STRUCTURAL AND BIOPHYSICAL INVESTIGATION OF CALMODULIN BINDING TO THE THIRD DOMAIN OF HUMAN PRP40A

By

Brian James Ferrer

Thesis
Submitted to the Faculty of the
Graduate School of Vanderbilt University
in partial fulfillment of the requirements
for the degree of

MASTER OF SCIENCE
in
Chemical and Physical Biology

June 30th, 2022
Nashville, Tennessee

Approved:

Walter J. Chazin, Ph.D.

Breann Brown, Ph.D.

To my mother for being both parents.

To my friends and colleagues, thank you for making me who I am today.

And to the friends I have yet to meet, I look forward to our encounter.

ACKNOWLEDGEMENTS

To begin, I would like to thank Dr. Walter Chazin. Getting a master's degree was never my plan, and I am not even sure what my plan is now. However, I'd be in a far more confusing place now if Dr. Chazin hadn't given me the opportunity to figure myself out over the past year and half. In that same vein, I have to say thank you to all the members of the Chazin Lab! And a special shout out to Dr. Randy Perera for being an awesome officemate. I also must thank Dr. Swati Balakrishnan for being my first mentor in the lab and an important person in helping me put the last few pieces of this project together! Lastly, I want to thank Kaitlyn Gallagher for dealing with my constant struggles, complaints, and lifelong learning experience with running the ITC instrument.

Next, I want to thank all the amazing friends I've made in my cohort. Starting graduate school in a pandemic was awful but having the opportunity to meet such amazing scientists and people made it slightly less awful. I have to give a special shoutout to my fellow CSB Megalab trainee Tyler Huth. Thank you for listening to all my awful ideas and somewhat useful advice, and you better keep texting me! I also want to mention Carl Stone, Matthew Dungan, and Megan Stanchfield for being some of my first friends at Vanderbilt. Thank you for continuing to be my friend even though I'm going in a different direction in life. I expect to see you all again at least once annually until we're all on separate continents or until one you of wins the Nobel prize.

I also couldn't resist making a separate paragraph for one of my closest friends Zach Sanchez. Thanks for braving every literal winter storm with me since we've been in Tennessee! You've made me a better scientist and have challenged me to be a better person outside the lab. I never expected to make a such a close friend in grad school, although I think cohabitating certainly encourages friendship. I can't wait to see how you grow as a scientist in the years to come.

Thank you to Dr. Breann Brown for being my first research mentor at Vanderbilt and an invaluable resource along my journey (as well as being a reviewer for my thesis!)

I would be remiss to not also thank my undergraduate advisor Dr. Sam Gellman and my post-bac advisor Dr. Jim Keck. You've both supported me along this journey and nothing about my adventure in academia has been straightforward or fair. I also haven't made the smartest decisions since venturing off on my own, but I will forever be grateful for your unconditional support.

Last of all, I have to say thank you to my mom and stepdad Dennis. Even in our most destitute times, you've never stopped supporting me in school or simply me as a person.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

1.1 Calcium Regulation and Signaling.

Calcium (Ca^{2+}) is a central element in numerous cellular processes, serving both as a ubiquitous second-messenger signal for receptor-mediated responses across the cell[1] and as a metal cofactor for many proteins. Ca^{2+}-dependent regulation and signaling are therefore important regulators for human health.[2,3] Ca^{2+} regulation modulates a range of specific functions across the cell (**Figure 1.1**). Notably, levels of Ca^{2+} fluctuate within cellular compartments and across the cell membrane. Consequently, Ca^{2+} signaling requires specific interacting partners including calcium-binding proteins (CaBPs), channels, and transporters.[4] The number and types of specific interactions involving Ca^{2+} highlight the importance of this metal in biological homeostasis and how studying these interactions within these processes can be a valuable tool for understanding normal physiology, pathology and therapeutic development.

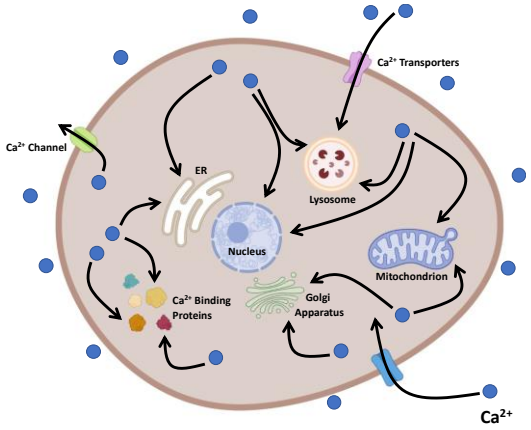

Figure 1.1. Diagram of Ca^{2+} signaling across the cell.
Ca^{2+} signaling occurs at various sites across the cell. Image created with BioRender.com

Ca^{2+}-binding proteins.

The interaction of Ca^{2+} with a signaling target or localization to its cellular compartment is often mediated by Ca^{2+}-binding proteins (CaBPs). The EF-hand family constitutes one major class of CaBP.[5] The term "EF-hand" was coined upon inspection of the x-ray crystal structure of parvalbumin, in which helix E and helix F are separated by a loop with partial β-sheet character that binds Ca^{2+} ions (and other divalent metals such as Mg^{2+}) (**Figure 1.2**).[6] The structure of this motif is highly conserved across homologs.[7]

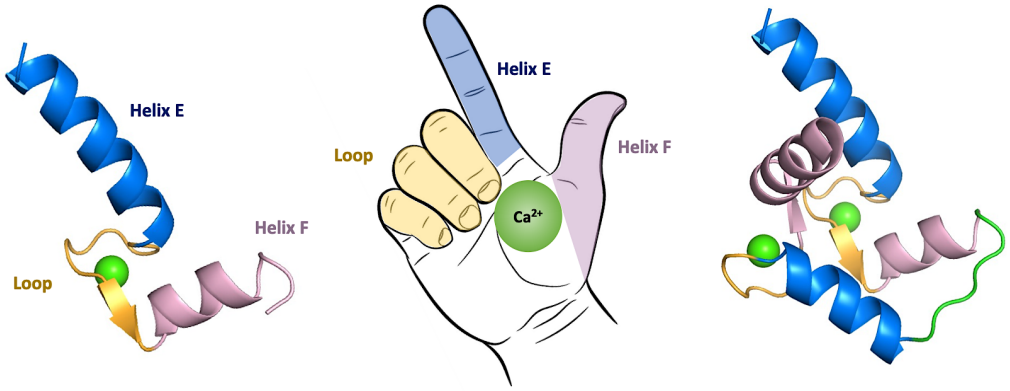

Figure 1.2. An example of an EF-hand motif from the human calmodulin crystal structure. (**Left**) An EF-hand motif extracted from the crystal structure of human calmodulin (PDB 1CLL). (**Center**) a cartoon diagram of EF-hand structure. (**Right**) An example of an EF-domain which is comprised of two EF-hand motifs. Coloring scheme: Helix E in **blue,** metal binding loop in **orange**, helix F in **pink,** and Ca^{2+} in **green.**

Although EF-hand proteins share many features, they serve specific functions and consequently, many EF-hand CaBPs exist. A short list of the most common EF-hand CaBPs is reported here (**Table 1.1**). In summary, EF-hand CaBPs regulate and buffer Ca^{2+} signals, or sense and transduce Ca^{2+} signals to regulate a wide range of biochemical pathways.

Table 1.1. A brief list of examples of EF-hand calcium-binding proteins.

Protein	Function	Reference
Calmodulin	Ubiquitous Ca^{2+}-sensor → Ca^{2+}-dependent enzyme activation and regulation of many biochemical pathways	[3, 8]
Troponin C	Specific Ca^{2+}-sensor → Ca2+-dependent activation of (cardiac) muscle contraction	[9, 10]
Calbindin-D_{28k}	Ca^{2+} signal modifier and sensor → Vitamin D dependent; shown to regulate signal transduction cascades	[11]
Calreticulin	Ca^{2+} metallochaperone in the endoplasmic reticulum	[12]
Centrin 2	Specific Ca^{2+}-sensor → regulation of the centrosome and involved in nucleotide excision repair	[13]
S100 Proteins	Ca^{2+} signal modifiers and sensors → intra- and extra-cellular functions including innate immune response, regulation of inflammation, cell metabolism, apoptosis, and Ca^{2+} homeostasis	[14-16]

1.2 Ca^{2+}-Binding Proteins in Human Health and Disease.

In addition to homeostatic control, many Ca^{2+}-binding proteins have also been linked to diseases. For example, S100A8/A9 (collectively known as calprotectin) is upregulated and is actively released from neutrophils to mediate certain immune responses and stimulate inflammation.[17] In Parkinson's Disease, there seems to be an effect on expression of Ca_v1 calcium ion channels as well as other CaBPs that are sensitive to changes in these channels.[18] In cancer, expression of S100P, S100A8, and S100A14 appears to be upregulated and contributes to anticancer drug resistance.[19-21] Additionally, the penta-EF-hand protein sorcin appears to be involved in multidrug resistance in certain cancer cells.[22] In summary, CaBPs appear to play in a role in disease progression, serve as biomarkers, and represent a potential target for therapeutic intervention. Consequently, there is a growing list of therapies that act upon CaBPs or intersect with pathways affected by CaBPs (**Table 1.2**).

Table 1.2. Examples of therapeutics acting upon calcium-binding proteins in human diseases.

Therapeutic	Target & Use	Reference
Chlorpromazine	Calmodulin and calmodulin-like protein inhibitor; primary use as an antipsychotic; off-label use as anti-nausea drug	[23]
Diltiazam	Calcium ion channel blocker in cardiac muscle tissue; primary use is to lower blood pressure, hypertension drug	[24]
Purfalcamine	Ca^{2+}-dependent kinase inhibitor in P. falciparum; antiparasitic	[25]
Cyclosporin A	Inhibits phosphatase activity of calcineurin; inhibits T cell activation; immunosuppressant	[26]
Thapsigargin	Inhibits sarco/endoplasmic reticulum Ca^{2+}-ATPase (SERCA); prevents influx of Ca^{2+}; ER stressor → cell death; anti-cancer	[27]
Levetiracetam	Inhibits N-type calcium ion channels; prevents excitatory neurotransmission; antiepileptic	[28]
Flunarizine	Calcium ion channel blocker in neuronal cells; migraine prophylaxis; some evidence for calmodulin inhibition to	[29]
Calcitonin	Inhibits osteoclasts; decreases bone loss in osteoporosis	[30]

Therapeutic intervention by targeting proteins involved in these pathways appears to be an effective approach. There are ongoing efforts to discover new therapeutics targeting CaBPs or associated proteins by high-throughput screening and drug discovery.[31,32] In addition, there have been efforts to engineer CaBP variants directly as therapeutics.[33,34] For example, targeting CaBPs in neurodegenerative diseases (NDs) remains underexplored, despite the fact that calcium dysregulation is a major hallmark of ND progression.[35] This is in part due to the complex nature of protein-protein interactions (PPIs) in these diseases. However, the mechanism of disease progression in neurodegenerative diseases is dependent on more than one PPI, which makes identifying a critical target complicated.[36] At the same time, if a coordinated network of multiple proteins is responsible for disease progression, targeting a single interaction by therapeutic intervention may not ameliorate a disease phenotype. Hence, a more detailed understanding of molecular interactions of disease interactomes may help improve therapeutic and diagnostic development. In particular, the focus for this thesis is the investigation of PPIs between **calcium sensors and other proteins associated with Huntington's Disease** (**HD**).

1.3 An Introduction to Huntington's Disease.

Huntington's Disease is a rare neurodegenerative disease that leads to decreased cognitive function, memory loss, and impaired motor control.[37] On a molecular level, the disease is triggered by a mutant form of the human huntingtin protein (mHTT). mHTT is prone to aggregation within neuronal cells and augments PPIs normally regulated by wildtype HTT (WT HTT). It is thought that WT HTT is important for long-term memory formation, although this has not been well characterized in humans.[37]

```
      10        20        30        40        50
MATLEKLMKA FESLKSFQQQ QQQQQQQQQQ QQQQQQQQPP PPPPPPPPPQ
      60        70        80        90       100
LPQPPPQAQP LLPQPQPPPP PPPPPPGPAV AEEPLHRPKK ELSATKKDRV
```

Figure 1.3. N-terminal WT HTT Amino Acid Sequence.
The first 100 residues of WT HTT are shown. Native glutamine residues are highlighted in **green**.

A key structural feature of HTT is that it has a polyglutamine repeat tract at its N-terminus (**Figure 1.3**). Mutations can occur in this region due to the polyglutamine sequence being encoded by a long repeat sequence of cytosine-adenine-guanine (CAG) bases.[38] These repeats are prone to DNA

replication errors and expansion of the CAG repeats is plausible.[39] The number of CAG repeats ultimately determines the likelihood of developing HD (**Table 1.3**). Normally, WT HTT has $\leq$ 26 CAG repeats. There is some level of tolerability for expanded CAG repeats in patients with 27-35 repeats. However, after this threshold, the risk of developing HD is severe. Notably, patients with 36-39 repeats are classified as low-penetrance phenotype in which they may or may not develop symptoms later in life, but almost certainly will pass on to their offspring. Lastly, patients with $\geq$ 39 CAG repeats/polyglutamine residues is the clinical cutoff for being diagnosed with HD.[38,39]

Table 1.3. Risk of HD Development: Tolerability of CAG/Glutamine Repeats in HTT.

Number of CAG Repeats	HD Likelihood	Reference
$\leq$ 26	Homeostatic phenotype; WT HTT carrier; unlikely to develop HD	[40]
27-35	Intermediate *HTT* allele; WT HTT carrier; won't show symptoms but possibly at-risk for HD development if enough mutations or replication errors occur; more susceptible at older ages	
$\geq$ 36	HD phenotype; mHTT carrier; high likelihood for inheritance; generally, this is the clinical cutoff for diagnosis	
36-39	Low penetrance phenotype; mHTT carrier; less-severe HD symptoms but still high likelihood for inheritance	
$\geq$ 80	Rare, but possible to have excessive CAG expansions; atypical clinical manifestation; most severe phenotypes especially in younger patients; if present, deleterious juvenile HD development likely	

On a clinical level, HD currently affects approximately 0.1-0.7 per 100,000 people (approximately 0.001-0.007% of the general population) as of 2021.[41] HD is typically seen in older patients but can occur at any age depending on how the length of expanded polyglutamine tract in mHTT, as noted previously. However, the major contribution to develop HD is genetics since the disease is an autosomal dominant condition and inheriting an at-risk or mutant allele of the HTT gene from a single parent is the major determinant for disease development.[42] Additionally, genetic screening can help assess a patient's risk for developing symptoms.[43] In terms of treatment, no direct cure for HD exists, although some therapeutics exist for reducing symptoms and there are several therapeutics in the pipeline currently.[44] In brief, target therapies are being developed to tackle HD at different molecular levels of expression and signaling (**Figure 1.4**). To date, no clinically approved PPI inhibitor for HD has been developed.

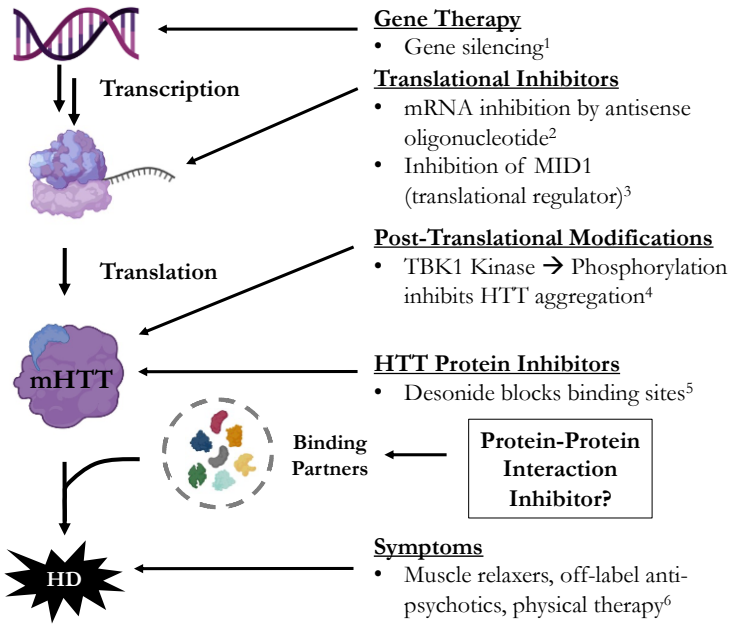

Figure 1.4. Current therapeutic strategies targeting the expression of mHTT and treatment of HD. Existing therapies are designed to target expression of mHTT at the DNA and mRNA levels.[45-48] One recent study identified a novel mHTT inhibitor.[49] Some therapies are used to treat symptoms of HD but cannot cure the disease directly.[50] Images created with BioRender.com

A study conducted in 2014 reported that HTT has about 100 interacting partners directly interacting with HTT and about 2000 secondary partners interacting with primary partners.[51] This network consists of direct interactions with the HTT protein as well as with other proteins interacting downstream of HTT. In a pool of thousands of direct HTT-interacting proteins and associated proteins, disruption of a single PPI could be difficult to target and simultaneously may not elicit a therapeutic response. To that end, research into the HTT interactome may help identify the molecular interactions underpinning HD and potentially lead to novel therapeutic development.

Huntingtin Protein.

The WT HTT protein is ~3000 residues with a mass of 346.7 kDa. The domain architecture of HTT is broadly defined as having N- and C-terminal domains of Huntingtin, Elongation factor 3, protein phosphatase 2A, TOR1 (HEAT) helical repeats, separated by a bridge domain linker.[52-53] Additionally, a cryoEM structure of the WT HTT protein in complex with Huntingtin Associated Protein 40 (HAP40) has been determined, although the authors weren't able to get resolution for the N-terminal glutamine rich region (**Figure 1.5**).[52] Based on this structure, there appears to be some

fairly ordered regions constituting the N- and C-terminal HEAT domains interacting with HAP40. Due to the size and structure of this protein, it is unsurprising that HTT can bind to numerous interacting partners. Furthermore, the dynamic nature of the N-terminus can be further promoted by polyglutamine expansions, which would contribute to the aggregation-prone behavior of mHTT.

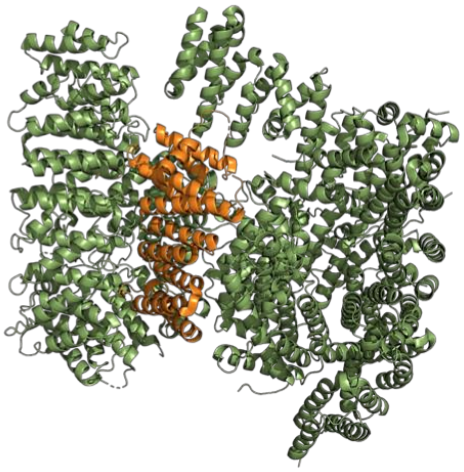

Figure 1.5. CryoEM structural model of WT HTT.
A structure of WT HTT was determined by cryoEM. HTT shown in **green.** HAP40 shown in **orange**. (PDB 6X9O).

