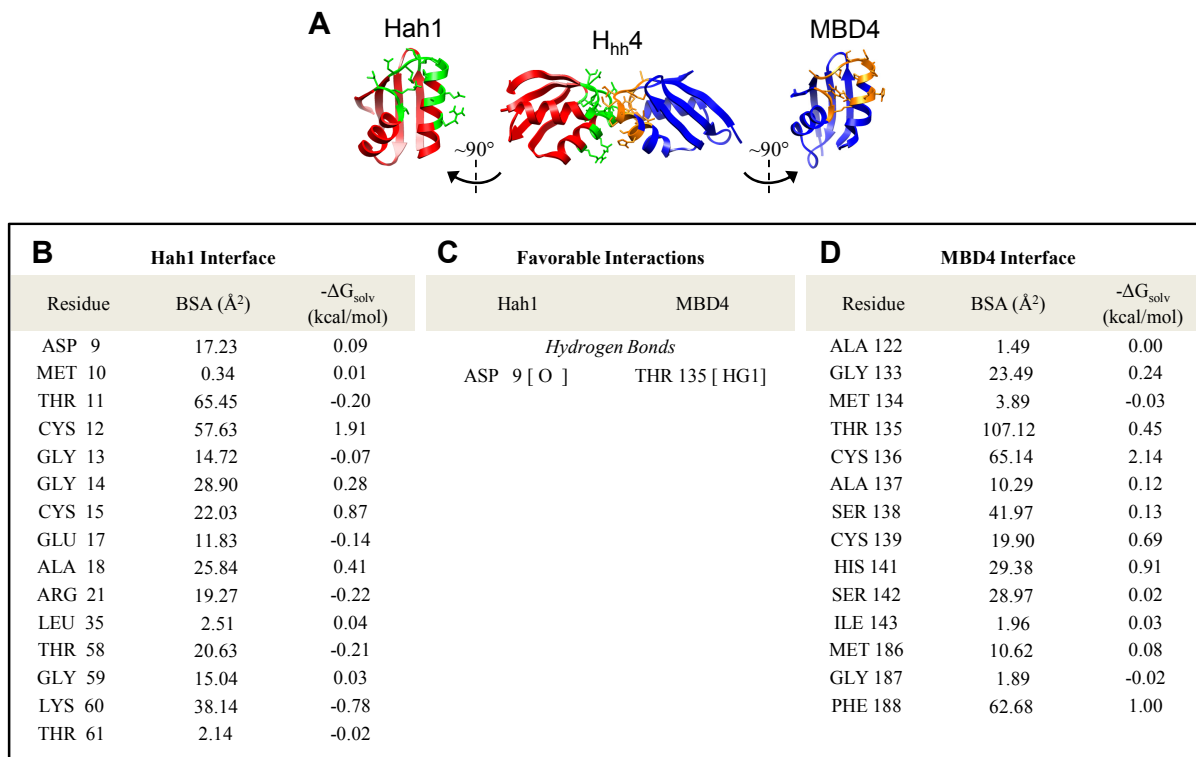| D	MBD4 Interface Residues		
Residue	BSA (Å²)	-ΔG$_{solv}$ (kcal/mol)	
ALA 132	21.33	0.00	
GLY 133	9.64	0.14	
MET 134	9.82	-0.07	
THR 135	116.94	0.38	
CYS 136	89.90	3.32	
ALA 137	52.07	0.76	
SER 138	49.64	0.12	
CYS 139	15.53	0.59	
HIS 141	12.50	0.70	
SER 142	20.30	-0.03	
LEU 160	16.05	0.26	
MET 186	15.54	0.02	
GLY 187	20.91	0.01	
PHE 188	61.04	0.98	
GLU 189	2.95	-0.03	

Figure S9. (A) Illustration of interface residues for 3$_{hh}$4. MBD3's interface residues are shown in purple while MBD4's interface residues are in orange. In addition, each protein is rotated ~90° to illustrate the interface residues more clearly. (B) MBD3 interface residues with corresponding BSA and ΔG$_{solv}$. (C) Favorable interactions – hydrogen bonds. (D) MBD4 interface residues and corresponding BSA and ΔG$_{solv}$.
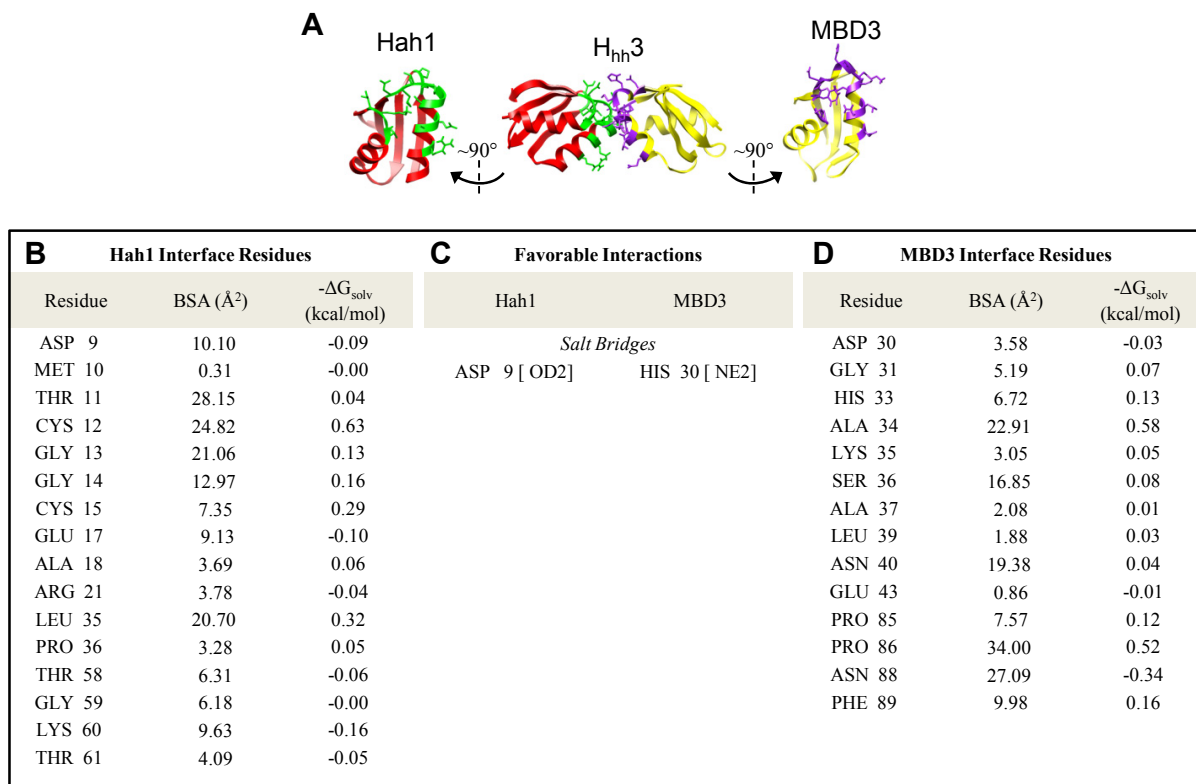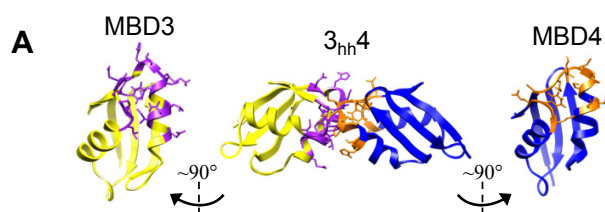
S4.3. Face-to-Back Interaction Models. The face-to-back interaction interfaces are illustrated and described for the $H_{fb}4$ (Figure S10), $H_{fb}3$ (Figure S11), $3_{fb}4$ (Figure S12) models. The "face" interface residues are in a region similar to those observed in the face-to-face interactions, while the "back" interface residues are found in the β-sheet regions next to the N-terminus. The face-to-back interaction models are also summarized in Chapter 4, Section 4.1 and Table 2.