1.4 Protein-Protein Interactions Associated with HTT.

As noted above, ~100 different proteins appear to have interactions with HTT. A select number of these interacting proteins is shown in a functional protein network generated by the STRINGS server algorithm which uses a combination of predicted and experimental data (**Figure 1.6**).[54] Among these are an RNA splicing factor Prp40, which is predicted to have a strong interaction with HTT, and two structurally related EF-hand Ca^{2+} sensors: human centrin 2 (hCen2), which is predicted to have a weak interaction with HTT, and calmodulin (CaM), which has been reported to interact with mHTT with high affinity[55,56]. Previous work in the Chazin Lab demonstrated that hCen2 interacts with Prp40A through its FF_3 domain.[57] In brief, our lab used isothermal titration calorimetry (ITC) and a combination of spectroscopic techniques to establish that Prp40A-FF_3 interacts with hCen2. The data revealed a stoichiometry of one-to-one and a binding affinity (K_d) of 278 ± 31 nM. Building off this work and the similarities of EF-hand Ca^{2+} sensors, we developed our hypothesis that CaM and hPrp40A interact with each other in a calcium-dependent manner.

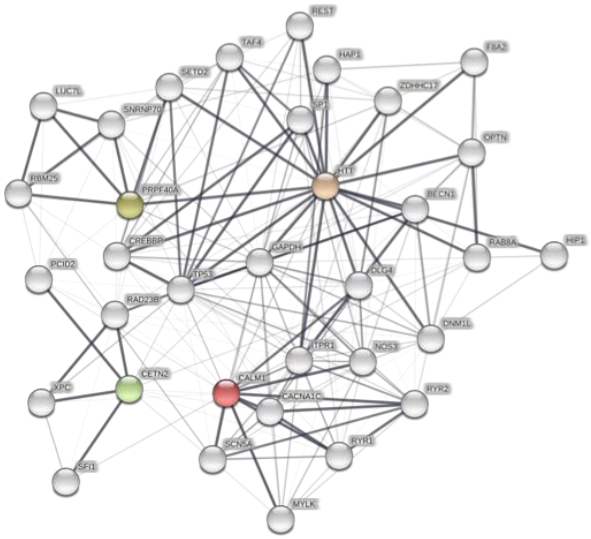

Figure 1.6. A selection of the predicted HTT interactome.
HTT in **light orange**, calmodulin (encoded by CALM1) in **red**, Prp40A (encoded by PRPF40A) in **olive green**, and centrin 2 (encoded by CETN2) in **light green**. Thickness of the connecting line indicates confidence in the interaction. Generated with the algorithm-based STRING Server (https://string-db.org).

The Calcium Sensor Calmodulin.

CaM is the prototype EF-hand Ca^{2+} sensor, comprised of four EF-hand motifs organized into N- and C-terminal globular domains (**Figure 1.7**). CaM is encoded by three genes (*CALM1*, *CALM2*, *CALM3*), which have slightly different nucleotide sequences but ultimately code for the same amino acid sequence, emphasizing the biological importance of this protein.[58] CaM has many binding partners across the cell that facilitate a wide range of calcium-dependent processes as alluded to above. Importantly, the two domains of CaM are connected by a flexible linker, which enables CaM to bind partner proteins in multiple modes: engaging the N-terminus, the C-terminus, or both domains simultaneously.[59]

Figure 1.7. Crystal structure of Ca^{2+}-loaded CaM.
CaM in **blue**, Ca^{2+} ions in **grey**. (PDB 1CLL).

CaM binding to Munc13-1 is an example of CaM engaging the target with both domains (and one that will be revisited later.) Munc13-1 is a protein involved in SNARE assembly and important for regulating vesicle fusion in the cell. A crystal structure of a truncated Munc13-1 peptide binding has been reported in which CaM binds in a typical extended configuration to two different motifs in the target.[60] **Figure 1.8** shows the structure of the complex and highlights two phenylalanine residues in the Munc13-1 peptide that anchor the target by inserting the side chains into the hydrophobic clefts of the two CaM domains.

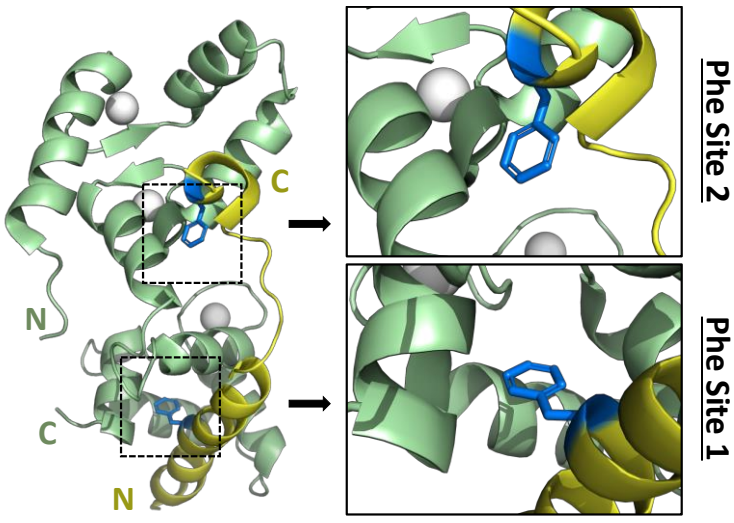

Figure 1.8. Crystal structure of Munc13-1 peptide bound to CaM.
CaM in **light green**, Ca^{2+} ions in **grey**, Munc13-1 peptide in **yellow**, and Munc-13 Phe residues in **blue**. (PDB 2KDU).

Human Pre-mRNA Processing Factor Homolog A (hPrp40A).

hPrp40 is a pre-mRNA processing factor with isoforms A and B. The domain architecture of the A isoform includes two N-terminal WW domains with two conserved Trp residues followed by six FF domains (**Figure 1.9**). Previous work has shown there is an interaction between the yeast homolog of Prp40A and HTT.[60] It has also been suggested that the WW domains interact with HTT and other mutated proteins involved in neurodegenerative disease progression.[61-64]

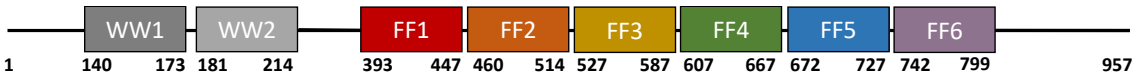

Figure 1.9. Domain architecture of hPrp40A.

Recently, several cryoEM stuctures of the yeast spliceosome have been determined.[65-67] The protein mixture used to prepare the sample (in all cases) included Prp40, but the authors were unable to assign Prp40 to the density. However, Prp40 can crosslink with Luc7 and Snu71 (other pre-mRNA processing factors) and the authors inferred that the partial density where Luc7 and Snu71 colocalize in the modeled structure also included Prp40A (**Figure 1.10**).

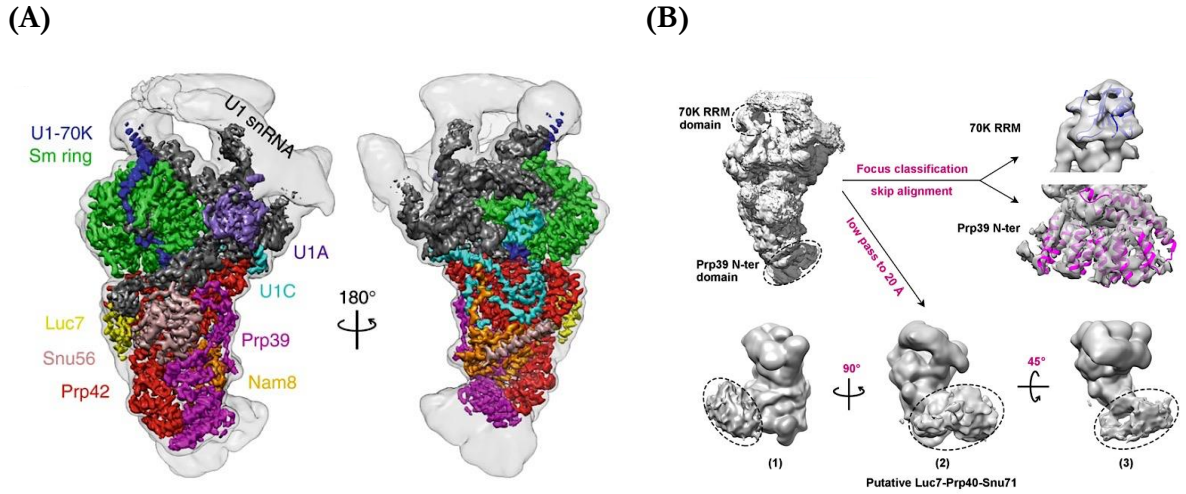

Figure 1.10. CryoEM structure of the yeast U1 snRNP.
(**A**) Various splicing factors are shown assembled bound to mRNA. (B) Unassigned cryoEM density showing where Luc7, Snu71, and Prp40A assemble within the spliceosome.[66] Used with permission.*

*Images from select publications used with permission in accordance with the Creative Commons license. Unless noted, no changes were made to the originally published figures. https://creativecommons.org/licenses/by/4.0/

To date, there have been no structures determined directly for the full-length hPrp40A, but the structures of individual FF domains have been reported. In particular, the structure of the FF_1 domain of yeast Prp40A has been determined by solution NMR (**Figure 1.12**).[68] The structure consists of a small three-helix bundle with an extended linker between the 2nd and 3rd helix.

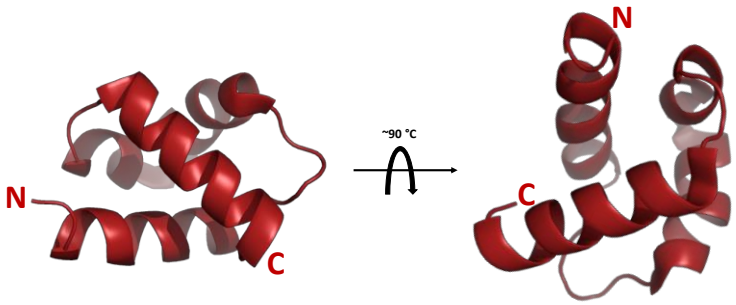

Figure 1.11. NMR solution structure of the FF_1 domain from yeast Prp40A.
(PDB 2B7E).

Calcium Sensors and hPrp40A Interactions

Previous work from the Chazin Laboratory demonstrated that two huntingtin-interacting proteins, Cen2 and Prp40A, interact with each other.[57] The strength of the interaction of hCen2 and the third domain of hPrp40A (hPrp40A-FF_3) measured by isothermal titration calorimetry, yielding a binding affinity (K_d) of 278 nM. Since hCen2 and CaM are functional and structural homologs as well as putative HTT-interacting proteins, it seemed plausible to hypothesize that CaM interacts with hPrp40A-FF_3 in a similar fashion as hCen2. Therefore, the major focus of this thesis was to test the hypothesis that these two proteins interact.

CHAPTER 2: INTERACTION OF THE THIRD DOMAIN OF HPRP40A AND CALMODULIN†

2.1 Introduction

Pre-mRNA processing protein 40 (Prp40) is an essential RNA splicing factor with two homologs in humans, A and B (hPrp40A and hPrp40B).[69,70] Prp40 proteins are composed of two WW domains (WW_1-WW_2) followed by six FF domains (FF_1-FF_6), with a flexible linker between the WW_2 and FF_1 domains.[71] The tandem WW domains have been implicated in severe neurological diseases, including Huntington's disease (HD)[62-64] and Rett syndrome[72] on the basis of interactions with both huntingtin (HTT) and a disease-associated mutant of methyl-CpG-binding protein 2, respectively. Furthermore, Prp40 has been reported to interact with other proteins via its FF domains, including RNA splicing factors. For example, the N-terminal crooked neck-like tetratricopeptide repeat (crn-TRP) of the splicing factor Clf1 interacts with yeast Prp40 (yPrp40) during the early steps of the spliceosome assembly.[68,73,74] Luc7 and Snu71 interact with yPrp40 FF domains simultaneously as components of the U1 small nuclear ribonucleoprotein (snRNP).[66,71,75] In addition, our group showed that the C-terminal domain of the EF-hand Ca^{2+} sensor human centrin 2 (hCen2) interacts in a Ca^{2+}-dependent manner with a 33 residue hPrp40A FF_3 peptide containing the centrin-binding consensus motif $W_1xxL_4xxxL_8$,[57] suggesting that Prp40A function may be modulated by intracellular Ca^{2+} signaling.

Calmodulin (CaM), the ubiquitous and archetypal EF-hand CaBP involved in regulating nearly 100 different pathways, is closely related to hCen2 and has also been found to interact with Prp40; protein microarray experiments on yeast CaM showed that it interacts with yPrp40.[76] Interestingly, human CaM has been shown to interact with HTT with a higher affinity for a disease-associated mutant than the wild-type protein,[56,77-79] and centrin has been found to colocalize with HTT at the centrosome and in the photoreceptor cilium.[80,81] Moreover, CaM is known to play a role in cross-linking mutant HTT to transglutaminase 2, an enzyme important for mediating Ca^{2+}-dependent protein modifications, and these proteins co-localize in the brain of HD patients.[79] Together, these observations suggest that hPrp40A and HTT are part of a coordinated network regulated by Ca^{2+} signaling via Cen2 and CaM.

† The majority of this chapter was recently submitted as a manuscript to *Biophysical Journal* as Díaz Casas A., Ferrer BJ, Cordoba JJ, Balakrishnan S, Wurm JE, Pastrana-Ríos B, and Chazin WJ., Binding by calmodulin requires unfolding of the Prp40A FF_3 domain, 2022.

The need for two calcium sensors acting on the same target simultaneously is puzzling as many Ca^{2+}-dependent signaling pathways typically only rely on one sensor.[82] This suggests that mHTT either uses one of these sensors in different functional contexts or uses both sensors in an unusual, coordinated fashion. In one model, interaction with one of the two sensors is significantly weaker than the other and competition could lead to preferential binding of mHTT to a specific sensor. More broadly, the observation that a pre-mRNA processing factor and calcium sensors are associated with mHTT suggests that mHTT might affect RNA splicing in a Ca^{2+}-dependent manner.

Given the observations that hCen2 and hPrp40A-FF_3 interact with each other and that hCen2 and CaM have high homology, we hypothesized that CaM and hPrp40A-FF_3 may also interact in a similar fashion. To test this proposal, a series of biophysical and structural analyses were performed to investigate the interaction of CaM with the hPrp40A FF_3 domain. We find Ca^{2+}-dependent binding by CaM in an extended configuration with both domains engaged. The characteristic hydrophobic anchors for binding into the two CaM domains are assigned to two Trp residues that are packed into the hydrophobic core of the folded FF_3 domain. The results are discussed in the context of the coupled unfolding and binding of the FF_3 domain and suggest a potential role for Ca^{2+} modulation of the HTT-Prp40-Cen2-CaM network in RNA splicing.

2.2 Experimental Methods

Expression and Purification of CaM.

The purification of CaM has been described elsewhere.[83,84] A pET-15b expression vector containing *CALM3* (encoding human calmodulin), inserted between 5'-NcoI and 3'-BamHI restriction sites, was obtained from the Chazin Laboratory. (See **Appendix** for plasmid sequence and biochemical properties of this construct.)

pET-15b-CaM was transformed by heat shock into *E. coli* BL21(DE3) cells. Successful transformants were inoculated into 2 x 1 L lysogeny broth (LB) supplemented with 100 μg/mL ampicillin. Cells were grown in an incubated shaker at 230 rpm at 37 °C. The optical density at 600 nm (OD_{600}) was measured until the culture reached OD_{600} of ~0.6-0.8. To the culture was added 1 mM isopropylthio-β-galactoside (IPTG). The culture was grown for an additional 3-4 hours at 37 °C. Cells were centrifuged at 6,500 RPM for 20 minutes at 4 °C. Supernatant decanted and the remaining cell pellet was either stored at -80 °C for long-term storage or immediately purified as described below.

Cell pellets were resuspended in ice cold lysis buffer [50 mM tris(hydroxymethyl)-aminomethane hydrochloric acid (Tris-HCl) (pH 7.4), 500 mM KCl, 1 mM ethylenediamine-tetraacetic acid (EDTA), 1 mM phenylmethylsulfonyl fluoride (PMSF), and 1 protease inhibitor tablet (Roche)] and sonicated for 3 cycles at 70% duty cycle at 4 °C. Crude lysate was centrifuged at 20,000 RPM for 20 minutes at 4 °C. The pellet was discarded and supernatant was filtered through a 0.45 μm PVDF syringe filter and kept on ice.

Supernatant was loaded onto a prepacked phenyl sepharose column (HiTrap Capto Phenyl hi sub, Cytiva, 1 column volume, CV = 5 mL) equilibrated in Buffer A1 [50 mM Tris (pH 7.4), 150 mM KCl, 1 mM EDTA] and CaM was collected in the flow-through. The column was washed with 2 CVs of Buffer A1, 2 CVs MQ H_2O, 2 CVs 20% EtOH, and 4 CVs MQ H2O. The column was equilibrated in Buffer B [50 mM Tris (pH 7.4), 500 mM KCl, 10 mM $CaCl_2$]. The flow-through containing CaM was spiked up to 10 mM $CaCl_2$ and loaded back onto the column. The column was washed 4 CV Buffer B. CaM was eluted with a 0-100% gradient of Buffer A2 [50 mM Tris (pH 7.4), 150 mM KCl, 1 mM EDTA]. Eluate was collected in 2 mL fractions. Fractions were run on SDS-PAGE and fractions containing CaM were pooled, concentrated to ~500 μM, flash-frozen in liquid N_2, and stored at -80 °C. The identity of all proteins was validated by mass spectrometry and SDS-PAGE (**Figure S1, S2; Table S1, S2**.).

Expression and Purification of ^{15}N-CaM.

^{15}N-CaM was prepared by Dr. Adalberto Díaz-Casas. pET-15b-CaM was transformed by heat shock into *E. coli* BL21(DE3) cells. Successful transformants were inoculated into 2 x 1 L ^{15}N-enriched M9 minimal media supplemented with 50 μg/mL ampicillin. Cells were grown in an incubated shaker at 230 rpm at 37 °C. The OD_{600} was measured until the culture reached OD_{600} of ~0.6-0.8. To the culture was added 0.5 mM IPTG. The culture was grown overnight at 25 °C. Cells were centrifuged at 6,500 RPM for 20 minutes at 4 °C. Supernatant decanted and the remaining cell pellet was either stored at -80 °C for long-term storage or immediately purified as described for **WT CaM**. The identity of all proteins was validated by mass spectrometry and SDS-PAGE (**Figure S1, S2; Table S1, S2**.).

Expression and Purification of WT hPrp40A-FF$_3$.

Cloning was performed by Brian Ferrer. A gene fragment of hPrp40A-FF$_3$ (encoding an N-terminal Met and hPrp40A residues Ser516 to Glu593) was inserted between 5'-NdeI and 3'-NheI restriction

sites using seamless In-Fusion Cloning (TakaraBio). (See **Appendix** for plasmid sequence and biochemical properties of this construct.)

pBG100-FF_3 was transformed by heat shock into *E. coli* BL21(DE3) cells. Successful transformants were inoculated into 2 x 1 L LB supplemented with 50 μg/mL kanamycin. Cells were grown in an incubated shaker at 230 rpm at 37 °C. The OD_{600} was measured until the culture reached OD_{600} ~0.6-0.8. To the culture was added 1 mM IPTG. The culture was grown for an additional 3-4 hours at 37 °C. Cells were centrifuged at 6,500 RPM for 20 minutes at 4 °C. The supernatant was decanted and the remaining cell pellet was either stored at -80 °C for long-term storage or immediately purified as described below.

Cell pellets were resuspended in ice cold lysis buffer [20 mM N-(2-Hydroxyethyl)piperazine-N'-(2-ethanesulfonic acid) (HEPES) (pH 7.4), 100 mM NaCl, 1 mM ethylenediaminetetraacetic acid (EDTA), 10 mM dithiothreitol (DTT), and 1 protease inhibitor tablet (Roche)] and sonicated for 3 cycles at 70% duty cycle at 4 °C. Crude lysate was centrifuged at 20,000 RPM for 20 minutes at 4 °C. The pellet was discarded and supernatant was saved. Ammonium sulfate powder was slowly added to the stirring supernatant to a final concentration of 0.450 g/mL at 4 °C. The sample was centrifuged at 20,000 RPM for 20 minutes at 4 °C. The pellet was discarded. The supernatant was spiked up to 750 mM NaCl, filtered through a 0.45 μm PVDF syringe filter, and kept on ice.

The supernatant was loaded onto a prepacked phenyl sepharose column (HiTrap Capto Phenyl hi sub, Cytiva, 1 column volume (CV) = 5 mL) equilibrated in 75% Buffer B [20 mM HEPES (pH 7.4), 1 M NaCl, 10 mM DTT] and 25% Buffer A [20 mM HEPES (pH 7.4), 100 mM NaCl, 10 mM DTT]. The column was washed with 5 CVs of 75% Buffer B and 25% Buffer A. Protein was eluted with a 75-0% Buffer B gradient. Eluate was collected in 2 mL fractions. Fractions were run on SDS-PAGE and fractions containing hPrp40A-FF_3 were pooled, concentrated to ~5 mL, and loaded in 2 x ~2.5 mL batches onto a prepacked gel filtration column (HiLoad 16/600 Superdex 75 pg, Cytiva, 1 CV = 120 mL) equilibrated in SEC Buffer [50 mM HEPES (pH 7.4), 150 mM NaCl, 4 mM $CaCl_2$, 10 mM DTT]. The protein was eluted with 1 CV of SEC Buffer. Eluate was collected in 1 mL fractions. Fractions were run on SDS-PAGE and fractions containing WT hPrp40A-FF_3 were pooled, concentrated to ~500 μM, flash-frozen in liquid N_2, and stored at -80 °C. The identity of all proteins was validated by mass spectrometry and SDS-PAGE (**Figure S1, S2; Table S1, S2**.).

Mutagenesis of hPrp40A-FF_3.

Cloning was performed by Brian Ferrer. pBG100-FF_3 was used as a template to generate tryptophan to alanine substitutions at our proposed binding sites. Non-overlapping single-stranded DNA (ssDNA) primers were designed to mutate these sites using site-directed mutagenesis. Primers were obtained from IDT. Primers used included the following:

W531A Forward Primer: 5'-CAAACGTAAC**GCG**GAAGCATTAAAAAATATTCTG-3'

W531A Reverse Primer: 5'-CGGAGTTGCTTTGCTTGT-3'

W550A Forward Primer: 5'-TAGCACCACC**GCA**TCAGAAGCACAGC-3'

W550A Reverse Primer: 5'-TAGGTAACATTGGCCATATTATC-3'

Mutagenesis was carried out using a commercial kit (Q5 SDM Kit, New England Biolabs). To obtain the double substitution variant, pBG100-FF_3 W531A was used as a template and the W550A substitution was generated using the corresponding primers. Successful constructs were validated by Sanger sequencing (Genewiz). (See **Appendix** for plasmid sequence and biochemical properties of these constructs.)

Expression and Purification of hPrp40A-FF_3 Variants.

pBG100-FF_3 W531A, W550A, and W531A/W550A were expressed and purified in the same manner as WT hPrp40A-FF_3 as described above.

Isothermal Titration Calorimetry.

Experiments were performed using a TA Instruments Affinity ITC. The initial round of ITC was performed by Dr. Adalberto Díaz Casas. Subsequent rounds of ITC with FF_3 variants were performed by Brian Ferrer. Samples for isothermal titration calorimetry were prepared by incubating CaM and WT hPrp40A-FF_3 with 100 mM β-mercaptoethanol (βME) added to the stock protein solution at 4 °C for 1 hour. Proteins were buffer exchanged into 50 mM HEPES (pH 7.4), 150 mM NaCl, 4 mM $CaCl_2$, and 5 mM TCEP-HCl. For analysis in the absence of Ca^{2+}, the samples were buffer exchanged into the same buffer except $CaCl_2$ was replaced with 2 mM EDTA. All samples were passed through a 0.45 μM PVDF filter and degassed for at least 10 min before use. Experiments were run with 30 μM CaM in the sample cell and 180 μM FF_3 (wildtype or variant) in the injection syringe. Titrations were performed by automatic injection with an initial injection of 0.5 μL FF_3 followed by 30 x 3 μL injections at 250-400 second intervals. For oversaturation experiments, 75 x 3 μL injections were made. Injections were made with 250-400 second intervals.

The data were fit with an independent binding model using the instrument's NanoAnalyze software. Experiments were performed in triplicate and the results are reported as the mean and standard deviation (**Table 2.1**).

Differential Scanning Calorimetry (DSC).

DSC experiments were performed by Dr. Adalberto Díaz Casas. A sample of hPrp40A FF_3 (60 μM) in 50 mM HEPES (pH 7.4), 150 mM NaCl, 4 mM $CaCl_2$, 4 $MgCl_2$, and 1 mM TCEP-HCl was analyzed using a MicroCal VP-DSC microcalorimeter. The protein sample and the reference were degassed for at least 10 min before use. Thermograms were collected at 25 psi and a scan rate of 60 °C/h over a temperature range of 10-95 °C, with a 16 s filtering period. The data analysis was performed using the instrument's Origin software. The thermogram was reference subtracted and a progressive baseline was obtained for accurate determination of the thermal denaturation mid-point (T_m).

Homology Modeling of hPrp40A-FF_3

Homology modeling of the hPrp40A-FF_3 domain alone was performed by Dr. Adalberto Díaz Casas. The model was generated *ab initio* using the QUARK server.[85,86] In brief, the sequence of FF_3 (Ser516 to Glu593) was input into an online server (https://zhanggroup.org/QUARK/). A knowledge-based Monte Carlo simulation was run to generate a predicted model independent of a PDB template. The top 5 final structures based on the highest predicted confidence score calculated by the algorithm were generated from this simulation. The best scoring model was selected. The model was overlaid with the *ab initio* electron map generated from experimental SAXS data for hPrp40A-FF_3 alone.

Threading of hPrp40A-FF_3 bound to CaM

Threading was performed by Johnny Cordoba and Dr. Adalberto Díaz Casas. The model of hPrp40A FF_3 bound to CaM was generated using Modeller v10.[87] In brief, the crystal structure of Munc13-1 peptide bound to CaM (PDB 2KDU) was used as a template. The sequence of FF_3 (Ser516 to Glu593) was aligned to the Munc13-1 peptide sequence. Additional emphasis was placed on key hydrophobic residues within the program. The threading simulation generated 100 models of potential fits of the FF_3 domain bound to an extended configuration of CaM. The best scoring model was selected. The

model was overlaid with the *ab initio* electron map generated from experimental SAXS data for CaM in solution with hPrp40A-FF_3.

Size-Exclusion Chromatography and Small Angle X-ray Scattering (SEC-SAXS)

The experiments were performed at the SIBYLS beamline using previously described protocols.[88] Sample preparation, experiments, and data processing were performed by Johnny Cordoba and Dr. Adalberto Díaz Casas. In brief, Ca^{2+}-loaded CaM and hPrp40A FF_3 were prepared in a 1:1.5 CaM:FF_3 ratio. The sample (10 mg/mL) was loaded onto a Shodex 802.5 SEC column connected to a multi-angle light scattering system and the CaM-FF_3 complex was eluted in 50 mM HEPES (pH 7.4), 100 mM NaCl, 2 mM $CaCl_2$, and 1 mM TCEP-HCl. Data were also collected for isolated CaM (8 mg/mL) and hPrp40A FF_3 (8 mg/mL) in the same buffer conditions. The sample volume run through the SEC was 50 µl and the eluent was split between SEC-MALS and SAXS channels. 3.0-s X-ray exposures were collected continuously during a ~35 min elution. The SAXS frames recorded prior to the protein elution peak were used as buffer blanks to subtract from all other frames. The subtracted frames were examined by radius of gyration (R_g) and scattering intensity at $q = 0$ Å^{-1} ($I(0)$), derived using the Guinier approximation $I(q) = I(0)e^{-qR_g^2/3}$ with the limits $q*R_g < 1.5$. $I(0)$ and R_g values were compared for each collected SAXS curve across the entire elution peak. The elution peak was mapped by plotting the scattering intensity at $q = 0$ Å^{-1} ($I(0)$), relative to the recorded frame. Uniform R_g values across an elution peak represent a homogenous assembly. The merged experimental SAXS data were additionally investigated for aggregation by inspecting Guinier plots. The SAXS data were processed using Scatter 4.0. Back-calculated scattering data from previously determined structural coordinates were generated using the FoXS server.[88,89] The *ab initio* electron maps obtained from the SAXS data of CaM, hPrp40A FF_3, and the CaM-FF_3 complex were determined using DENSS.[90]

^{15}N-^{1}H HSQC NMR.

Samples were prepared by Brian Ferrer. Stock concentrations of ^{15}N-labeled calmodulin (^{15}N-CaM) and hPrp40A-FF_3 (WT or variants) were prepared by buffer exchanging into ITC Buffer [50 mM HEPES (pH 7.4), 150 mM NaCl, 4 mM $CaCl_2$, 5 mM TCEP] at 4 °C. The samples were concentrated to be ≥ 80 µM for ^{15}N-CaM and ≥ 80 µM hPrp40A-FF_3 (WT or variants). Samples were prepared in a 1:2 ratio (^{15}N-CaM/hPrp40A-FF_3) with 50 µM ^{15}N-CaM, 100 µM hPrp40A-FF_3, and 10% (v/v) D_2O in ITC Buffer [50 mM HEPES (pH 7.4), 150 mM NaCl, 4 mM $CaCl_2$, 5 mM

TCEP]. For oversaturation experiments, samples were prepared in a 1:20 ratio with 50 μM ^{15}N-CaM, 1000 μM hPrp40A-FF_3, and 10% (v/v) D_2O in ITC Buffer.

NMR experiments and data processing were performed by Dr. Swati Balakrishnan. 2D ^{15}N-^{1}H HSQC experiments were recorded in 5-mm tubes at 37 °C using a Bruker AVANCE 900 MHz spectrometer equipped with a TCI cryoprobe. The pulse sequence hsqcetfpf3gpsi2 from the standard suite of Bruker pulse programs was used. In brief, this pulse program includes data acquisition for ^{15}N and ^{1}H signals allowing peak acquisition in two dimensions as well as solvent suppression to decrease background noise.[91] All spectra were acquired with 32 scans and 2048 in the direct dimension and 128 points in the indirect dimension. Data processing and analysis were carried out using NMRPipe and CcpNmr version 2.5.1. Peak assignments for specific regions were made and labeled on the spectra.

Fluorescence Spectroscopy

Intrinsic fluorescence experiments were performed by Jennifer Wurm. A sample of hPrp40A FF_3 (5 μM) in 50 mM HEPES (pH 7.4), 150 mM NaCl, 2 mM $CaCl_2$, and 1 mM TCEP-HCl was titrated with a solution of 100 μM CaM in the same buffer. The fluorescence emission of the two Trp residues in the FF_3 was measured at 37 °C using a Horiba Jobin Yvon (Edison, NJ) Fluoromax-3 fluorometer. The fluorescence excitation wavelength was 295 nm, and the fluorescence emission spectra were recorded between 300 and 500 nm, using quartz cuvettes.

2.3 Results

CaM binds to hPrp40A FF_3 domain in a Ca^{2+}- and temperature-dependent manner

CaM and Cen2 are EF-hand Ca^{2+} sensors with similar structures that have both been shown to interact with Prp40, suggesting that the activity of this splicing factor is modulated by a complex Ca^{2+} signaling network. The binding of hCen2 to hPrp40A has been mapped to the FF_3 domain, and since CaM and Cen2 are similar, we used isothermal titration calorimetry experiments to determine if CaM could also bind to this domain. The ITC experiments were performed by titrating CaM into the FF_3 domain, for both Ca^{2+}-free and Ca^{2+}-loaded CaM (**Figure 2.1**). The data show interaction is Ca^{2+}-dependent with a K_d value of 25 ± 3 μM and a stoichiometry of 1:1. Further support was provided by the observation that a mixture of CaM and the FF_3 co-purify by SEC at room temperature but not at 4 °C, consistent with temperature-dependent binding (**Figure 2.2**).

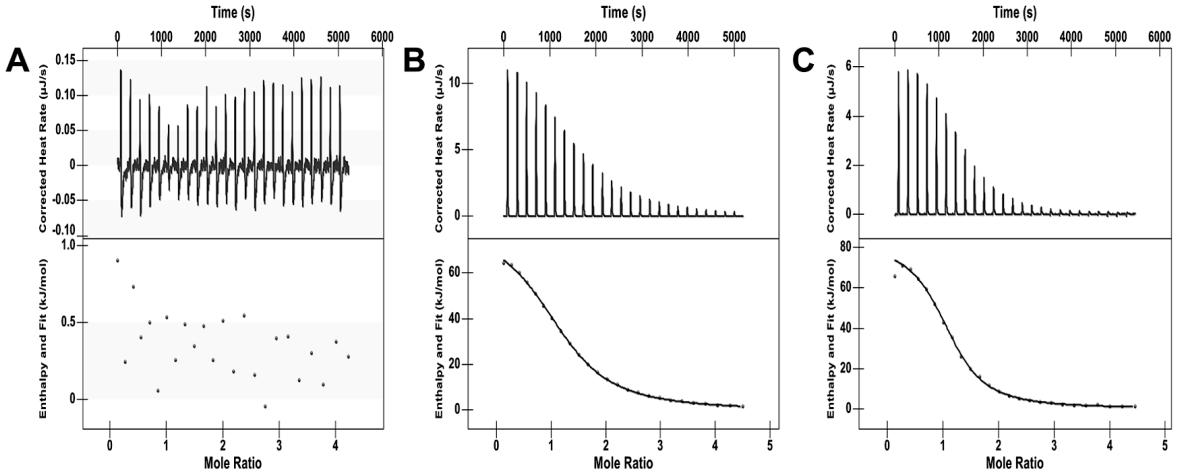

Figure 2.1. Analysis of the interaction of hPrp40A FF_3 with CaM.
Representative ITC thermograms and plots of heat changes in titrations of the hPrp40A FF_3 with **(A)** Ca^{2+}-free CaM, **(B)** Ca^{2+}-loaded CaM at 25 °C and **(C)** Ca^{2+}-loaded CaM at 37 °C.

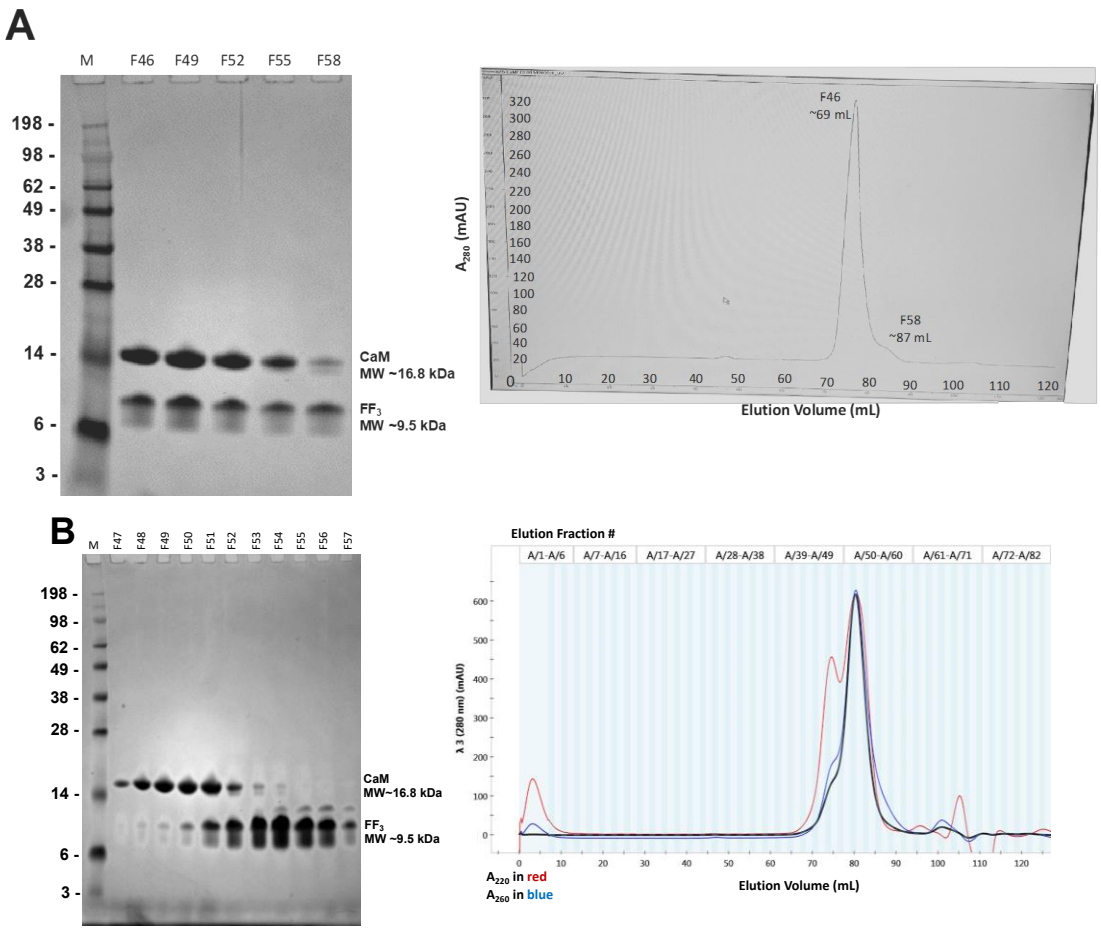

Figure 2.2. Co-elution analysis between CaM and hPrp40A FF_3 by SEC.
(A) A sample of CaM with a 1.5 molar excess of hPrp40A-FF_3 was mixed at 25 °C and loaded onto a 120 mL HiLoad 16/600 Superdex 75 column (Cytiva). The sample was eluted at 25 °C on a GE ÄKTA FPLC system. 1.5 mL fractions were manually collected.
(B) A sample of CaM with a 1.5 molar excess of hPrp40A-FF_3 was mixed at 4 °C and loaded onto a 120 mL HiLoad 16/600 Superdex 75 column (Cytiva). The sample was eluted at 4 °C on a Bio-Rad NGC Medium-Pressure LC system. 1.5 mL fractions were automatically collected.
The bands that correspond to CaM (16.8 kDa) and hPrp40A FF_3 (9.5 kDa) are identified. In each gel, M corresponds to the SeeBlue Plus 2 Prestained molecular weight standard from Invitrogen (Carlsbad, CA).