B	Hah1 Interface		C	Favorable Interactions		D	MBD4 Interface	
Residue	BSA (Å²)	$-\Delta G_{solv}$ (kcal/mol)	Hah1		MBD4	Residue	BSA (Å²)	$-\Delta G_{solv}$ (kcal/mol)
ASP 9	18.30	-0.06	*Hydrogen Bonds*			THR 123	45.95	0.30
THR 11	55.16	-0.03	CYS 15 [SG]		SER 124 [O]	SER 124	81.49	2.21
CYS 12	29.13	0.92	CYS 15 [SG]		SER 124 [OG]	SER 125	0.50	0.01
GLY 14	1.61	0.03	THR 11 [O]		THR 123 [N]	THR 126	23.92	0.00
CYS 15	22.33	0.90				GLN 154	45.81	-0.47
ALA 18	5.42	0.09				GLN 155	65.37	-0.82
LYS 56	3.52	-0.04				THR 166	6.80	0.11
LYS 57	5.84	-0.07				LEU 168	34.00	0.54
THR 58	17.05	-0.19				PRO 171	0.70	0.01
GLY 59	58.37	0.07						
LYS 60	35.73	-0.55						
THR 61	9.43	-0.08						

Figure S10. (A) Illustration of interface residues for $H_{fb}4$. Hah1's interface residues are shown in green while MBD4's interface residues are in orange. In addition, each protein is rotated ~90° to illustrate the interface residues more clearly. (B) Hah1 interface residues with corresponding BSA and ΔG_{solv}. (C) Favorable interactions – hydrogen bonds. (D) MBD4 interface residues and corresponding BSA and ΔG_{solv}.
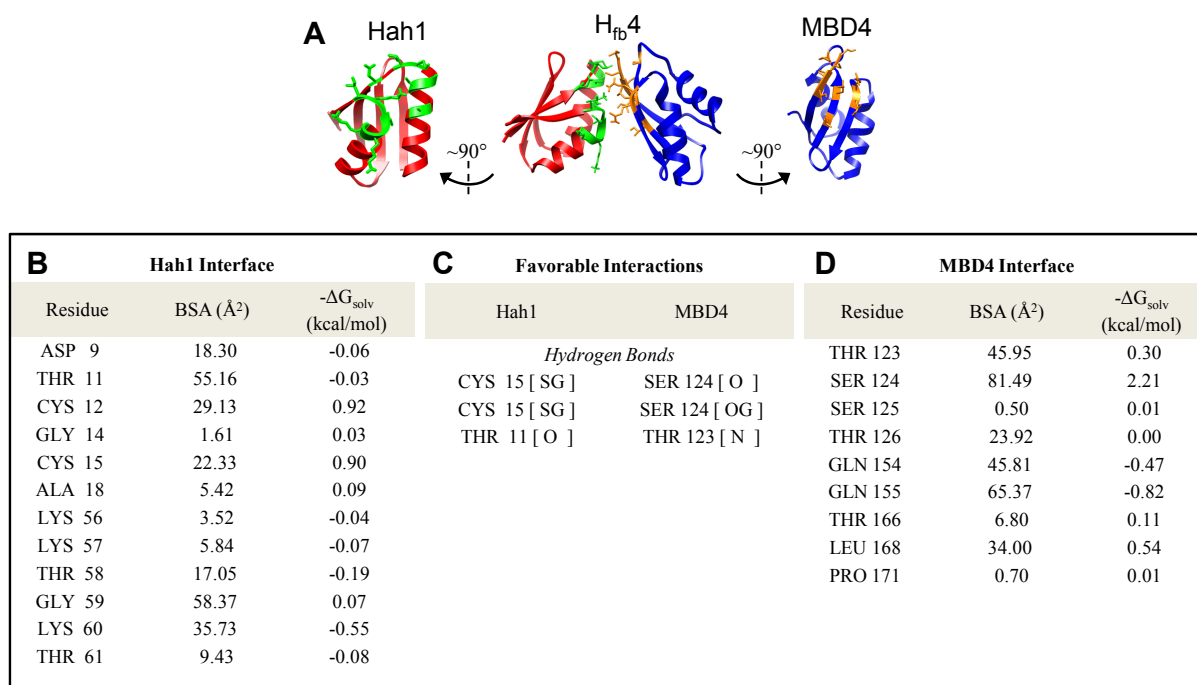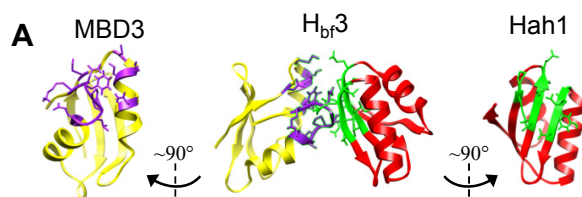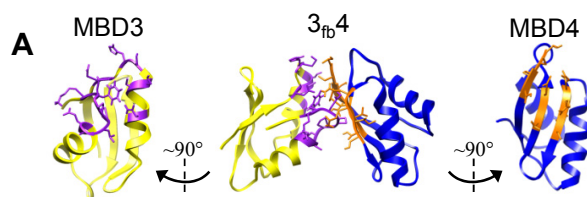
Figure S11. (A) Illustration of interface residues for $H_{bf}3$. Hah1's interface residues are shown in green while MBD3's interface residues are in orange. In addition, each protein is rotated ~90° to illustrate the interface residues more clearly. (B) MBD3 interface residues with corresponding BSA and ΔG_{solv}. (C) Favorable interactions – hydrogen bonds. (D) Hah1 interface residues and corresponding BSA and ΔG_{solv}.