Considering the relatively modest affinity at room temperature and knowing that CaM generally binds linear peptide motifs of 15-25 residues[59] yet the FF_3 domain has a globular fold, we surmised that the binding by CaM of the FF_3 may require unfolding of the domain. To assess the potential for the domain to unfold, the T_m of the FF_3 was measured by DSC. A value of 50 °C was measured (**Figure 2.3**), consistent with previously reported studies by circular dichroism and 2D IR correlation spectroscopy.[92]

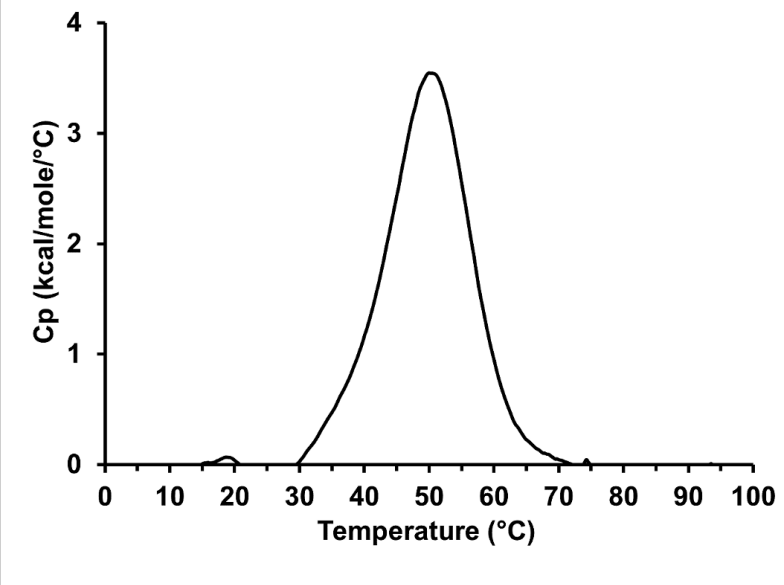

Figure 2.3. DSC thermogram plot of hPrp40A FF_3.
The T_m value measured is 50 °C.

The T_m of hPrp40A FF_3 is also similar to the T_m of other FF domains reported in the literature, which include the p190 RhoGAP FF_1 (T_m = 52 °C), and the hTCERG1 FF_2 (T_m = 55 °C)[93] and FF_4 (T_m = 49 °C) domains.[94] The T_m value of 50 °C for the FF3 implies a relatively low equilibrium concentration of unfolded protein at 25 °C, consistent with the relatively weak binding measured by ITC. To obtain further insight, the ITC measurement was repeated at 37 °C (**Figure 2.1 C; Table 2.1**) where the equilibrium concentration of unfolded protein will be higher. As anticipated, a five-fold higher affinity (K_d = 4.9 ± 0.3 μM) was observed, consistent with binding to CaM occurring via the equilibrium population of unfolded FF_3. Taken together, these results imply the binding of hPrp40A FF_3 to CaM is Ca^{2+}-dependent and coupled to unfolding of the FF_3 domain.

Table 2.1. Summary of the thermodynamic parameters for the interaction of CaM and hPrp40A-FF_3.

Protein	Temperature (° C)	K_d (error) (μM)	ΔH (error) (kJ/mol)	-TΔS (error) (kJ/mol)	*n* (error)
Ca^{2+}-CaM + FF_3	25	25 (3)	82 (2)	-108 (2)	1.25 (0.05)
Ca^{2+}-CaM + FF_3	37	4.9 (0.3)	85 (4)	-116 (4)	1.13 (0.04)
Apo-CaM + FF_3	25	N.B.	-	-	-

N.B., no binding was observed.

The error reported is the standard deviation of triplicate measurements.

Structure of hPrp40A FF_3

Structures have been determined previously for the FF_1 and FF_6 of hPrp40A and yPrp40,[77,95-97] but not for hPrp40A-FF_3. Alignment of FF domain sequences from three different classes of proteins show there are eleven conserved residues that constitute the hydrophobic core of FF domains (**Figure. 2.4**).

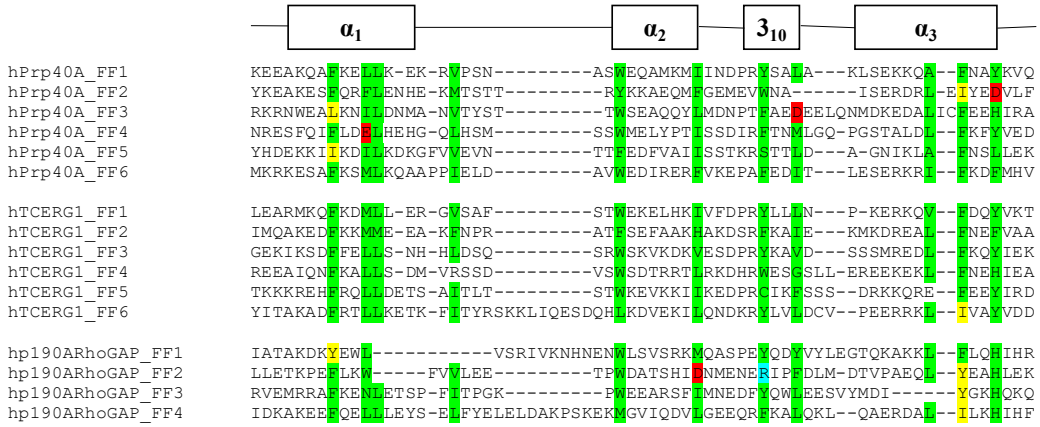

Figure 2.4. Sequence alignment of FF domains.
FF domains from hPrp40A (Uniprot ID: O75400), transcription factor hTCERG1 (Uniprot ID: O14776), and Rho GTPase human p190A RhoGAP (hp190ARhoGAP) (Uniprot ID: Q9NRY4). Residues known to be part of the hydrophobic core of FF domains are highlighted in **green**. Substitutions of these conserved residues with hydrophobic, acidic, and basic residues are highlighted in **yellow** and **red**, and **blue**, respectively. The alignments were generated using the online tool Clustal Omega.[98]

Of note, hPrp40A FF_3 has two non-consensus residues: the conserved Phe in α_1 is substituted by a Leu (Leu534) and one of the conserved hydrophobic residues in the 3_{10}-helix is substituted by an Asp (Asp566). Despite these differences, the high degree of sequence similarity overall meant that an accurate structural model could be generated using a homology modeling approach.

A structural model of hPrp40A FF_3 was generated using the online QUARK server.[85,86] (**Figure 2.5 A**). Overlay of this model on the human transcription elongation regulator 1 (hTCERG1) FF_1 shows the two key 'F' residues in the FF_3 domain (Leu534 and Phe581) are well aligned with the two conserved phenylalanine residues (**Figure 2.5 B-D**). Previous structural comparisons of FF domains revealed that the loop-3_{10}-loop region is the most variable.[95] This is also the case for hPrp40A FF_3. We believe the reorientation of this loop is driven largely by the non-consensus residue Asp566, which faces the solvent rather than inward towards the core as observed for the consensus Leu in hTCERG1 FF_1 (**Figure 2.5 E**).

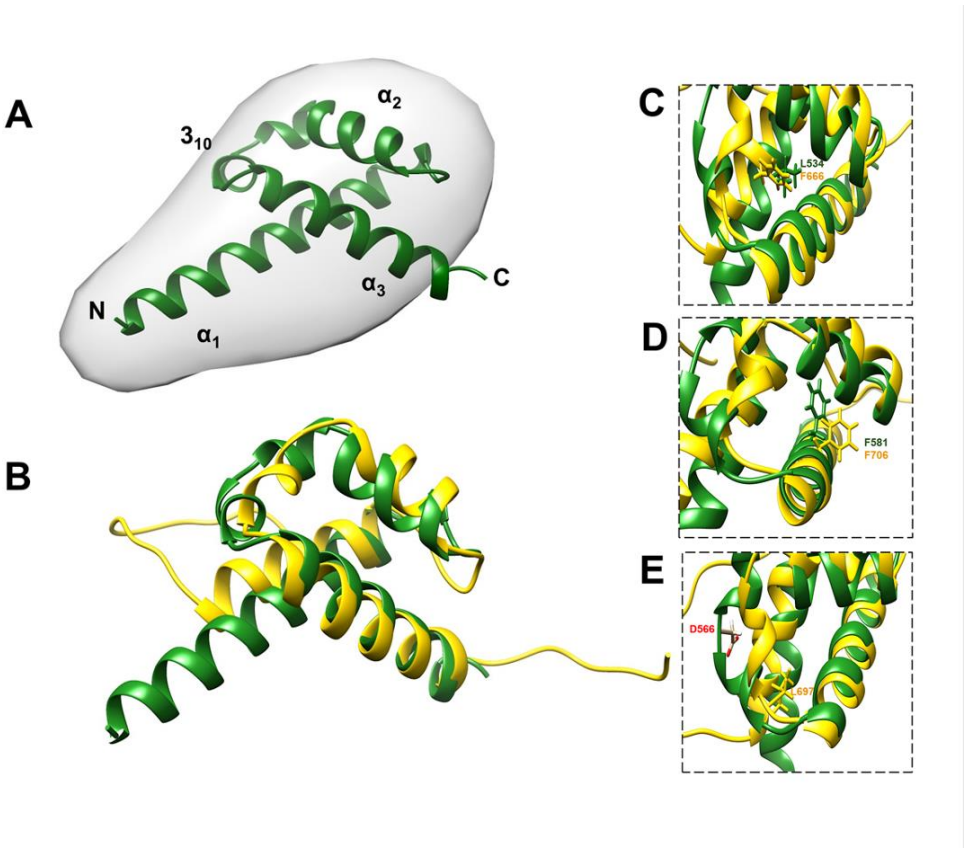

Figure 2.5. Structural analysis of hPrp40A FF_3.
(A) Overlay of the *ab initio* electron map of hPrp40A FF_3 generated from SAXS and the 3D homology model (**green**). **(B)** Superposition of the homology model of the hPrp40A FF_3 and the NMR structure of hTCERG1 FF_1 (PDB 2DOD, **yellow**). **(C, D)** Close-up view of the two conserved "FF domain" phenylalanine residues located in the middle of α1 and α3, respectively. **(E)** Close-up view of the non-conserved Asp566 residue in the loop-3_{10}-loop region of hPrp40A FF_3, compared to the conserved Leu697 in hTCERG1. Note that Asp566 faces the solvent whereas Leu697 extends into the hydrophobic core. The figures were generated using UCSF Chimera.[99]

To validate the structural model, SAXS data were collected for isolated hPrp40A FF_3 (**Figure 2.6 A-C**). The linearity observed in the Guinier region of the $\log_{10}$ intensity plot indicates the sample was free of aggregation. The Kratky plot and the Porod exponent (P_x) of 3.6 are indicative of a stable globular particle, but with some degree of flexibility retained, likely at the N- and/or C-terminus. The experimental probability distribution function P(r), which reflects the distribution of inter-atomic distances, was consistent with a somewhat spherical globular domain.

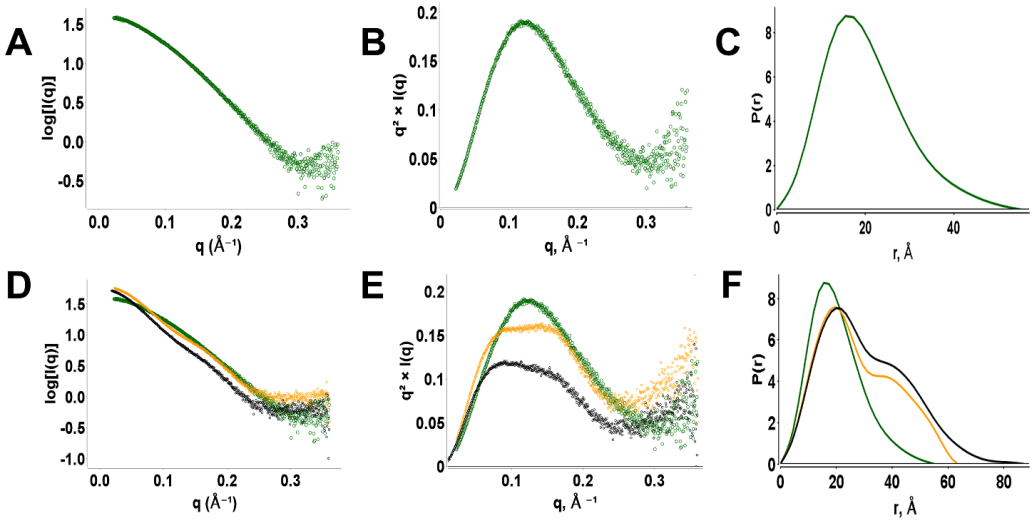

Figure 2.6. SAXS data for the hPrp40A FF_3, CaM, and the CaM-FF_3 complex.
(A) Log_{10} intensity plot, **(B)** Kratky plot, and the **(C)** P(r) plot of hPrp40A FF_3. **(D-F)** Overlay of the log_{10} intensity plot, Kratky plot, and P(r) plot of hPrp40A FF_3 (**green**), CaM (**orange**), and the CaM-FF_3 complex (**black**).

Figure 2.7 compares the experimental P(r) function to those back-calculated from the structures of three different FF domains. These show great similarity overall, with the closest match to hTCERG1 FF_1 (PDB 2DOD; $\chi^2 = 1.97$). The P(r) function back-calculated from the homology model is also shown; the high χ^2 value can be attributed to the random positioning of the disordered N- and C-termini. Thus, the SAXS analysis is consistent with the hPrp40A FF_3 possessing the common fold found in all FF domains and supports the validity of the homology model.

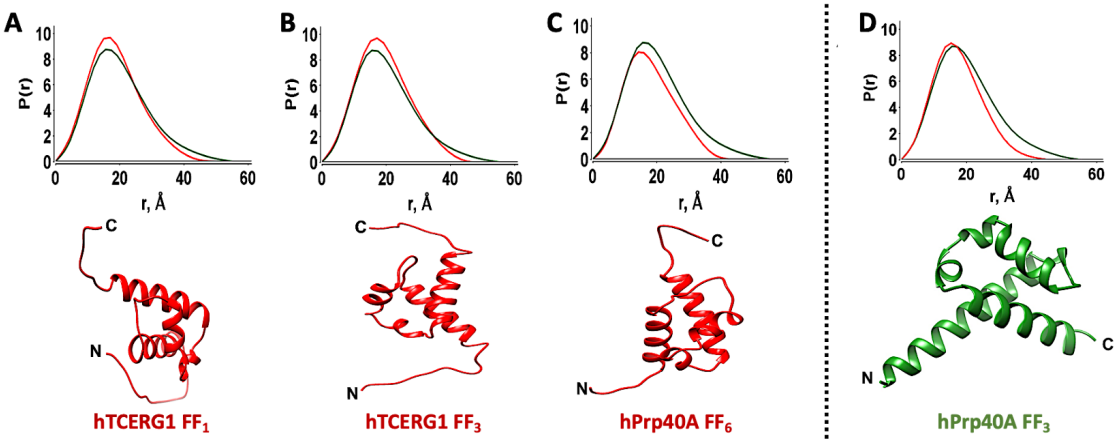

Figure 2.7. SAXS analysis of the structure of the hPrp40A FF_3 domain.
Overlays of the P(r) derived from the SAXS data for hPrp40A-FF_3 (**green**) with the back-calculated P(r) (**red**) extracted from coordinates of NMR structures of **(A)** hTCERG1 FF_1 (PDB: 2DOD model 5, $\chi^2 = 1.97$), **(B)** hTCERG1 FF_3 (PDB: 2DOE model 15, $\chi^2 = 2.12$), **(C)** hPrp40A FF_6 (PDB. 2CQN model 5, $\chi^2 = 3.23$), and **(D)** the homology model of hPrp40A FF_3 ($\chi^2 = 13.15$). The structure of the hPrp40A FF_3 is colored **green** and other FF domains are colored **red**. Structures were generated using UCSF Chimera.[99]

CaM interacts with hPrp40A FF_3 in an extended binding mode

We set out to determine the X-ray crystal structure of the CaM-FF_3 complex. However, despite the ability to isolate the complex by SEC, we were unable to find conditions that produced diffraction quality crystals. We therefore considered alternate approaches and chose to generate a structural model using an NMR and SAXS-based approach as applied previously in our laboratory and elsewhere.[100-102] Although this approach does not directly provide complete high-resolution structures, it can accurately determine molecular shapes and topologies for globular domains and multi-domain complexes.

CaM is comprised of two EF-hand domains connected by a flexible linker, an organization that provides it with a wide range of possibilities to interact with target proteins.[103] These different modes span the range from utilizing just one of its two domains, to wrapping both domains around a short linear motif in the target, to using its two domains to contact distant points in the target or even two different target molecules.[104] Here we used NMR to determine if one or both domains of CaM are engaged in the complex with FF_3. Samples of ^{15}N-enriched CaM were prepared in the presence of Ca^{2+}, and 2D ^{15}N-^{1}H HSQC NMR spectra were acquired in the absence and presence of the FF_3 (**Figure 2.8**).

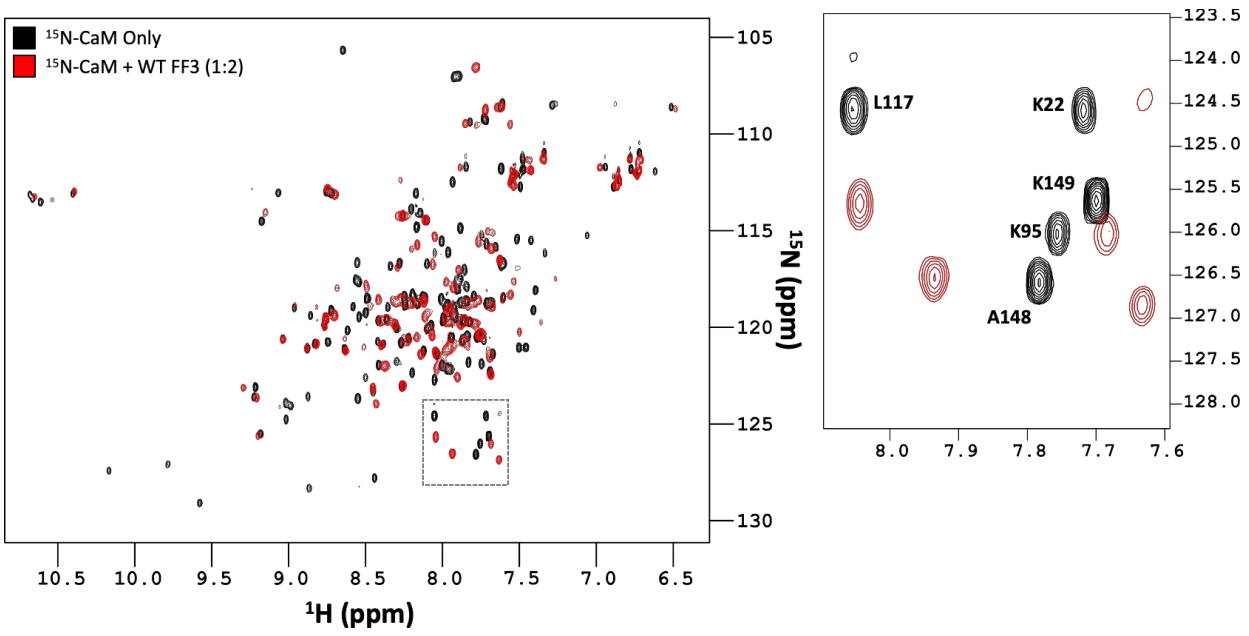

Figure 2.8. NMR analysis of the interaction between hPrp40A FF_3 and CaM.
900 MHz 2D ^{15}N-^{1}H HSQC NMR spectra at 37 °C of ^{15}N-enriched Ca^{2+}-loaded CaM alone (**black**) and with 2-fold excess of WT FF_3 (**red**). The sequence-specific assignment is provided for crosspeaks from residues Lys22, Lys95, and Leu117 in both the N- and C-terminal domains of CaM to highlight that both domains engage the FF_3 domain.

Binding of FF_3 resulted in a general increase in linewidths due to the increased mass of the complex and likely reduction in the inter-domain flexibility. Moreover, numerous chemical shift perturbations were observed in both the N- *and* C-terminal domains of CaM, validating that both are involved in the interaction with the FF_3 domain. In addition, a select number of crosspeaks in the spectrum were broadened beyond detection due to intermediate exchange between the free and bound state, which precluded structure determination by traditional NMR approaches.

SAXS was used to determine the topology and shape of the CaM-FF_3 complex, using data acquired for isolated FF_3 and CaM as well as for the complex (**Figure 2.6; Table 2.2**). Similar to the isolated FF_3, the Guinier region of the $\log_{10}$ intensity plots of isolated CaM and the CaM-FF_3 complex indicate that the samples were free of aggregation (**Figure 2.6 D**). The high quality of the data is also reflected in the excellent agreement between the R_g values obtained from the Guinier and P(r) analyses (**Table 2.2**). The R_g, Porod volume (V_p), and D_{max} values correlate well with globular structures being present and are consistent with the respective masses of hPrp40A FF_3 (9.3 kDa), CaM (16.7 kDa), and the CaM-FF_3 complex (26.0 kDa). The Kratky plots and P_x values indicate a substantial degree of ordered globular structure (**Figure 2.6 E; Table 2.2**).

Table 2.2. Summary of the parameters obtained from the SAXS analysis of the hPrp40A FF_3, CaM, and the CaM-FF_3 complex.
The values in bold were obtained from SAXS analysis (See Figure 2.6). $R_{g\ Guinier}$ and $R_{g\ P(r)}$ are the R_g values obtained from SAXS Guinier and P(r) plot analyses, respectively.

Protein	MW (kDa)	$R_{g\ Guinier}$ (Å)	$R_{g\ P(r)}$ (Å)	V_p (Å^3)	D_{max} (Å)	P_x
hPrp40A-FF_3	9.3	**16.09**	**15.65**	**17490**	**55**	**3.6**
CaM	16.7	**21.66**	**21.99**	**26889**	**67.5**	**3.8**
CaM + hPrp40A-FF_3	26.0	**23.96**	**24.32**	**34626**	**87**	**3.9**

CaM has been studied extensively by SAXS and the values obtained in our study are consistent with past reports.[105-107] The P(r) of free CaM contains a peak with a maximum at ~21 Å and a shoulder at ~40 Å indicating that on average it occupies an extended configuration, which is similar to the crystal structure of Ca^{2+}-loaded CaM (**Figure 2.6 F and Figure 2.9**). Interestingly, the P(r) of the CaM-FF_3 complex showed the same two features (**Figure 2.6 F**).

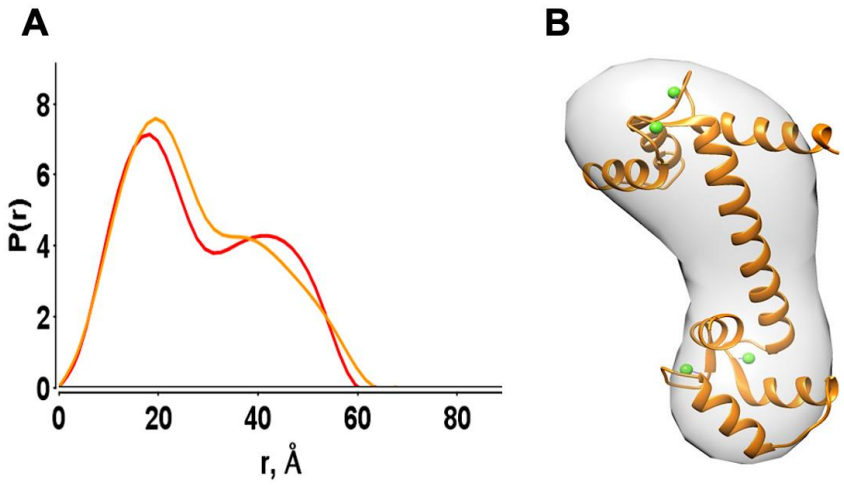

Figure 2.9. Analysis of the architecture of CaM by SAXS.
(A) Overlay of the experimental SAXS P(r) for Ca^{2+} saturated CaM (orange) with the back-calculated P(r) (red) extracted from the coordinates of the crystal structure of the Ca^{2+} bound CaM (PDB: 1CLL, $\chi^2 = 5.05$). (**B**) The structure of CaM is superimposed with the *ab initio* electron map generated from SAXS.

In order to further analyze the molecular architecture of the complex, we compared the experimental scattering of the CaM-FF_3 complex with the theoretical scattering back-calculated from structural coordinates of three CaM-target complexes with three different binding modes (**Figure 2.10**). High values for the χ^2 fitting parameter and poor recapitulation of the shape of the curve were obtained for structures in which CaM engages its target in a "wrap-around" mode (χ^2=147.88) or with only one CaM domain bound (χ^2 = 8.65) (**Figure 2.10 A, B**). The greatest similarity between the experimental and back-calculated P(r) was found for the structure of CaM bound to a Munc13-1 peptide (χ^2=6.10),[61] in which CaM engages in an extended mode with each domain contacting separate motifs in the target (**Figure 2.10 C**).

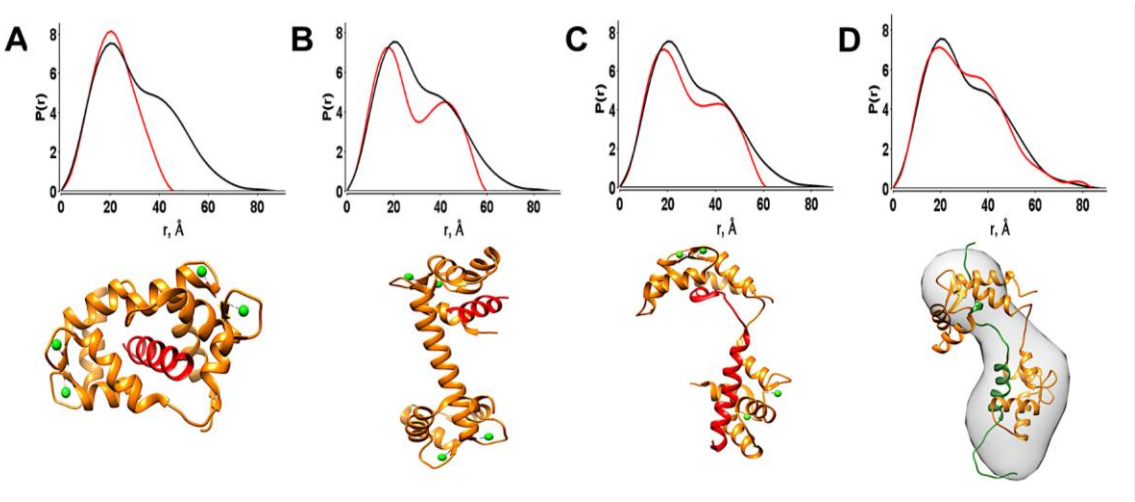

Figure 2.10. Comparative SAXS analysis of the CaM-hPrp40A FF_3 complex and CaM-target complexes in different binding modes.
(**A-C**) Overlays of the P(r) derived from the SAXS data for the CaM-FF_3 complex (**black**) on the back-calculated P(r) (**red**) extracted from coordinates of Ca^{2+} bound CaM in complex with a CaM-binding peptides from **(A)** smooth muscle light chain kinase (PDB: 2O5G, $\chi^2 = 147.88$), **(B)** human cardiac sodium channel ($Na_V1.5$) (PDB 4DJC, $\chi^2 = 8.65$), **(C)** Munc13-1 (PDB 2KDU model 5, $\chi^2 = 6.10$) peptides.
(D) Overlay of the P(r) derived from SAXS data for the CaM-FF3 complex (**black**) and back-caclculated P(r) (**red**) extracted from our threaded model of the CaM-FF_3 complex ($\chi^2 = 7.04$).
CaM is colored **orang**e, the target peptides **red**, and the FF_3 domain **green**. Ca^{2+} ions are represented with **green spheres**. The structures were rendered with UCSF Chimera.[99]

Based on this analysis we generated a threaded model of the CaM-FF_3 complex using Modeller v10.[87] First, the sequence of the FF_3 domain was aligned to the Munc13-1 peptide (**Figure 2.11**). An emphasis was placed on the key hydrophobic residues that are known to serve as critical anchors that extend into the hydrophobic pockets of Ca^{2+}-loaded CaM. Second, the FF_3 sequence was threaded through the structure of the CaM-Munc13-1 complex. Since the P(r) is dominated by the position of the CaM domains, the fit to the experimental data for the CaM-FF_3 complex is very similar to that of the CaM-Munc13-1 complex (**Figure 2.10 D**). The model also fits well to the *ab initio* electron map of the CaM-FF_3 complex generated from the SAXS data. Together, the structural data indicate that CaM binds in an extended mode with both domains engaged with an unfolded state of hPrp40A FF_3.

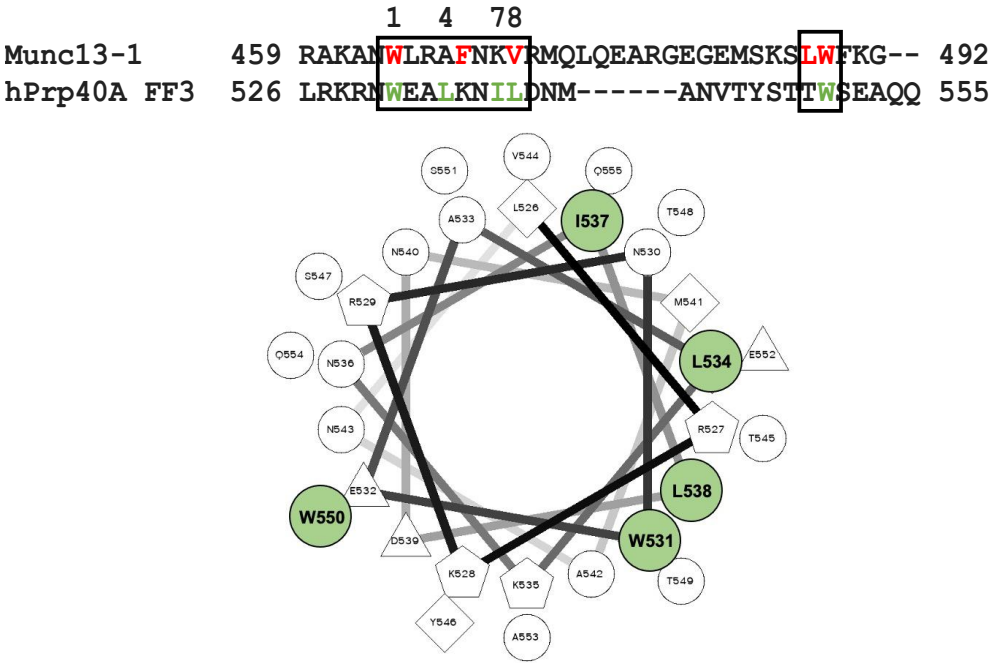

Figure 2.11. Identification of the two CaM-binding sites of the hPrp40A FF_3 domain.
(A) Sequence alignment of the CaM-binding motifs of hPrp40A FF_3 with Munc13-1. Hydrophobic residues of Munc13-1 known to interact with CaM based on the 3D structure are highlighted in **red**. The two CaM-binding motifs of Munc13-1 are highlighted by boxes. The residues of the hPrp40A FF_3 predicted to interact with CaM are highlighted in **green**. The alignment of hPrp40A FF_3 residues 526-555 (Uniprot ID: O75400) with Munc13-1 residues 459-492 (Uniprot ID: Q62768) was generated using the online tool Clustal Omega. **(B)** Helical wheel of the FF_3 comprising the sequence from Leu526 to Gln555. The predicted hydrophobic residues of the $W_1xxL_4xxI_7L_8$ motif are identified in **green**. The Trp550 residue is also highlighted in green. The helical wheel representation was generated using the online resource at https://www.donarmstrong.com/cgi-bin/wheel.pl.

CaM binding is anchored by two tryptophan residues in the hPrp40A FF_3 domain.

Examination of the alignment of the sequences of hPrp40A FF_3 and the Munc13-1 peptide (**Figure 2.11**) provided further insights into the molecular details of the CaM-FF_3 complex. Instead of the $W_1xxxF_5xxV_8$ motif found in Munc13-1, the first binding site predicted to mediate the interaction with CaM is comprised of a $W_1xxL_4xxI_7L_8$ motif, which overlaps with the Cen2-binding site.[57] A second binding site is predicted by aligning a tryptophan in the FF_3 to the tryptophan of Munc13-1 that anchors binding to CaM N-terminal domain. Considering the possibility that both CaM-binding motifs in the FF_3 preserve the helical confirmations observed in the Munc13-1 complex upon their binding with CaM, we proceeded to generate a helical wheel to visualize the spatial distribution of the $W_1xxL_4xxI_7L_8$ motif. In this projection, the hydrophobic residues part of the $W_1xxL_4xxI_7L_8$ motif are located on one face of the α-helix (**Figure 2.11**). In the homology model, the predicted hydrophobic residues including this motif and an additional Trp residue are facing toward the hydrophobic pockets

of the CaM domains (**Figure 2.12**), which include two FF_3 Trp residues (Trp531, Trp550) primed to serve as the key hydrophobic anchors for binding.

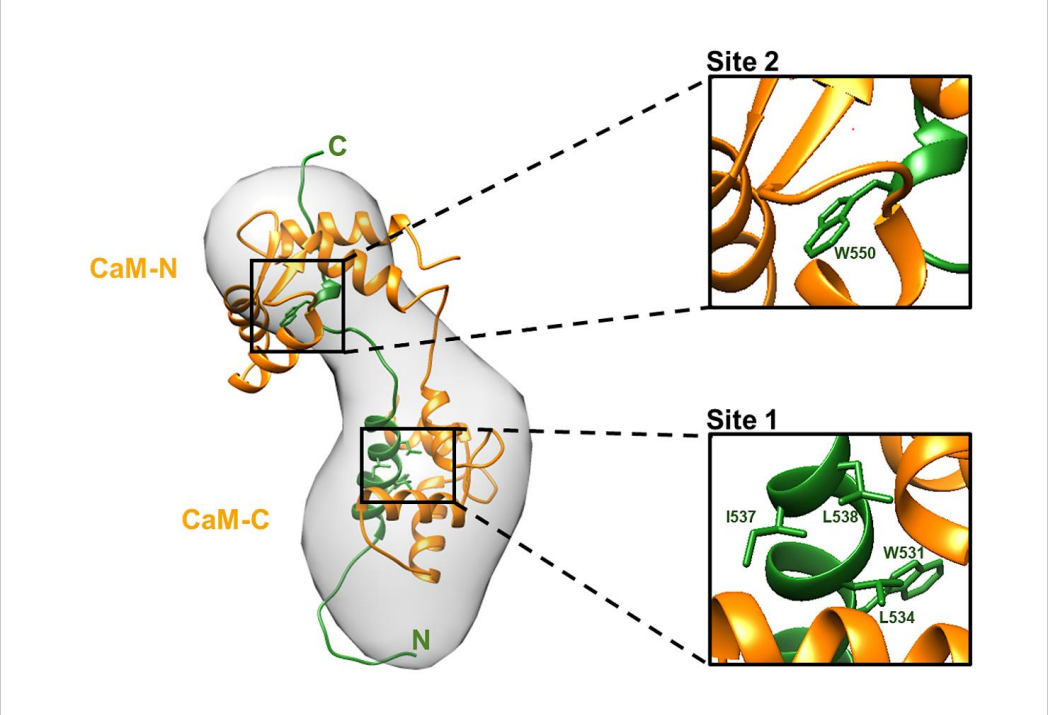

Figure 2.12. Structure of the CaM-FF_3 complex.
Overlay of the *ab initio* electron map of the CaM-FF_3 complex generated from SAXS and the 3D homology model. Black boxes present a close-up view of the two binding sites. CaM and FF_3 are colored orange and green, respectively. Ca^{2+} ions are represented with green spheres. The structures were generated using UCSF Chimera.[99]

To test the hypothesis that the two Trp residues serve as anchors for binding to CaM, we performed an intrinsic Trp fluorescence experiment. This approach was feasible because the FF_3 contains only these two Trp residues whereas CaM has none. For this experiment, Trp fluorescence emission spectra were recorded in the absence and presence of increasing amounts of Ca^{2+}-loaded CaM (**Figure 2.13**). Observation of maximum fluorescence emission at 354 nm for the isolated FF_3 indicates that the Trp side chains are at least partially exposed to the aqueous solvent. Additions of Ca^{2+}-loaded CaM result in smoothly increasing emission intensity and a blue shift, which reaches 337 nm in saturating conditions. These changes in the fluorescence spectrum indicate both Trp residues are sequestered from solvent upon binding to CaM, consistent with both Trp residues binding into hydrophobic pockets. The combination of the NMR, SAXS and these data support a model in which the binding to CaM of the hPrp40A FF_3 domain is coupled to unfolding and mimics the extended structure of CaM in its complex with Munc13-1.[61]

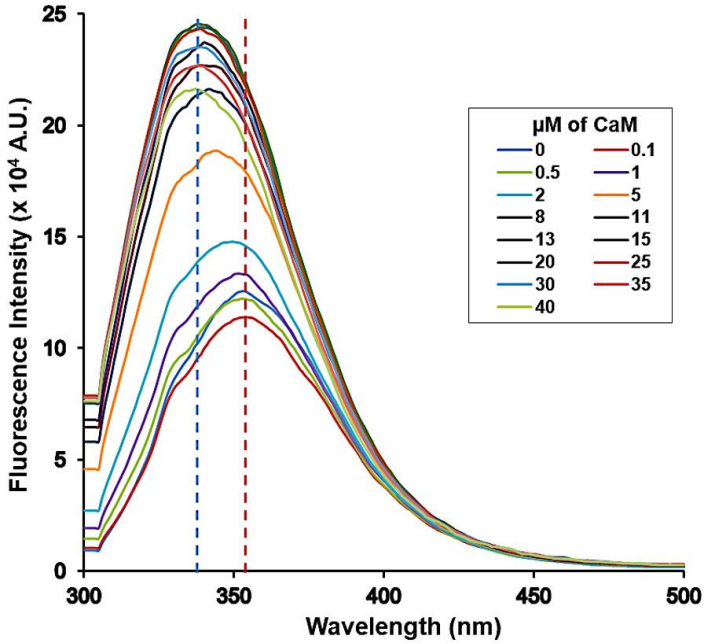

Figure 2.13. Titration of the hPrp40A FF_3 domain with CaM monitored by intrinsic fluorescence. Tryptophan fluorescence emission of FF_3 (5 μM) recorded at 37 °C in the absence and presence of increasing amount of CaM. Note the shift in the maximum emission of the FF_3 from 354 nm (red lines) to 337 nm (blue lines) upon saturation with CaM.