B — MBD3 Interface Residues

Residue	BSA (Å^2)	$-\Delta G_{solv}$ (kcal/mol)
ASP 30	0.49	-0.01
ALA 34	3.79	0.15
SER 36	6.99	-0.00
ALA 37	5.93	0.24
ASN 40	5.51	-0.09
GLU 82	3.82	0.02
ALA 83	7.91	0.03
LEU 84	17.90	-0.20
PRO 85	13.14	0.11
PRO 86	126.32	1.51
GLY 87	55.80	0.06
ASN 88	72.11	-0.47
PHE 89	17.75	0.25
LYS 90	11.35	0.18

C — Favorable Interactions

MBD3	Hah1
Hydrogen Bonds	
LEU 84 [O]	LYS 3 [HZ2]
GLY 87 [O]	TYR 31 [H]

D — Hah1 Interface Residues

Residue	BSA (Å^2)	$-\Delta G_{solv}$ (kcal/mol)
PRO 2	23.40	0.14
LYS 3	45.32	-0.61
GLY 28	1.58	-0.00
VAL 29	0.79	0.01
LYS 30	130.94	1.05
TYR 31	15.85	-0.05
ASP 32	20.95	0.12
SER 41	1.58	0.02
ILE 42	0.34	0.01
GLU 43	68.53	-0.37
SER 44	6.75	0.08
GLU 45	3.58	0.06

B	MBD3 Interface Residues		C	Favorable Interactions		D	MBD4 Interface Residues	
Residue	BSA (Å^2)	$-\Delta G_{solv}$ (kcal/mol)	MBD3	MBD4		Residue	BSA (Å^2)	$-\Delta G_{solv}$ (kcal/mol)
ASP 30	10.50	-0.12	*Hydrogen Bonds*			THR 123	43.87	0.25
GLY 31	17.99	0.20	GLY 31 [O]	SER 124 [SG]		SER 124	75.22	1.89
HIS 33	6.46	0.10	ASP 30 [OD2]	GLN 154 [NE2]		SER 125	2.68	0.04
ALA 34	48.69	1.91	ASN 88 [O]	GLN 154 [NE2]		THR 126	37.99	0.14
SER 36	10.26	-0.02	GLU 82 [O]	GLN 155 [NE2]		GLN 154	45.33	-0.39
ALA 37	18.60	0.76	ASN 88 [ND2]	SER 125 [O]		GLN 155	86.71	-0.91
ASN 40	9.62	-0.10				ILE 156	0.49	0.01
GLU 82	9.90	0.02				SER 157	6.49	0.10
ALA 83	2.15	-0.02				THR 166	28.01	0.12
LEU 84	3.90	-0.04				VAL 167	2.39	0.04
PRO 85	14.06	0.12				LEU 168	45.82	0.72
PRO 86	117.00	1.25						
GLY 87	58.51	0.43						
ASN 88	117.40	-0.76						
PHE 89	22.80	0.35						
LYS 90	9.37	0.15						

Figure S12. (A) Illustration of interface residues for 3$_{fb}$4. MBD3's interface residues are shown in purple while MBD4's interface residues are in orange. In addition, each protein is rotated ~90° to illustrate the interface residues more clearly. (B) MBD3 interface residues with corresponding BSA and ΔG_{solv}. (C) Favorable interactions – hydrogen bonds. (D) MBD4 interface residues and corresponding BSA and ΔG_{solv}.

S5. Details of Molecular Dynamics Simulation

All molecular dynamics simulations were performed by Linghao Zhong (Pennsylvania State University, Mont Alto, PA). The simulations were carried out on four possible models of the Hah1-MBD4 complex ($H_{ff}4$, $H_{fb}4$, $H_{bf}4$, and $H_{hh}4$). All calculations were performed with CHARMM[104] using the CHARMM27[105] protein force field. For each model, the complex was placed in the center of a TIP3P[106] periodic water box. A minimum of 10 Å solvation shell was allowed for each protein atom, resulting in a box of about $80 \times 50 \times 50$ Å^3 in size. Five Cl$^-$ and 13 Na$^+$ ions were added to neutralize the system, so particle mesh Ewald could be used for the long-range electrostatics.[107] The cutoff for non-bonded interactions was set to be 12 Å. The SHAKE algorithm[108] was applied to fix lengths for all bonds with hydrogen attached, and a 2-fs step size was used for the integration. The system was first heated up stepwise at 100 K, 200 K and 300 K, each for 20 ps, followed by 40 ps of equilibration time. To collect the production trajectory, the system was integrated as a microcanonical ensemble (NVE) at 300 K with periodic boundary conditions (PBC). For the $H_{ff}4$, $H_{fb}4$, and $H_{hh}4$ models, the production trajectories were collected for 10 ns. Only 1.5 ns of production were collected for the $H_{bf}4$ model due to the rapid dissociation.

REFERENCES

(1) Byung-Eun Kim, T. N., & Dennis J Thiele: Mechanisms for copper acquisition, distribution and regulation. *Nat. Chem. Biol.* **2008**, *4*, 176.

(2) Klaudia Jomova, D. V., Michael Lawson, Marian Valko: Metals, oxidative stress and neurodegenerative disorders. *Mol. Cell. Biochem.* **2010**, *345*, 91-104.

(3) Robinson, N. J.; Winge, D. R.: Copper Metallochaperones. *Ann. Rev. Biochem.* **2010**, *79*, 537-562.

(4) Pufahl, R. A.; Singer, C. P.; Peariso, K. L.; Lin, S.-J.; Schmidt, P. J.; Fahrni, C. J.; Culotta, V. C.; Penner-Hahn, J. E.; O'Halloran, T. V.: Metal Ion Chaperone Function of the Soluble Cu(I) Receptor Atx1. *Science* **1997**, *278*, 853-856.

(5) Hamza, I.; Schaefer, M.; Klomp, L. W. J.; Gitlin, J. D.: Interaction of the Copper Chaperone Hah1 with the Wilson Disease Protein Is Essential for Copper Homeostasis. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 13363-13368.

(6) Larin, D.; Mekios, C.; Das, K.; Ross, B.; Yang, A.-S.; Gilliam, T. C.: Characterization of the Interaction between the Wilson and Menkes Disease Proteins and the Cytoplasmic Copper Chaperone, Hah1p. *J. Biol. Chem.* **1999**, *274*, 28497-28504.

(7) O'Halloran, T. V.; Culotta, V. C.: Metallochaperones, An Intracellular Shuttle Service for Metal Ions. *J. Biol. Chem.* **2000**, *275*, 25057-25060.

(8) Lutsenko, S.; Barnes, N. L.; Bartee, M. Y.; Dmitriev, O. Y.: Function and regulation of human copper-transporting ATPases. *Physiol. Rev.* **2007**, *87*, 1011-46.

(9) Terada, K.; Nakako, T.; Yang, X. L.; Iida, M.; Aiba, N.; Minamiya, Y.; Nakai, M.; Sakaki, T.; Miura, N.; Sugiyama, T.: Restoration of holoceruloplasmin synthesis in LEC rat after infusion of recombinant adenovirus bearing WND cDNA. *J. Biol. Chem.* **1998**, *273*, 1815-20.

(10) Petris, M. J.; Strausak, D.; Mercer, J. F.: The Menkes copper transporter is required for the activation of tyrosinase. *Hum. Mol. Genet.* **2000**, *9*, 2845-51.

(11) Hung, I. H.; Suzuki, M.; Yamaguchi, Y.; Yuan, D. S.; Klausner, R. D.; Gitlin, J. D.: Biochemical characterization of the Wilson disease protein and functional expression in the yeast Saccharomyces cerevisiae. *J. Biol. Chem.* **1997**, *272*, 21461-6.

(12) Petris, M. J.; Mercer, J. F.; Culvenor, J. G.; Lockhart, P.; Gleeson, P. A.; Camakaris, J.: Ligand-regulated transport of the Menkes copper P-type ATPase efflux pump from the Golgi apparatus to the plasma membrane: a novel mechanism of regulated trafficking. *EMBO J.* **1996**, *15*, 6084-95.

(13) Lutsenko, S.; LeShane, E. S.; Shinde, U.: Biochemical Basis of Regulation of Human

Copper-Transporting ATPase. *Arch. Biochem. Biophys.* **2007**, *463*, 134-148.

(14) Wilson, S. A. K.: Progressive lenticular degeneration: a familial nervous disease associated with cirrhosis of the liver. *Brain* **1912**, *34*, 296-484.

(15) Menkes, J. H.; Alter, M.; Steigleder, G. K.; Weakley, D. R.; Sung, J. H.: A sex-linked recessive disorder with retardation of growth, peculiar hair, and focal cerebral and cerebellar degeneration. *Pediatrics* **1962**, *29*, 764-79.

(16) Mercer, J. F. B.: The Molecular Basis of Copper-Transporting Disease. *Trends Mol. Med.* **2001**, *7*, 64-69.

(17) Arguello, J. M.: Identification of ion-selectivity determinants in heavy-metal transport P1B-type ATPases. *J. Membr. Biol.* **2003**, *195*, 93-108.

(18) Arguello, J. M.; Eren, E.; Gonzalez-Guerrero, M.: The structure and function of heavy metal transport P1B-ATPases. *Biometals* **2007**, *20*, 233-48.

(19) Boal, A. K.; Rosenzweig, A. C.: Structural Biology of Copper Trafficking. *Chem. Rev.* **2009**, *109*, 4760–4779.

(20) Yatsunyk, L. A.; Rosenzweig, A. C.: Copper(I) Binding and Transfer by the N-terminus of the Wilson Disease Protein. *J. Biol. Chem.* **2007**, *282*, 8622-8631.

(21) Xiao, Z.; Brose, J.; Schimo, S.; Ackland, S. M.; La Fontaine, S.; Wedd, A. G.: Unification of the copper(I) binding affinities of the metallo-chaperones Atx1, Atox1, and related proteins: detection probes and affinity standards. *J. Biol. Chem.* **2011**, *286*, 11047-55.