Mutational analysis suggests a dual tryptophan anchor motif is vital for FF_3 binding to CaM. To further test the hypothesis that the two tryptophan residues serve as the anchors for binding into the hydrophobic pockets of the two CaM domains, we generated three FF_3 constructs with tryptophan to alanine substitutions and assayed their effect on CaM binding with ITC and NMR: W531A, W550A, and the double mutant W531A,W550A. To directly monitor if the mutations result in differences in binding affinity, ITC experiments were performed under conditions optimized for the WT FF_3 domain (**Figure 2.14**). The thermograms show that the single-site mutants W531A and W550A bind more weakly to CaM than the WT, to the point that saturation is not yet reached at 4 equivalents of the mutated FF_3 (**Figure 2.14 B, C**). Moreover, the binding of the double mutant is so weak that there is effectively no evidence of binding to CaM under these conditions (**Figure 2.14 D**).

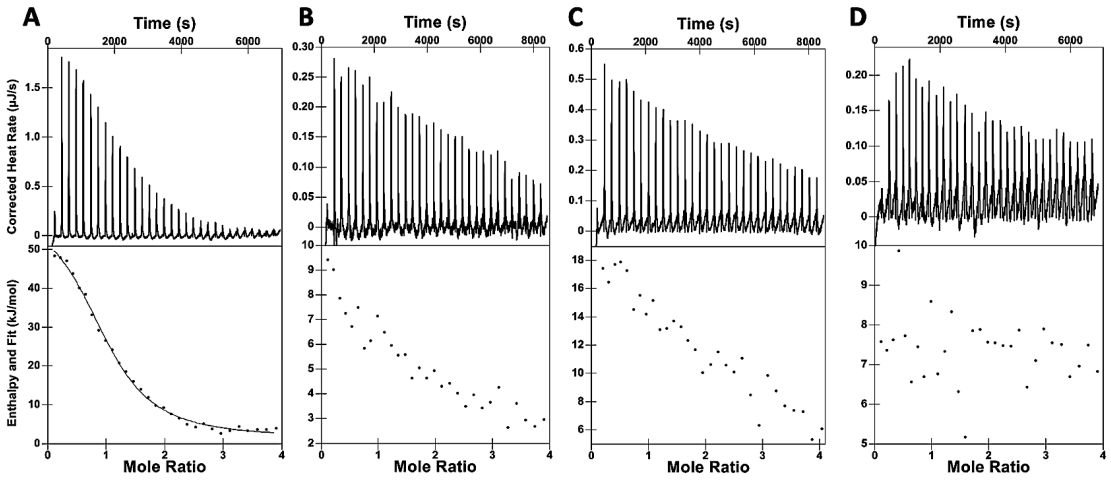

Figure 2.14. ITC analysis of the interaction of hPrp40A-FF_3 variants and CaM.
Representative plots for ITC runs are shown for the titration of (A) wt hPrp40A-FF_3 with CaM, (B) hPrp40A-FF_3 W531A with CaM, (C) hPrp40A-FF_3 W550A with CaM, and (D) hPrp40A-FF_3 W531A/W550A with CaM.

The effects of these FF_3 variants on binding to CaM was also tested by 2D ^{15}N-^{1}H HSQC NMR. As noted above, titration of WT FF_3 into ^{15}N-enriched CaM results in general line broadening and a substantial number of chemical shift perturbations in the spectrum. **Figure 2.15** shows a comparison of a region of the CaM spectrum containing signals for 8 residues (E7, E8, F17, L33, E46, F66, E120 and E141) before and after addition of two molar equivalents of the WT FF_3 domain or one of the mutants. In the case of the WT protein, four signals shift, one is broadened beyond detection and four remain unchanged. After addition of either the W531A or W550A mutant the effects mimic those observed for the WT FF_3. However, the extent of the perturbations of the spectrum is far more modest with very small chemical shift perturbations for two residues, and the rest remaining unchanged with no significant line broadening (**Figure 2.15 B,C**). Moreover, addition of the W531A/W550A double mutant has no effect on the spectrum of CaM indicating there is no binding under these conditions (**Figure 2.15 D**).

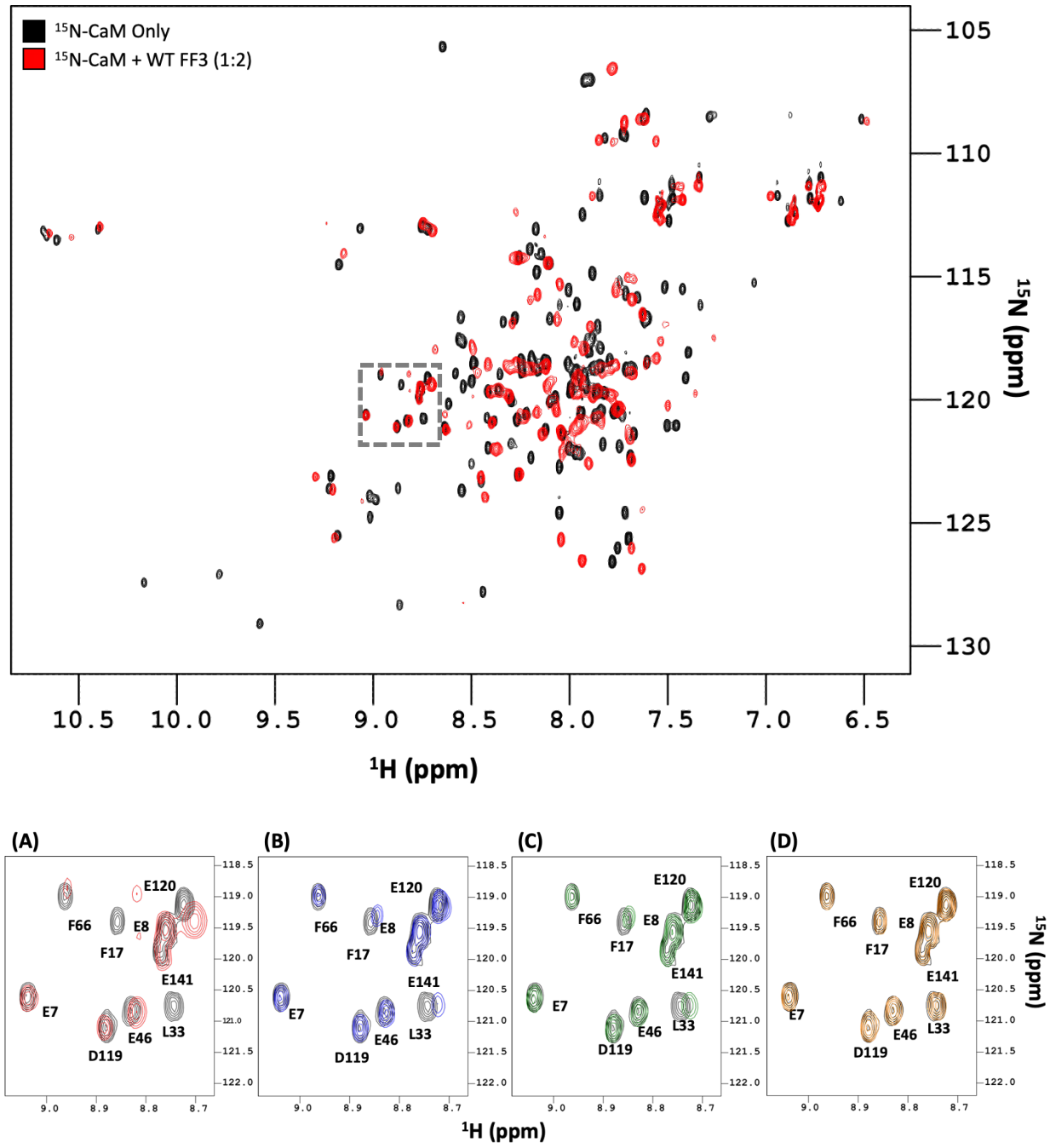

Figure 2.15. NMR analysis of the interaction between hPrp40A-FF$_3$ variants and CaM.
900 MHz 2D ^{15}N-^{1}H HSQC NMR spectra at 37 °C of ^{15}N-enriched Ca^{2+}-loaded CaM alone (**black**) and with 2-fold excess of (**A**) WT FF$_3$ (**red**), (**B**) FF$_3$ W531A (**blue**), (**C**) FF$_3$ W550A (**green**), and (**D**) FF$_3$ W531A/W550A (**orange**). The sequence-specific assignment is provided for crosspeaks from residues in both the N- and C-terminal domains of CaM to highlight changes in line broadening and peak shifts across different FF$_3$ variants. For clarity, full spectra only shown for ^{15}N-CaM and WT FF$_3$.

Even increasing the W531A or W550A to CaM ratio to 20:1 does not produce perturbations of the CaM spectrum comparable to the 2:1 ratio of WT FF_3 to CaM (**Figure 2.16**), fully consistent with the ITC data which show that the binding of the FF_3 mutants is substantially weaker. Together, the NMR and ITC data confirm that Trp531 and Trp550 serve as the key hydrophobic anchors for the binding of the Prp40A FF3 domain to CaM.

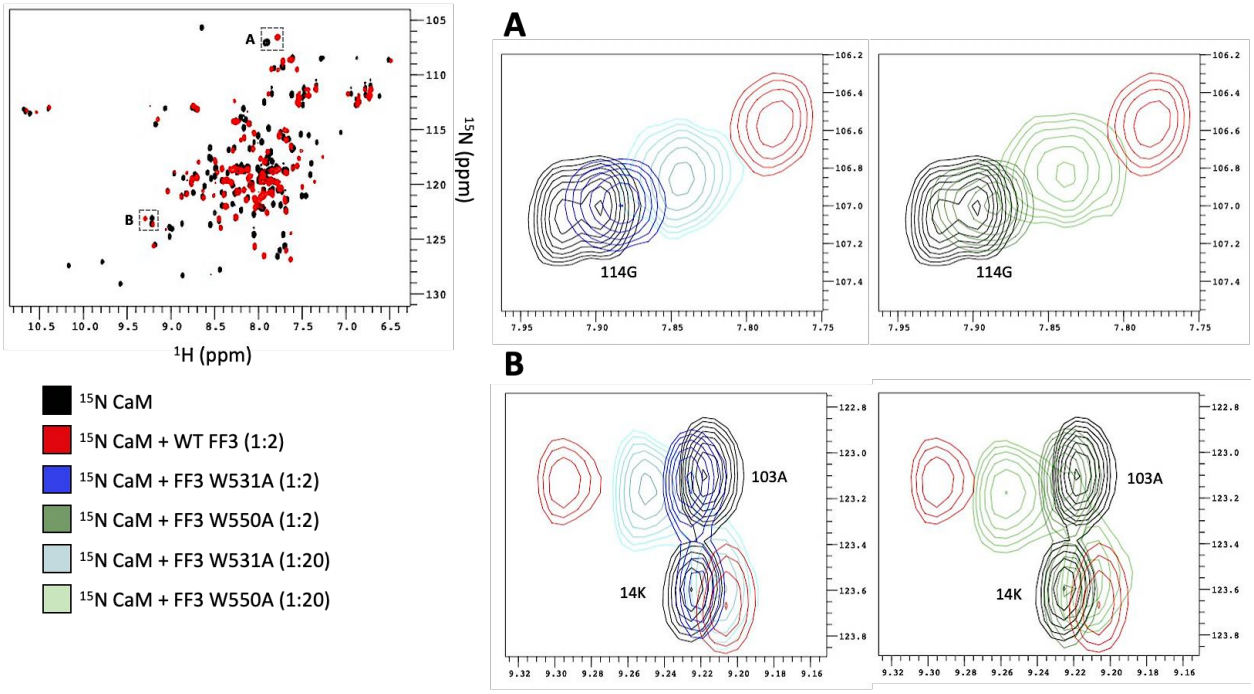

Figure 2.16. NMR analysis of the 20-fold excess of hPrp40A-FF_3 single-residue variants and CaM. 900 MHz 2D ^{15}N-^{1}H HSQC NMR spectra at 37 °C of ^{15}N-enriched Ca^{2+}-loaded CaM alone (**black**) and with 2-fold excess of WT FF_3 (**red**), FF_3 W531A (**blue**), and FF_3 W550A (**green**). Additional spectra were collected with 20-fold excess of FF_3 W531A (**cyan**) and FF_3 W550A (**sea green**). The sequence-specific assignment for two selected regions (**A, B**) is provided for crosspeaks from residues in both the N- and C-terminal domains of CaM to highlight changes in line broadening and peak shifts across different FF_3 variants.

2.4 Discussion

The interaction between Prp40 and CaM was first reported in a *Saccharomyces cerevisiae* protein microarray.[76] Our results show that like hCen2, CaM binds the hPrp40A FF_3 domain in a Ca^{2+}-dependent manner. As anticipated, the overall structure of hPrp40A FF_3 is similar to other FF domains. However, sequence alignment of hPrp40A FF_3 reveals a unique feature in the loop-3_{10}-loop region, i.e. substitution in hPrp40A FF_3 of an aspartate residue for a conserved hydrophobic residue whose side chain packs into the hydrophobic core and presumably stabilizes the loop conformation. Moreover, this aspartate is surrounded by three glutamatic acid residues. The

combination of the strong electrostatic field associated with this cluster and the loss of the hydrophobic interaction implies this loop contributes to a specific and distinct function of this FF domain.

Like most Ca^{2+}-dependent interactions of CaM with targets, both domains of CaM are engaged in binding hPrp40A FF_3. Sequence alignment of the FF_3 domain with Munc13-1 allowed us to predict the two putative CaM-binding motifs containing key hydrophobic anchors. The FF_3 residues Trp531, Leu534, Ile537, and Leu538 are the key residues in the $W_1xxL_4xxI_7L_8$ motif previously determined as the interaction site for hCen2.[57] A second CaM-binding site could also be identified containing Trp550. Based on this alignment and the structural similarities with the CaM-Munc13-1 complex, we assume hPrp40A FF_3 binds to CaM in the same antiparallel mode. The antiparallel orientation is common, having been reported for other CaM-target complexes that exhibit a bipartite binding mode, including the HIV-1 MA[107] and Na^+/H^+ exchanger 1 (NHE1).[108]

In contrast to CaM binding to the FF_3, hCen2 binds solely through its C-terminal domain, which correlates with the absence of functional Ca^{2+}-binding sites in the N-terminal domain.[57,59] Despite these differences, in both cases access to key hydrophobic resides requires unfolding of the FF3 domain. The need for unfolding for functional interaction with an FF domain has been reported previously in one case, for the p190A RhoGAP FF_1, in which access to Tyr308 for phosphorylation by the PDGF receptor α kinase requires unfolding of the FF domain.[94]

CaM binds its targets in a range of modes.[59] The most common is termed the wrap-around mode in which both domains envelop a common motif of ~15 residues, first identified in binding to the myosin light chain kinase.[109] Over the years binding in extended modes have been increasingly observed. These includes cases where the CaM domains engage different motifs within the same domain or subunit of the target, the same motif in two target molecules, or two different subunits of a complex.[59] Analysis of our SAXS data revealed that the Prp40A FF_3 domain binds to CaM in a very similar mode as observed for Munc13-1 and several other complexes including the CaM complex with the HIV-1 MA protein.[107] However, to our knowledge, the requirement for unfolding of the binding partner has only been seen in the binding of HIV-1 MA to CaM.[107] Remarkably, CaM binding to MA clearly disrupts its structure, which in turn inhibits its interactions with other proteins critical to HIV-1 replication. In contrast, our findings suggest that unfolding of Prp40A FF_3 stimulates its interaction with partner proteins at the core of the putative Ca^{2+}-dependent regulatory network of Prp40.

The observation that two different EF-hand Ca^{2+} sensor proteins have different binding modes to the FF_3 domain implies that the mechanism of Ca^{2+}-dependent regulation of Prp40 is complex. Coupled response of two sensors to intracellular Ca^{2+} signals has been characterized for the complex modulation of voltage gated Na^+ channels by CaM, with models proposing the involvement of multiple CaM molecules or CaM shifting between different sites on the channel in response to Ca^{2+} signals.[110] With both Cen2 and CaM binding to hPrp40A FF_3, a similar mechanism may be at play with Cen2 and CaM bound to Prp40 at different times over the course of a cycle of Ca^{2+} signaling.

Dysregulated protein-protein interactions of huntingtin (HTT) are often observed in Huntington's Disease (HD) patients. Previous studies have implied that both hPrp40A and CaM are involved in HD based on their interactions with WT and disease-associated mutants of HTT.[61-63,76-78] hPrp40A binds to both HTT and CaM utilizing its two WW domains, the FF_3 domain and possibly other FF domains. CaM is also known to bind directly to HTT. We have shown CaM uses both of its domains to engage hPrp40A-FF_3, so either multiple CaM molecules are recruited to this network, or a single CaM molecule is exchanging between different sites over the dynamic trajectory of the HTT-Prp40-CaM-Cen2 machinery. Clearly, further studies are required to clarify the action of CaM and Cen2 within this interaction network and establish if this machinery represents a viable potential therapeutic and diagnostic target for HD.

CHAPTER 3: CONCLUSIONS AND FUTURE DIRECTIONS‡

3.1 Overall Summary of the Biophysical Investigation of CaM and hPrp40A-FF_3

Our studies show that CaM binds in an extended mode to an unfolded state of the globular hPrp40A FF_3 domain. The unfolding of the hPrp40A FF_3 allows the exposure of hydrophobic residues buried within its core in the folded state that mediate binding of two distinct motifs with Trp anchors into the hydrophobic pockets of the two CaM domains. The favorable entropy of binding detected by ITC suggests that binding to CaM helps drive the FF_3 to unfold. Our results imply that binding to a lowly populated unfolded state may be a more general regulatory mechanism than previously appreciated to regulate FF domain-mediated interaction and function.