(22) Banci, L.; Bertini, I.; Ciofi-Baffoni, S.; Kozyreva, T.; Zovo, K.; Palumaa, P.: Affinity gradients drive copper to cellular destinations. *Nature* **2010**, *465*, 645-8.

(23) Badarau, A.; Dennison, C.: Copper trafficking mechanism of CXXC-containing domains: insight from the pH-dependence of their Cu(I) affinities. *J. Am. Chem. Soc.* **2011**, *133*, 2983-8.

(24) Banci, L.; Bertini, I.; Cantini, F.; Massagni, C.; Migliardi, M.; Rosato, A.: An NMR Study of the Interaction of N-terminal Cytoplasmic Tail of the Wilson Disease Protein with Copper(I)-Hah1. *J. Biol. Chem.* **2009**, *284*, 9354-9360.

(25) Wernimont, A. K.; Huffman, D. L.; Lamb, A. L.; O'Halloran, T. V.; Rosenzweig, A. C.: Structural Basis for Copper Transfer by the Metallochaperone for the Menkes/Wilson Disease Proteins. *Nat. Struct. Biol.* **2000**, *7*, 766-771.

(26) Arnesano, F.; Banci, L.; Bertini, I.; Bonvin, M. J. J.: A Docking Approach to the Study of Copper Trafficking Proteins: Interactions between Metallochaperones and Soluble Domains of Copper ATPases. *Structure* **2004**, *12*, 669-676.

(27) Huffman, D. L.; O'Halloran, T. V.: Function, Structure, and Mechanism of Intracellular

Copper Trafficking Proteins. *Ann. Rev. Biochem.* **2001**, *70*, 677-701.

(28) Rodriguez-Granillo, A.; Crespo, A.; Estrin, D. A.; Wittung-Stafshede, P.: Copper-Transfer Mechanism from the Human Chaperone Atox1 to a Metal-Binding Domain of Wilson Disease Protein. *J. Phys. Chem. B* **2010**, *114*, 3698–3706.

(29) Banci, L.; Bertini, I.; Calderone, V.; Della-Malva, N.; Felli, I. C.; Neri, S.; Pavelkova, A.; Rosato, A.: Copper(I)-mediated protein-protein interactions result from suboptimal interaction surfaces. *Biochem. J.* **2009**, *422*, 37-42.

(30) Walker, J. M.; Huster, D.; Ralle, M.; Morgan, C. T.; Blackburn, N. J.; Lutsenko, S.: The N-Terminal Metal-Binding Site 2 of the Wilson's Disease Protein Play a Key Role in the Transfer of Copper from Atox1. *J. Biol. Chem.* **2004**, *279*, 15376-15384.

(31) Achila, D.; Banci, L.; Bertini, I.; Bunce, J.; Ciofi-Baffoni, S.; Huffman, D. L.: Structure of human Wilson protein domains 5 and 6 and their interplay with domain 4 and the copper chaperone HAH1 in copper uptake. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 5729-5734.

(32) Rodriguez-Granillo, A.; Crespo, A.; Wittung-Stafshede, P.: Conformational Dynamics of Metal-Binding Domains in Wilson Disease Protein: Molecular Insights into Selective Copper Transfer. *Biochemistry* **2009**, *48*, 5849-5863.

(33) Rodriguez-Granillo, A.; Crespo, A.; Wittung-Stafshede, P.: Interdomain interactions modulate collective dynamics of the metal-binding domains in the Wilson disease protein. *J. Phys. Chem. B* **2010**, *114*, 1836-48.

(34) van Dongen, E. M.; Klomp, L. W.; Merkx, M.: Copper-dependent protein-protein interactions studied by yeast two-hybrid analysis. *Biochem. Biophys. Res. Commun.* **2004**, *323*, 789-95.

(35) Banci, L.; Bertini, I.; Cantini, F.; Della-Malva, N.; Migliardi, M.; Rosato, A.: The Different Intermolecular Interactions of the Soluble Copper-Binding Domains of the Menkes Protein, ATP7A. *J. Biol. Chem.* **2007**, *282*, 23140-23146.

(36) Fatemi, N.; Korzhnev, D. M.; Velyvis, A.; Sarkar, B.; Forman-Kay, J. D.: NMR Characterization of Copper-Binding Domains 4-6 of ATP7B. *Biochemistry* **2010**, 8468–8477.

(37) Iida, M.; Terada, K.; Sambongi, Y.; Wakabayashi, T.; Miura, N.; Koyama, K.; Futai, M.; Sugiyama, T.: Analysis of functional domains of Wilson disease protein (ATP7B) in Saccharomyces cerevisiae. *FEBS Lett.* **1998**, *428*, 281-5.

(38) Forbes, J. R.; Hsi, G.; Cox, D. W.: Role of the Copper-Binding Domain in the Copper Transport Function of ATP7B, the P-type ATPase Defective in Wilson Disease. *J. Biol. Chem.* **1999**, *274*, 12408-12413.

(39) Cater, M. A.; Forbes, J.; La Fontaine, S.; Cox, D.; Mercer, J. F.: Intracellular trafficking of the human Wilson protein: the role of the six N-terminal metal-binding sites. *Biochem. J.* **2004**, *380*, 805-13.

(40) Gonzalez-Guerrero, M.; Arguello, J. M.: Mechanism of Cu+-transporting ATPases: soluble Cu+ chaperones directly transfer Cu+ to transmembrane transport sites. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 5992-7.

(41) DiDonato, M.; Hsu, H. F.; Narindrasorasak, S.; Que, L., Jr.; Sarkar, B.: Copper-induced conformational changes in the N-terminal domain of the Wilson disease copper-transporting ATPase. *Biochemistry* **2000**, *39*, 1890-6.

(42) Tsivkovskii, R.; MacArthur, B. C.; Lutsenko, S.: The Lys1010-Lys1325 fragment of the Wilson's disease protein binds nucleotides and interacts with the N-terminal domain of this protein in a copper-dependent manner. *J. Biol. Chem.* **2001**, *276*, 2234-42.

(43) Leshane, E. S.; Shinde, U.; Walker, J. M.; Barry, A. N.; Blackburn, N. J.; Ralle, M.; Lutsenko, S.: Interactions between copper-binding sites determine the redox status and conformation of the regulatory N-terminal domain of ATP7B. *J. Biol. Chem.* **2010**, *285*, 6327-6336.

(44) Wu, C. C.; Rice, W. J.; Stokes, D. L.: Structure of a copper pump suggests a regulatory role for its metal-binding domain. *Structure* **2008**, *16*, 976-85.

(45) Gourdon, P.; Liu, X. Y.; Skjorringe, T.; Morth, J. P.; Moller, L. B.; Pedersen, B. P.; Nissen, P.: Crystal structure of a copper-transporting PIB-type ATPase. *Nature* **2011**, *475*, 59-64.

(46) Vanderwerf, S. M.; Cooper, M. J.; Stetsenko, I. V.; Lutsenko, S.: Copper specifically regulates intracellular phosphorylation of the Wilson's disease protein, a human copper-transporting ATPase. *J. Biol. Chem.* **2001**, *276*, 36289-94.

(47) Bartee, M. Y.; Ralle, M.; Lutsenko, S.: The loop connecting metal-binding domains 3 and 4 of ATP7B is a target of a kinase-mediated phosphorylation. *Biochemistry* **2009**, *48*, 5573-81.

(48) Pilankatta, R.; Lewis, D.; Adams, C. M.; Inesi, G.: High yield heterologous expression of wild-type and mutant Cu+-ATPase (ATP7B, Wilson disease protein) for functional characterization of catalytic activity and serine residues undergoing copper-dependent phosphorylation. *J. Biol. Chem.* **2009**, *284*, 21307-16.

(49) Veldhuis, N. A.; Valova, V. A.; Gaeth, A. P.; Palstra, N.; Hannan, K. M.; Michell, B. J.; Kelly, L. E.; Jennings, I.; Kemp, B. E.; Pearson, R. B.; Robinson, P. J.; Camakaris, J.: Phosphorylation regulates copper-responsive trafficking of the Menkes copper transporting P-type ATPase. *Int. J. Biochem. Cell Biol.* **2009**, *41*, 2403-12.

(50) Pilankatta, R.; Lewis, D.; Inesi, G.: Involvement of protein kinase D in expression and trafficking of ATP7B (copper ATPase). *J. Biol. Chem.* **2011**, *286*, 7389-96.

(51) Walker, J. M.; Tsivkovskii, R.; Lutsenko, S.: Metallochaperone Atox1 Transfers Copper to the NH2-Terminal Domain of the Wilson's Disease Protein and Regulates Its Catalytic Activity. *J. Biol. Chem.* **2002**, *277*, 27953-27959.

(52) Multhaup, G.; Strausak, D.; Bissig, K.-D.; Solioz, M.: Interaction of the CopZ Copper Chaperone with the CopA Copper ATPase of *Enterococcus hirae* Assessed by Surface Plasmon Resonance. *Biochem. Biophys. Res. Commun.* **2001**, *288*, 172-177.

(53) Strausak, D.; Howies, M. K.; Firth, S. D.; Schlicksupp, A.; Pipkorn, R.; Multhaup, G.; Mercer, J. F. B.: Kinetic Analysis of the Interaction of the Copper Chaperone Atox1 with the Metal Binding Sites of the Menkes Protein. *J. Biol. Chem.* **2003**, *278*, 20821-20827.

(54) Arnesano, F.; Banci, L.; Bertini, I.; Cantini, F.; Ciofi-Baffoni, S.; Huffman, D. L.; O'Halloran, T. V.: Characterization of the Binding Interfaces between the Copper Chaperone Atx1 and the First Cytosolic Domain of Ccc2 ATPase. *J. Biol. Chem.* **2001**, *276*, 41365-41376.

(55) Banci, L.; Bertini, I.; Cantini, F.; Felli, I. C.; Gonnelli, L.; Hadjiliadis, N.; Pierattelli, R.; Rosato, A.; Voulgaris, P.: The Atx1-Ccc2 Complex is a Metal-Mediated Protein-Protein Interaction. *Nat. Chem. Biol.* **2006**, *2*, 367-368.

(56) Banci, L.; Bertini, I.; Cantini, F.; Chasapis, C. T.; Hadjiliadis, N.; Rosato, A.: A NMR Study of the Interactions of a Three-Domain Construct of ATP7A with Copper(I) and Copper(I)-Hah1: The Interplay of Domains. *J. Biol. Chem.* **2005**, *280*, 38259-38263.

(57) Banci, L.; Bertini, I.; Francesca, C.; Rosenzweig, A. C.; Yatsunyk, L. A.: Metal Binding Domains 3 and 4 of the Wilson Disease Protein: Solution Structure and Interaction with the Copper(I) Chaperone Hah1. *Biochemistry* **2008**, *47*, 7423-7429.

(58) Badarau, A.; Firbank, S. J.; McCarthy, A. A.; Banfield, M. J.; Dennison, C.: Visualizing the Metal-Binding Versatility of Copper Trafficking Sites. *Biochemistry* **2010**, 7798–7810.

(59) Zimmerman, M.; Clarke, O.; Gulbis, J. M.; Keizer, D. W.; Jarvis, R. S.; Cobbett, C. S.; Hinds, M. G.; Xiao, Z.; Wedd, A. G.: Metal Binding Affinities of *Arabidopsis* Zinc and Copper Transporters: Selectivities Match the Relative, but Not the Absolute, Affinities of their Amino-Terminal Domains. *Biochemistry* **2009**, 11640–11654.