In summary, these findings have helped elucidate the biophysical and structural features needed to facilitate CaM binding to hPrp40A-FF_3. Broadly, this study has begun to look at potential interactions between calcium sensors and hPrp40A. Previous work in the Chazin Lab had demonstrated that hCen2 also interacts hPrp40A-FF_3.[57] The calcium sensors CaM and hCen2 as well as the pre-mRNA processing factor hPrp40A have been reported to interact with huntingtin protein (HTT).[55-57]A mutant form of HTT (mHTT) is responsible for causing pathogenicity in Huntington's Disease (HD).[38,39] However, the molecular interactions of mHTT in HD are not well understood, so a better understanding of these interactions occurring at the HTT interface may help determine how the disease manifests biochemically and which pathways are affected by mHTT. Looking at this more holistically, this study has laid the foundation for further work investigating potential CaM and Cen2 interactions with the remaining Prp40A FF domains. The longer-term goal would be to characterize the CaM-hCen2-hPrp40A-FF_3 network and its interactions with HTT, defining how calcium sensors are coordinating at and modulating the huntingtin (HTT) protein interactome. There may also be some implications for studying the effects of HTT on mRNA editing since hPrp40A appears to be an HTT interactor.[62-64]

‡ Section 3.1 from this chapter is adapted from a recent submission to *Biophysical Journal* as Díaz Casas A., Ferrer BJ, Cordoba JJ, Balakrishnan S, Wurm JE, Pastrana-Ríos B, and Chazin WJ., Binding by calmodulin requires unfolding of the Prp40A FF_3 domain, 2022.

3.2 Future Directions

Given that this work has focused on the isolated domains of Prp40A and CaM, it is important to extend these studies to a broader structural and cellular context. Certainly, future work will need to address the full-length Prp40A as well as incorporate the other calcium sensor hCen2. Additionally, to achieve the ultimate goal, the physical and functional correlation with HTT will need be to explored. Interestingly, the observation from this study that CaM engages hPrp40A-FF_3 with both domains suggests that there needs to be some type of exchange of CaM or a second CaM molecule in order to also interact with HTT. One model would imply hPrp40A assists in recruitment of CaM to HTT. A similar inference could be made for hCen2 as well. However, given that most Ca^{2+}-dependent pathways utilize only one calcium sensor, the first aspect to be investigated is if one Ca^{2+} sensor is more vital to regulation of HTT. Perhaps even more important than these studies of the physical interactions, it is imperative to establish the biological relevance of the proposed Ca^{2+}-dependent regulation in cells and *in vivo* using techniques well outside the scope of those in this thesis.

For example, this work could be expanded to include cell-based assays. Initially, we could engineer a cell line that stably expresses endogenous WT HTT or mHTT (with varying degrees of CAG repeats in order to measure the effects of mHTT polyglutamine length), as well as CaM, hCen2, and Prp40A. Using this cell line, we could treat the cells with a CRISPR interference (CRISPRi)[111] screen in order to knockdown CaM or hCen2 and measure the effects of gene expression, cell growth, and cell morphology. Transcription levels could be measured using real-time quantitative PCR (RT-qPCR). Cell growth and cell morphology could be monitored using confocal microscopy. The knockdown of CaM or hCen2 might have other effects beyond disrupting putative interactions with mHTT given the importance of Ca^{2+}-binding proteins within the cell, so the experiment would most likely need to be optimized to reduce effects of disrupting calcium sensors. This could be accomplished, for example, by CRISPR editing residues critical to interaction with Prp40A or complementing after knock-down with the corresponding mutant proteins. Another idea would be to supplement additional calcium sensors that don't interact with mHTT but can still regulate Ca^{2+} without CaM or hCen2. A fundamentally different approach would be to further engineer the proposed cell line and include fluorescent fusions of each of the proteins in our proposed system [mHTT, CaM, hCen2, and Prp40A]. Using orthogonal fluorescent probes, we could trace if these proteins are interacting in real time using confocal microscopy to visualize colocalization of the proteins.[112]

Beyond the work presented here, it is interesting to speculate how other structural tools could be used to probe this interaction network further. Testing to determine if the other FF domains of hPrp40A interact with CaM is a natural next step. After identification of other FF CaM-interacting domains, further work would be needed to characterize if they interact in a similar fashion as the FF_3 domain. Another driving question is how might the results obtained for fragments or full-length hPrp40A as well as for CaM and Cen2 be altered by the presence of huntingtin (HTT)?

The relative mass of each of these proteins makes them well poised for other structural studies: hPrp40A MW ~180.8 kDa, WT HTT MW ~347.6 kDa, CaM MW ~16.8 kDa, hCen2 MW ~19.7 kDa (monomeric). In particular, does function require that more than one CaM interacts with hPrp40A and HTT? The mass of these potential complexes would be much larger than those studied so far, well within the ideal range for cryogenic electron microscopy (cryoEM) and SAXS. We should be able to use both approaches to analyze not only the various permutations of complexes, but also the HTT and hPrp40A individual components.

SAXS does not provide high resolution structures and the resolution obtainable by cryoEM varies from system to system, typically not reaching atomic resolution (< 2.5 A). Since it is common to fit SAXS and CryoEM electron density with high resolution structures of single proteins and sub-complexes, co-crystallization of the individual FF domains with CaM would be useful as templates for modeling in data obtained for the full-length complexes. Attempts at crystallizing and determining structures of the hPrp40A FF_3 domain with CaM were performed, but I was unable to obtain diffracting crystals. However, I think with further optimization, obtaining co-crystal structures of multiple FF domains in complex with CaM or hCen2 should be possible. Individual structures of the binary complexes of CaM with each FF domain should be able to assist with modeling of the putative ternary CaM-hPrp40A-HTT (or hCen2-hPrp40A-HTT) complex by cryoEM.

APPENDIX

Supplemental Figures and Tables.

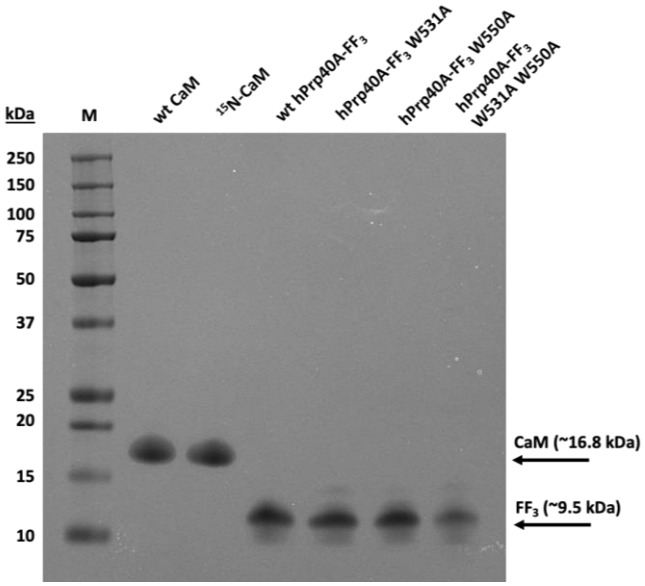

Figure S1. SDS-PAGE gel of CaM and FF_3 protein stocks.
20 μL of each stock (at 50 μM) was mixed 1:1 with 2x SDS-PAGE loading buffer and 10 μL was run on a 4-12% SDS denaturing gel. We noticed that the FF_3 double variant had a lower band intensity. We used this construct at lower and higher concentrations and noticed no impact on binding affinity via ITC and NMR and concluded that this estimation was sufficient for our mutational analysis. Note that FF_3 constructs are running higher than expected.

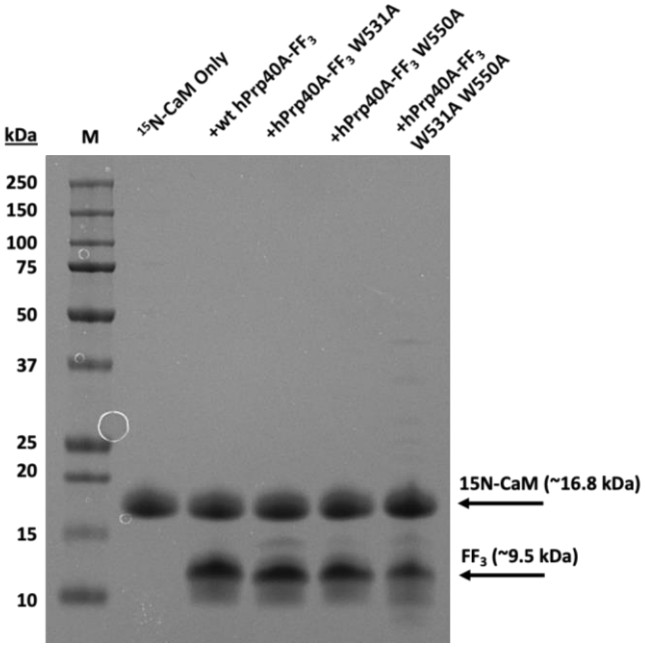

Figure S2. SDS-PAGE gel of ^{15}N-CaM and FF_3 NMR samples.
NMR samples were prepared with 50 μM CaM + 100 μM FF_3 construct. 10 μL of the prepared NMR sample was mixed 1:1 with 2x SDS-PAGE loading buffer and 10 μL was run on a 4-12% SDS denaturing gel.

Table S1. Plasmids used in this thesis.
Annotated plasmid maps are available by following the provided links.

Name	Use	Resistance Marker
pET-15b-CaM https://benchling.com/s/seq-csn8hBUbVdpbXAgRIS0f	Calmodulin expression	Ampicillin
pBG100-FF_3 https://benchling.com/s/seq-7xk58E7iShrMUZRFqTGM?m=slm-zsYgWtmKbRMwkPSZ2rzp	hPrp40A-FF_3 expression	Kanamycin
pBG100-FF_3 W531A https://benchling.com/s/seq-XwjTApNsoGB7PbVM24t4?m=slm-vFaGPmigEFTdDucNsUUi	FF_3 Site 1 Substitution expression	Kanamycin
pBG100-FF_3 W550A https://benchling.com/s/seq-kdtdFpN4HV4ZSfLuLMnM?m=slm-mmRgRmGivU87dUxQfSoM	FF_3 Site 2 substitution expression	Kanamycin
pBG100-FF_3 W531A W550A https://benchling.com/s/seq-jAiMJSyggz5lP6oWZh9x?m=slm-rxFh1bsRvMy7KfEk41LD	FF_3 double site substitution expression	Kanamycin

Table S2. Biochemical parameters for proteins used in this thesis.
For Trp to Ala mutants, the substitutions are **bold** in the sequence. Parameters were calculated with ProtParam (Expasy).[113]

Name	Amino Acid Sequence (N- to C-terminus)	MW (Da)	pI	ε (M^{-1} cm^{-1})
WT CaM	MADQLTEEQIAEFKEAFSLFDKDGDGTITTKELGTVMRSLGQN PTEAELQDMINEVDADGNGTIDFPEFLTMMARKMKDTDSEEEI REAFRVFDKDGNGYISAAELRHVMTNLGEKLTDEEVDEMIREA DIDGDGQVNYEEFVQMMTAK	16837.59	4.09	2980
^{15}N-CaM	(above)	(above)	(above)	(above)
WT hPrp40A-FF_3	MSKKEKEQAKQLRKRNWEALKNILDNMANVTYSTTWSEAQQY LMDNPTFAEDEELQNMDKEDALICFEEHIRALEKEEE	9469.54	4.59	13980
hPrp40A-FF_3 W531A	MSKKEKEQAKQLRKRN**A**EALKNILDNMANVTYSTTWSEAQQY LMDNPTFAEDEELQNMDKEDALICFEEHIRALEKEEE	9354.40	4.59	8480
hPrp40A-FF_3 W550A	MSKKEKEQAKQLRKRNWEALKNILDNMANVTYSTT**A**SEAQQY LMDNPTFAEDEELQNMDKEDALICFEEHIRALEKEEE	9354.40	4.59	8480
hPrp40A-FF_3 W531A W550A	MSKKEKEQAKQLRKRN**A**EALKNILDNMANVTYSTT**A**SEAQQYL MDNPTFAEDEELQNMDKEDALICFEEHIRALEKEEE	9239.27	4.59	2980

REFERENCES

1. Berridge, M. J., Lipp, P., & Bootman, M. D. (2000). The versatility and universality of calcium signalling. Nature Reviews Molecular Cell Biology, 1(1), 11–21. https://doi.org/10.1038/35036035
2. Chin, D., & Means, A. R. (2000). Calmodulin: A prototypical calcium sensor. Trends in Cell Biology, 10(8), 322–328. https://doi.org/10.1016/s0962-8924(00)01800-6
3. Junho, C. V., Caio-Silva, W., Trentin-Sonoda, M., & Carneiro-Ramos, M. S. (2020). An overview of the role of calcium/calmodulin-dependent protein kinase in cardiorenal syndrome. Frontiers in Physiology, 11. https://doi.org/10.3389/fphys.2020.00735
4. Monteith, G. R., Prevarskaya, N., & Roberts-Thomson, S. J. (2017). The calcium–cancer signalling nexus. Nature Reviews Cancer, 17(6), 373–380. https://doi.org/10.1038/nrc.2017.18
5. Lewit-Bentley, A., & Réty, S. (2000). EF-hand calcium-binding proteins. Current Opinion in Structural Biology, 10(6), 637–643. https://doi.org/10.1016/s0959-440x(00)00142-1
6. Kretsinger, R. H., & Nockolds, C. E. (1973). Carp muscle calcium-binding protein. Journal of Biological Chemistry, 248(9), 3313–3326. https://doi.org/10.1016/s0021-9258(19)44043-x
7. Kawasaki, H., Nakayama, S., & Kretsinger, R. H. (1998). Classification and evolution of EF-hand proteins. Biometals, 11(4), 277–295. https://doi.org/10.1023/a:1009282307967
8. Kakiuchi, S., & Yamazaki, R. (1970). Calcium dependent phosphodiesterase activity and its activating factor (PAF) from brain. Biochemical and Biophysical Research Communications, 41(5), 1104–1110. https://doi.org/10.1016/0006-291x(70)90199-3
9. Li, M. X., & Hwang, P. M. (2015). Structure and function of cardiac troponin C (TNNC1): Implications for heart failure, cardiomyopathies, and troponin modulating drugs. Gene, 571(2), 153–166. https://doi.org/10.1016/j.gene.2015.07.074
10. Parmacek, M. S., & Leiden, J. M. (1991). Structure, function, and regulation of troponin C. Circulation, 84(3), 991–1003. https://doi.org/10.1161/01.cir.84.3.991
11. Gross, M., & Kumar, R. (1987). The physiology and biochemistry of vitamin D-dependent calcium binding proteins. Steroid and Sterol Hormone Action, 371–394. https://doi.org/10.1007/978-1-4613-2073-9_20
12. Michalak, M., Corbett, E. F., Mesaeli, N., Nakamura, K., & Opas, M. (1999). Calreticulin: One protein, one gene, many functions. Biochemical Journal, 344(2), 281–292. https://doi.org/10.1042/bj3440281
13. Nishi, R., Sakai, W., Tone, D., Hanaoka, F., & Sugasawa, K. (2013). Structure-function analysis of the EF-hand protein centrin-2 for its intracellular localization and nucleotide excision repair. Nucleic Acids Research, 41(14), 6917–6929. https://doi.org/10.1093/nar/gkt434
14. Donato, R., R. Cannon, B., Sorci, G., Riuzzi, F., Hsu, K., J. Weber, D., & L. Geczy, C. (2012). Functions of S100 proteins. Current Molecular Medicine, 13(1), 24–57. https://doi.org/10.2174/1566524011307010024
15. Xia, C., Braunstein, Z., Toomey, A. C., Zhong, J., & Rao, X. (2018). S100 proteins as an important regulator of macrophage inflammation. Frontiers in Immunology, 8. https://doi.org/10.3389/fimmu.2017.01908

16. Holzinger, D., Foell, D., & Kessel, C. (2018). The role of S100 proteins in the pathogenesis and monitoring of autoinflammatory diseases. Molecular and Cellular Pediatrics, 5(1). https://doi.org/10.1186/s40348-018-0085-2
17. Jukic, A., Bakiri, L., Wagner, E. F., Tilg, H., & Adolph, T. E. (2021). Calprotectin: From Biomarker to biological function. Gut, 70(10), 1978–1988. https://doi.org/10.1136/gutjnl-2021-324855
18. Hurley, M. J., Brandon, B., Gentleman, S. M., & Dexter, D. T. (2013). Parkinson's disease is associated with altered expression of Cav1 channels and calcium-binding proteins. Brain, 136(7), 2077–2097. https://doi.org/10.1093/brain/awt134
19. Jiang, H., Hu, H., Tong, X., Jiang, Q., Zhu, H., & Zhang, S. (2011). Calcium-binding protein s100p and cancer: Mechanisms and clinical relevance. Journal of Cancer Research and Clinical Oncology, 138(1), 1–9. https://doi.org/10.1007/s00432-011-1062-5
20. Hodeify, R., Siddiqui, S. S., Matar, R., Vazhappilly, C. G., Merheb, M., Al Zouabi, H., & Marton, J. (2021). Modulation of calcium-binding proteins expression and cisplatin chemosensitivity by calcium chelation in human breast cancer MCF-7 cells. Heliyon, 7(1). https://doi.org/10.1016/j.heliyon.2021.e06041
21. Xu, C., Chen, H., Wang, X., Gao, J., Che, Y., Li, Y., Ding, F., Luo, A., Zhang, S., & Liu, Z. (2014). S100A14, a member of the EF-hand calcium-binding proteins, is overexpressed in breast cancer and acts as a modulator of HER2 signaling. Journal of Biological Chemistry, 289(2), 827–837. https://doi.org/10.1074/jbc.m113.469718
22. Colotti, G., Poser, E., Fiorillo, A., Genovese, I., Chiarini, V., & Ilari, A. (2014). SORCIN, a calcium binding protein involved in the multidrug resistance mechanisms in cancer cells. Molecules, 19(9), 13976–13989. https://doi.org/10.3390/molecules190913976
23. Marshak, D. R., Lukas, T. J., & Watterson, D. M. (1985). Drug-protein interactions: Binding of chlorpromazine to calmodulin, calmodulin fragments, and related calcium binding proteins. Biochemistry, 24(1), 144–150. https://doi.org/10.1021/bi00322a020
24. Elliott, W. J., & Ram, C. V. (2011). Calcium channel blockers. The Journal of Clinical Hypertension, 13(9), 687–689. https://doi.org/10.1111/j.1751-7176.2011.00513.x
25. Green, J. L., Moon, R. W., Whalley, D., Bowyer, P. W., Wallace, C., Rochani, A., Nageshan, R. K., Howell, S. A., Grainger, M., Jones, H. M., Ansell, K. H., Chapman, T. M., Taylor, D. L., Osborne, S. A., Baker, D. A., Tatu, U., & Holder, A. A. (2016). IMIDAZOPYRIDAZINE inhibitors of plasmodium falciparum calcium-dependent protein kinase 1 also target cyclic GMP-dependent protein kinase and heat shock protein 90 to kill the parasite at different stages of intracellular development. Antimicrobial Agents and Chemotherapy, 60(3), 1464–1475. https://doi.org/10.1128/aac.01748-15
26. Matsuda, S., & Koyasu, S. (2000). Mechanisms of action of Cyclosporine. Immunopharmacology, 47(2-3), 119–125. https://doi.org/10.1016/s0162-3109(00)00192-2
27. Lindner, P., Christensen, S. B., Nissen, P., Møller, J. V., & Engedal, N. (2020). Cell death induced by the ER stressor thapsigargin involves death receptor 5, a non-autophagic function of MAP1LC3B, and distinct contributions from unfolded protein response components. Cell Communication and Signaling, 18(1). https://doi.org/10.1186/s12964-019-0499-z