(60) Alvarez, H. M.; Xue, Y.; Robinson, C. D.; Canalizo-Hernández, M. A.; Marvin, R. G.; Kelly, R. A.; Mondragón, A.; Penner-Hahn, J. E.; O'Halloran, T. V.: Tetrathiomolybdate Inhibits Copper Trafficking Proteins Through Metal Cluster Formation. *Science* **2009**, *327*, 331-334.

(61) Hussain, F.; Rodriguez-Granillo, A.; Wittung-Stafshede, P.: Lysine-60 in Copper Chaperone Atox1 Plays an Essential Role in Adduct Formation with a Target Wilson Disease Domain. *J. Am. Chem. Soc.* **2009**, *131*, 16371-16373.

(62) Ha, T.: Single-Molecule Fluorescence Resonance Energy Transfer. *Methods* **2001**, *25*, 78-86.

(63) Chiu, D. T.; Wilson, C. F.; Karlsson, A.; Danielsson, A.; Lundqvist, A.; Strömberg, A.; Ryttsén, F.; Davidson, M.; Nordholm, S.; Orwar, O.; Zare, R. N.: Manipulating the biochemical nanoenvironment around single molecules contained within vesicles. *Chem. Phys.* **1999**, *247*, 133-139.

(64) Boukobza, E.; Sonnenfeld, A.; Haran, G.: Immobilization in Surface-Tethered Lipid Vesicles as a New Tool for Single Biomolecule Spectroscopy. *J. Phys. Chem. B.* **2001**, *105*, 12165-12170.

(65) Okumus, B.; Wilson, T. J.; Lilley, D. M. J.; Ha, T.: Vesicle Encapsulation Studies Reveal that Single Molecule Ribozyme Heterogeneities Are Intrinsic. *Biophys. J.* **2004**, *87*, 2798-2806.

(66) Benitez, J. J.; Keller, A. M.; Chen, P.: Nanovesicle Trapping for Studying Weak Protein Interactions by Single-Molecule FRET. *Methods Enzymol.* **2010**, *472*, 41-60.

(67) Lakowicz, J. R.: *Principles of fluorescence spectroscopy*; 3rd ed.; Springer: New York, 2006.

(68) Benitez, J. J.; Keller, A. M.; Ochieng, P.; Yatsunyk, L. A.; Huffman, D. L.; Rosenzweig, A. C.; Chen, P.: Probing Real-time Transient Metallochaperone-Target Protein Interactions at the Single-Molecule Level with Nanovesicle Trapping. *J. Am. Chem. Soc.* **2008**, *130*, 2446-2447.

(69) Benitez, J. J.; Keller, A. M.; Ochieng, P.; Yatsunyk, L. A.; Huffman, D. L.; Rosenzweig, A. C.; Chen, P.: Correction/Addition: Probing Real-time Transient Metallochaperone-Target Protein Interactions at the Single-Molecule Level with Nanovesicle Trapping. *J. Am. Chem. Soc.* **2009**, *131*, 871.

(70) Benitez, J. J.; Keller, A. M.; Huffman, D. L.; Yatsunyk, L.; Rosenzweig, A. C.; Chen, P.: Relating Dynamic Protein Interactions of Metallochaperones with Metal Transfer at the Single-Molecule Level. *Faraday Discuss.* **2011**, *148*, 71-82.

(71) Wernimont, A. K.; Yatsunyk, L. A.; Rosenzweig, A. C.: Binding of Copper(I) by the Wilson Disease Protein and Its Copper Chaperone. *J. Biol. Chem.* **2004**, *279*, 12269-12276.

(72) Brenner, A. J.; Harris, E. D.: A Quantitative Test for Copper Using Bicinchoninic Acid. *Anal. Biochem.* **1995**, *226*, 80-84.

(73) Cordes, T.; Vogelsang, J.; Tinnefeld, P.: On the mechanism of Trolox as antiblinking and antibleaching reagent. *J. Am. Chem. Soc.* **2009**, *131*, 5018-9.

(74) Shi, X.; Lim, J.; Ha, T.: Acidification of the oxygen scavenging system in single-molecule fluorescence studies: in situ sensing with a ratiometric dual-emission probe. *Anal. Chem.* **2010**, *82*, 6132-8.

(75) Aitken, C. E.; Marshall, R. A.; Puglisi, J. D.: An oxygen scavenging system for improvement of dye stability in single-molecule fluorescence experiments. *Biophys. J.* **2008**, *94*, 1826-35.

(76) Sarkar, S. K.; Andoy, N. M.; Benitez, J. J.; Chen, P. R.; Kong, J. S.; He, C.; Chen, P.: Engineered Holliday Junctions as Single-Molecule Reporters for Protein-DNA Interactions with Application to a MerR-Family Regulator. *J. Am. Chem. Soc.* **2007**, *129*, 12461-12467.

(77) Chung, S. H.; Kennedy, R. A.: Forward-backward non-linear filtering technique for extracting small biological signals from noise. *J. Neurosci. Methods* **1991**, *40*, 71-86.

(78) Haran, G.: Noise reduction in single-molecule fluorescence trajectories of folding proteins. *Chem. Phys.* **2004**, *307*, 137-145.

(79) Nan, X.; Sims, P. A.; Chen, P.; Xie, X. S.: Observation of Individual Microtubule Motor Steps in Living Cells with Endocytosed Quantum Dots. *J. Phys. Chem. B* **2005**, *109*, 24220-24224.

(80) Anastassopoulou, I.; Banci, L.; Bertini, I.; Cantini, F.; Katsari, E.; Rosato, A.: Solution Structure of the Apo and Copper(I)-Loaded Human Metallochaperone HAH1. *Biochemistry* **2004**, *43*, 13046-13053.

(81) Krissinel, E.; Henrick, K.: Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **2007**, *372*, 774-97.

(82) Guex, N.; Peitsch, M. C.: SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis* **1997**, *18*, 2714-23.

(83) Scott, W. R. P.; Hünenberger, P. H.; Tironi, I. G.; Mark, A. E.; Billeter, S. R.; Fennen, J.; Torda, A. E.; Huber, T.; Krüger, P.; van Gunsteren, W. F.: The GROMOS Biomolecular Simulation Program Package. *J. Phys. Chem. A* **1999**, *103*, 3596-3607.

(84) *Protein-Protein Recognition*; Kleanthous, C., Ed.; Oxford University Press: Oxford, 2000.

(85) Portnoy, M. E.; Rosenzweig, A. C.; Rae, T. D.; Huffman, D. L.; O'Halloran, T. V.; Culotta, V. C.: Structure-Function Analyses of the ATX1 Metallochaperone. *J. Biol. Chem.* **1999**, *274*, 15041-15045.

(86) Hohng, S.; Joo, C.; Ha, T.: Single-Molecule Three-Color FRET. *Biophys. J.* **2004**, *87*, 1328-1337.

(87) Lee, N. K.; Kapanidis, A. N.; Koh, H. R.; Korlann, Y.; Ho, S. O.; Kim, Y.; Gassman, N.; Kim, S. K.; Weiss, S.: Three-color alternating-laser excitation of single molecules: monitoring multiple interactions and distances. *Biophys. J.* **2007**, *92*, 303-12.

(88) Lee, S.-K.; Lee, J. C.; Hohng, S.: Single-Molecule Three-Color FRET with Both Negligible Spectral Overlap and Long Observation Time. *PloS One* **2010**, *5*, e12270.

(89) Martinez-Finley, E. J.; Chakraborty, S.; Fretham, S. J.; Aschner, M.: Cellular transport and homeostasis of essential and nonessential metals. *Metallomics : integrated biometal science* **2012**.

(90) Padovani, D.; Labunska, T.; Palfey, B. A.; Ballou, D. P.; Banerjee, R.: Adenosyltransferase tailors and delivers coenzyme B12. *Nat. Chem. Biol.* **2008**, *4*, 194-6.

(91) Padovani, D.; Banerjee, R.: A Rotary Mechanism for Coenzyme B$_{12}$ Synthesis by Adenosyltransferase. *Biochemistry* **2009**, *48*, 5350-5356.

(92) Banerjee, R.; Gherasim, C.; Padovani, D.: The Tinker, Tailor, Soldier in Intracellular B$_{12}$ Trafficking. *Curr. Opin. Chem. Biol.* **2009**, *13*, 484-491.

(93) Moon, J. H.; Park, A. K.; Jang, E. H.; Kim, H. S.; Chi, Y. M.: Crystal structure of a PduO-type ATP:cobalamin adenosyltransferase from Burkholderia thailandensis. *Proteins* **2008**, *72*, 1066-70.

(94) Maurice, M. S.; Mera, P.; Park, K.; Brunold, T. C.; Escalante-Semerena, J. C.; Rayment, I.: Structural characterization of a human-type corrinoid adenosyltransferase confirms that coenzyme B-12 is synthesized through a four-coordinate intermediate. *Biochemistry* **2008**, *47*, 5755-5766.

(95) Schubert, H. L.; Hill, C. P.: Structure of ATP-bound human ATP : cobalamin adenosyltransferase. *Biochemistry* **2006**, *45*, 15188-15196.

(96) McKinney, S. A.; Joo, C.; Ha, T.: Analysis of single-molecule FRET trajectories using hidden Markov modeling. *Biophys. J.* **2006**, *91*, 1941-1951.

(97) Lee, T. H.: Extracting Kinetics Information from Single-Molecule Fluorescence Resonance Energy Transfer Data Using Hidden Markov Models. *J. Phys. Chem. B* **2009**, *113*, 11535-11542.

(98) Eichman, B. F.; Vargason, J. M.; Mooers, B. H. M.; Ho, P. S.: The Holliday junction in an inverted repeat DNA sequence: sequence effects on the structure of four-way junctions. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 3971-3976.

(99) Lilley, D. M. J.: Structures of Helical Junctions in Nucleic Acids. *Q. Rev. Biophys.* **2000**, *33*, 109-159.

(100) McKinney, S. A.; Declais, A. C.; Lilley, D. M. J.; Ha, T.: Structural Dynamics of Individual Holliday Junctions. *Nat. Struct. Biol.* **2003**, *10*, 93-97.

(101) Andoy, N. M.; Sarkar, S. K.; Wang, Q.; Panda, D.; Benitez, J. J.; Kalininskiy, A.; Chen, P.: Single-Molecule Study of Metalloregulator CueR-DNA Interactions Using Engineered Holliday Junctions. *Biophys. J.* **2009**, *97*, 97, 844-852.

(102) Sabanayagam, C. R.; Eid, J. S.; Meller, A.: Using fluorescence resonance energy transfer to measure distances along individual DNA molecules: Corrections due to nonideal transfer. *J. Chem. Phys.* **2005**, *122*.

(103) Schuler, B.; Lipman, E. A.; Steinbach, P. J.; Kumke, M.; Eaton, W. A.: Polyproline and the "spectroscopic ruler" revisited with single-molecule fluorescence. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 2754-9.

(104) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M.: Charmm - a Program for Macromolecular Energy, Minimization, and Dynamics Calculations. *J. Comput. Chem.* **1983**, *4*, 187-217.

(105) MacKerell, A. D.; Bashford, D.; Bellott, M.; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M.: All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B* **1998**, *102*, 3586-3616.

(106) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L.: Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926-935.

(107) Darden, T.; York, D.; Pedersen, L.: Particle Mesh Ewald - an N.Log(N) Method for Ewald Sums in Large Systems. *J. Chem. Phys.* **1993**, *98*, 10089-10092.

(108) Vangunsteren, W. F.; Berendsen, H. J. C.: Algorithms for Macromolecular Dynamics and Constraint Dynamics. *Mol. Phys.* **1977**, *34*, 1311-1327.