28. Lukyanetz, E. A., Shkryl, V. M., & Kostyuk, P. G. (2002). Selective blockade of N-type calcium channels by Levetiracetam. Epilepsia, 43(1), 9–18. https://doi.org/10.1046/j.1528-1157.2002.24501.x
29. Kubo, K., Matsuda, Y., Kase, H., & Yamada, K. (1984). Inhibition of calmodulin-dependent cyclic nucleotide phosphodiesterase by flunarizine, a calcium-entry blocker. Biochemical and Biophysical Research Communications, 124(2), 315–321. https://doi.org/10.1016/0006-291x(84)91555-9
30. Civitelli, R. (2013). Calcitonin in osteoporosis. Osteoporosis, 1839–1858. https://doi.org/10.1016/b978-0-12-415853-5.00079-0
31. Garcia, V., Perera, Y. R., & Chazin, W. J. (2022). A structural perspective on calprotectin as a ligand of receptors mediating inflammation and potential drug target. Biomolecules, 12(4), 519. https://doi.org/10.3390/biom12040519
32. Wilder, P. T., Varney, K. M., & Weber, D. J. (2019). Targeting S100 calcium-binding proteins with small molecule inhibitors. Methods in Molecular Biology, 291–310. https://doi.org/10.1007/978-1-4939-9030-6_19
33. Davis, J. P., Shettigar, V., Tikunova, S. B., Little, S. C., Liu, B., Siddiqui, J. K., Janssen, P. M. L., Ziolo, M. T., & Walton, S. D. (2016). Designing proteins to combat disease: Cardiac troponin C as an example. Archives of Biochemistry and Biophysics, 601, 4–10. https://doi.org/10.1016/j.abb.2016.02.007
34. Shettigar, V., Zhang, B., Little, S. C., Salhi, H. E., Hansen, B. J., Li, N., Zhang, J., Roof, S. R., Ho, H.-T., Brunello, L., Lerch, J. K., Weisleder, N., Fedorov, V. V., Accornero, F., Rafael-Fortney, J. A., Gyorke, S., Janssen, P. M., Biesiadecki, B. J., Ziolo, M. T., & Davis, J. P. (2016). Rationally engineered troponin C modulates in vivo cardiac function and performance in health and disease. Nature Communications, 7(1). https://doi.org/10.1038/ncomms10794
35. Calvo-Rodriguez, M., Kharitonova, E. K., & Bacskai, B. J. (2020). Therapeutic strategies to target calcium dysregulation in alzheimer's disease. Cells, 9(11), 2513. https://doi.org/10.3390/cells9112513
36. Haenig, C., Atias, N., Taylor, A. K., Mazza, A., Schaefer, M. H., Russ, J., Riechers, S.-P., Jain, S., Coughlin, M., Fontaine, J.-F., Freibaum, B. D., Brusendorf, L., Zenkner, M., Porras, P., Stroedicke, M., Schnoegl, S., Arnsburg, K., Boeddrich, A., Pigazzini, L., … Wanker, E. E. (2020). Interactome mapping provides a network of neurodegenerative disease proteins and uncovers widespread protein aggregation in affected brains. Cell Reports, 32(7), 108050. https://doi.org/10.1016/j.celrep.2020.108050
37. Saudou, F., & Humbert, S. (2016). The Biology of Huntingtin. Neuron, 89(5), 910–926. https://doi.org/10.1016/j.neuron.2016.02.003
38. Kremer, B., Goldberg, P., Andrew, S. E., Theilmann, J., Telenius, H., Zeisler, J., Squitieri, F., Lin, B., Bassett, A., Almqvist, E., Bird, T. D., & Hayden, M. R. (1994). A worldwide study of the Huntington's disease mutation: The sensitivity and specificity of measuring CAG repeats. New England Journal of Medicine, 330(20), 1401–1406. https://doi.org/10.1056/nejm199405193302001
39. Warby, S. C., Montpetit, A., Hayden, A. R., Carroll, J. B., Butland, S. L., Visscher, H., Collins, J. A., Semaka, A., Hudson, T. J., & Hayden, M. R. (2009). CAG expansion in the Huntington disease gene is associated with a specific and targetable predisposing haplogroup. The American Journal of Human Genetics, 84(3), 351–366. https://doi.org/10.1016/j.ajhg.2009.02.003

40. Migliore, S., Jankovic, J., & Squitieri, F. (2019). Genetic counseling in Huntington's disease: Potential new challenges on Horizon? Frontiers in Neurology, 10. https://doi.org/10.3389/fneur.2019.00453
41. Crowell, V., Houghton, R., Tomar, A., Fernandes, T., & Squitieri, F. (2021). Modeling manifest huntington's disease prevalence using diagnosed incidence and survival time. Neuroepidemiology, 55(5), 361–368. https://doi.org/10.1159/000516767
42. Aziz, N. A., Roos, R. A., Gusella, J. F., Lee, J.-M., & MacDonald, M. E. (2012). CAG repeat expansion in Huntington disease determines age at onset in a fully dominant fashion. Neurology, 79(9), 952–953. https://doi.org/10.1212/wnl.0b013e3182697986
43. Bourn, D. (2021). Autosomal dominant inheritance and Huntington disease. Diagnostic Genetic Testing, 23–35. https://doi.org/10.1007/978-3-030-85510-9_2
44. Komatsu, H. (2021). Innovative therapeutic approaches for huntington's disease: From nucleic acids to GPCR-targeting small molecules. Frontiers in Cellular Neuroscience, 15. https://doi.org/10.3389/fncel.2021.785703
45. Byun, S., Lee, M., & Kim, M. (2022). Gene therapy for Huntington's disease: The final strategy for a cure? Journal of Movement Disorders, 15(1), 15–20. https://doi.org/10.14802/jmd.21006
46. Tabrizi, S. J., Leavitt, B. R., Landwehrmeyer, G. B., Wild, E. J., Saft, C., Barker, R. A., Blair, N. F., Craufurd, D., Priller, J., Rickards, H., Rosser, A., Kordasiewicz, H. B., Czech, C., Swayze, E. E., Norris, D. A., Baumann, T., Gerlach, I., Schobel, S. A., Paz, E., … Lane, R. M. (2019). Targeting Huntingtin expression in patients with Huntington's disease. New England Journal of Medicine, 380(24), 2307–2316. https://doi.org/10.1056/nejmoa1900907
47. Heinz, A., Schilling, J., van Roon-Mom, W., & Krauß, S. (2021). The MID1 protein: A promising therapeutic target in Huntington's disease. Frontiers in Genetics, 12. https://doi.org/10.3389/fgene.2021.761714
48. Hegde, R. N., Chiki, A., Petricca, L., Martufi, P., Arbez, N., Mouchiroud, L., Auwerx, J., Landles, C., Bates, G. P., Singh-Bains, M. K., Dragunow, M., Curtis, M. A., Faull, R. L. M., Ross, C. A., Caricasole, A., & Lashuel, H. A. (2020). TBK1 phosphorylates mutant Huntingtin and suppresses its aggregation and toxicity in Huntington's disease models. The EMBO Journal, 39(17). https://doi.org/10.15252/embj.2020104671
49. Song, H., Wang, C., Zhu, C., Wang, Z., Yang, H., Wu, P., Cui, X., Botas, J., Dang, Y., Ding, Y., Fei, Y., & Lu, B. (2022). Suppression of toxicity of the mutant huntingtin protein by its interacting compound, desonide. Proceedings of the National Academy of Sciences, 119(10). https://doi.org/10.1073/pnas.2114303119
50. Flavell, L. (2018, July 26). Treatments for Huntington's disease. Huntington's Disease News. Retrieved May 29, 2022, from https://huntingtonsdiseasenews.com/treatments/
51. Tourette, C., Li, B., Bell, R., O'Hare, S., Kaltenbach, L. S., Mooney, S. D., & Hughes, R. E. (2014). A large scale huntingtin protein interaction network implicates Rho GTPase signaling pathways in Huntington disease. Journal of Biological Chemistry, 289(10), 6709–6726. https://doi.org/10.1074/jbc.m113.523696
52. Harding, R. J., Deme, J. C., Hevler, J. F., Tamara, S., Lemak, A., Cantle, J. P., Szewczyk, M. M., Begeja, N., Goss, S., Zuo, X., Loppnau, P., Seitova, A., Hutchinson, A., Fan, L., Truant, R., Schapira, M., Carroll, J. B., Heck, A. J., Lea, S. M., & Arrowsmith, C. H. (2021). Huntingtin structure is orchestrated by hap40 and shows a

polyglutamine expansion-specific interaction with exon 1. Communications Biology, 4(1). https://doi.org/10.1038/s42003-021-02895-4

53. Guo, Q., Huang, B., Cheng, J., Seefelder, M., Engler, T., Pfeifer, G., Oeckl, P., Otto, M., Moser, F., Maurer, M., Pautsch, A., Baumeister, W., Fernández-Busnadiego, R., & Kochanek, S. (2018). The cryo-electron microscopy structure of Huntingtin. Nature, 555(7694), 117–120. https://doi.org/10.1038/nature25502
54. Tabrizi, S. J., Leavitt, B. R., Landwehrmeyer, G. B., Wild, E. J., Saft, C., Barker, R. A., Blair, N. F., Craufurd, D., Priller, J., Rickards, H., Rosser, A., Kordasiewicz, H. B., Czech, C., Swayze, E. E., Norris, D. A., Baumann, T., Gerlach, I., Schobel, S. A., Paz, E., … Lane, R. M. (2019). Targeting Huntingtin expression in patients with Huntington's disease. New England Journal of Medicine, 380(24), 2307–2316. https://doi.org/10.1056/nejmoa1900907
55. Friedberg, F. (2006). Centrin isoforms in mammals. relation to Calmodulin. Molecular Biology Reports, 33(4), 243–252. https://doi.org/10.1007/s11033-006-9004-z
56. Dudek, N. L., Dai, Y., & Muma, N. A. (2008). Protective effects of interrupting the binding of calmodulin to Mutant Huntingtin. Journal of Neuropathology & Experimental Neurology, 67(4), 355–365. https://doi.org/10.1097/nen.0b013e31816a9e60
57. Díaz Casas, A., Chazin, W. J., & Pastrana-Ríos, B. (2017). Prp40 homolog A is a novel Centrin Target. Biophysical Journal, 112(12), 2529–2539. https://doi.org/10.1016/j.bpj.2017.03.042
58. Berchtold, M. W., Egli, R., Rhyner, J. A., Hameister, H., & Strehler, E. E. (1993). Localization of the human bona fide calmodulin genes CALM1, CALM2, and CALM3 to chromosomes 14q24-q31, 2p21.1-p21.3, and 19q13.2-q13.3. Genomics, 16(2), 461–465. https://doi.org/10.1006/geno.1993.1211
59. Tidow, H., & Nissen, P. (2013). Structural diversity of calmodulin binding to its target sites. FEBS Journal, 280(21), 5551–5565. https://doi.org/10.1111/febs.12296
60. Jiang, Y.-J., Che, M.-X., Yuan, J.-Q., Xie, Y.-Y., Yan, X.-Z., & Hu, H.-Y. (2011). Interaction with polyglutamine-expanded huntingtin alters cellular distribution and RNA processing of Huntingtin yeast two-hybrid protein A (HYPA). Journal of Biological Chemistry, 286(28), 25236–25245. https://doi.org/10.1074/jbc.m110.216333
61. Rodríguez-Castañeda, F., Maestre-Martínez, M., Coudevylle, N., Dimova, K., Junge, H., Lipstein, N., Lee, D., Becker, S., Brose, N., Jahn, O., Carlomagno, T., & Griesinger, C. (2009). Modular architecture of Munc13/calmodulin complexes: Dual regulation by ca2+ and possible function in short-term synaptic plasticity. The EMBO Journal, 29(3), 680–691. https://doi.org/10.1038/emboj.2009.373
62. Faber, P. (1998). Huntingtin interacts with a family of WW domain proteins. Human Molecular Genetics, 7(9), 1463–1474. https://doi.org/10.1093/hmg/7.9.1463
63. Passani, L. A. (2000). Huntingtin's WW domain partners in Huntington's disease post-mortem brain fulfill genetic criteria for direct involvement in Huntington's disease pathogenesis. Human Molecular Genetics, 9(14), 2175–2182. https://doi.org/10.1093/hmg/9.14.2175
64. Jiang, Y.-J., Che, M.-X., Yuan, J.-Q., Xie, Y.-Y., Yan, X.-Z., & Hu, H.-Y. (2011). Interaction with polyglutamine-expanded huntingtin alters cellular distribution and RNA processing of Huntingtin yeast two-

hybrid protein A (HYPA). Journal of Biological Chemistry, 286(28), 25236–25245. https://doi.org/10.1074/jbc.m110.216333

65. Plaschka, C., Lin, P.-C., Charenton, C., & Nagai, K. (2018). Prespliceosome structure provides insights into spliceosome assembly and Regulation. Nature, 559(7714), 419–422. https://doi.org/10.1038/s41586-018-0323-8
66. Li, X., Liu, S., Jiang, J., Zhang, L., Espinosa, S., Hill, R. C., Hansen, K. C., Zhou, Z. H., & Zhao, R. (2017). CryoEM structure of saccharomyces cerevisiae U1 snrnp offers insight into alternative splicing. Nature Communications, 8(1). https://doi.org/10.1038/s41467-017-01241-9
67. Choudhary, B., Marx, O., & Norris, A. D. (2021). Spliceosomal component PRP-40 is a central regulator of microexon splicing. Cell Reports, 36(5), 109464. https://doi.org/10.1016/j.celrep.2021.109464
68. Gasch, A., Wiesner, S., Martin-Malpartida, P., Ramirez-Espain, X., Ruiz, L., & Macias, M. J. (2006). The structure of Prp40 FF1 domain and its interaction with the CRN-TPR1 motif of CLF1 gives a new insight into the binding mode of FF domains. Journal of Biological Chemistry, 281(1), 356–364. https://doi.org/10.1074/jbc.m508047200
69. Kao, H-Y. and P.G. Siliciano. 1996. Identification of Prp40, a novel essential yeast splicing factor associated with the U1 small nuclear ribonucleoprotein particle. Mol. Cell. Biol. 16:960-967, doi: 10.1128/mcb.16.3.960.
70. Becerra, S., E. Andrés-León, S. Prieto-Sánchez, C. Hernández-Munain, and C. Suñé. 2016. Prp40 and early events in splice site definition. WIREs RNA. 7:17-32, doi: 10.1002/wrna.1312.
71. Li, X., S. Liu, L. Zhang, A. Issaian, R.C. Hill, S. Espinosa, S. Shi, Y. Cui, K. Kappel, R. Das, K.C. Hansen, Z.H. Zhou, and R. Zhao. 2019. A unified mechanism for intron and exon definition and back- splicing. Nature. 573: 375-380, doi: 10.1038/s41586-019-1523-6.
72. Buschdorf, J.P. and W.H. Strätling. 2004. A WW domain binding region in methyl-CpG-binding protein MeCP2: impact on Rett syndrome. J. Mol. Med. 82:135-143, doi: 10.1007/s00109-003-0497-9.
73. Chung, S., M.R. McLean, and B.C. Rymond. 1999. Yeast ortholog of the Drosophila crooked neck-protein promotes spliceosome assembly through stable U4/U6.U5 snRNP addition. RNA. 5: 1042-1054, doi: 10.1017/s1355838299990635.
74. Vincent, K., Q. Wang, S. Jay, K. Hobbs, and B.C. Rymond. 2003. Genetic interactions with CLF1 identify additional pre-mRNA splicing. Genetics. 164: 895-907, doi: 10.1093/genetics/164.3.895.
75. Ester, C. and P. Uetz. 2008. The FF domains of yeast U1 snRNP protein Prp40 mediate interactions with Luc7 and Snu71. BMC Biochemistry. 9: 1-11, doi: 10.1186/1471-2091-9-29.
76. Hesselberth, J.R., J.P. Miller, A. Golob, J.E. Stajich, G.A. Michaud, and S. Fields. 2006. Comparative analysis of Saccharomyces cerevisiae WW domains and their interacting proteins. Genome Biol. 7: R30, doi: 10.1186/gb-2006-7-4-r30.
77. Dudek, N. L., Dai, Y., & Muma, N. A. (2010). Neuroprotective effects of calmodulin peptide 76-121aa: Disruption of calmodulin binding to mutant Huntingtin. Brain Pathology, 20(1), 176–189. https://doi.org/10.1111/j.1750-3639.2008.00258.x

78. Bao, J., A.H. Sharp, M.V. Wagster, M. Becher, G. Schilling, C.A. Ross, V.L. Dawson, and T.M. Dawson. 1996. Expansion of polyglutamine repeat in huntingtin leads to abnormal protein interaction with calmodulin. Proc. Natl. Acad. Sci. USA. 93: 5037-5042, doi: 10.1073/pnas.93.10.5037.
79. Zainelli, G.M., C.A. Ross, J.C. Troncoso, J.K. Fitzgerald, and N.A. Muma. 2004. Calmodulin regulates transglutaminase 2 cross-linking of huntingtin. J. Neurosci. 24: 1954-1961, doi: 10.1523/JNEUROSCI.4424-03.2004.
80. Keryer, G., J.R. Pineda, G. Liot, J. Kim, P. Dietrich, C. Benstaali, K. Smith, F.P. Cordelières, N. Spassky, R.J. Ferrante, I. Dragatsis, and F. Saudou. 2011. Ciliogenesis is regulated by huntingtin-HAP1-PCM1 pathway and is altered in Huntington disease. J. Clin. Invest. 122: 4372-4382, doi: 10.1172/JCI57552.
81. Karam, A., L. Tebbe, C. Weber, N. Messaddeq, L. Morlé, P. Kessler, U. Wolfru, and Y. Trottier. 2015. A novel function of Huntingtin in the cilium and retinal ciliopathy in Huntington's disease mice. Neurobiol. Dis. 80: 15-28, doi: 10.1016/j.nbd.2015.05.008.
82. Burgoyne, R. D., & Haynes, L. P. (2015). Sense and specificity in neuronal calcium signalling. Biochimica Et Biophysica Acta (BBA) - Molecular Cell Research, 1853(9), 1921–1932. https://doi.org/10.1016/j.bbamcr.2014.10.029
83. Chagot B, Chazin WJ. Solution NMR structure of Apo-calmodulin in complex with the IQ motif of human cardiac sodium channel NaV1.5. J Mol Biol. 2011 Feb 11;406(1):106-19. doi: 10.1016/j.jmb.2010.11.046.
84. Damo SM, Feldkamp MD, Chagot B, Chazin WJ. NMR studies of the interaction of calmodulin with IQ motif peptides. Methods Mol Biol. 2013;963:173-86. doi: 10.1007/978-1-62703-230-811.
85. Xu, D., & Zhang, Y. (2012). Ab initio protein structure assembly using continuous structure fragments and optimized knowledge-based force field. Proteins: Structure, Function, and Bioinformatics, 80(7), 1715–1735. https://doi.org/10.1002/prot.24065
86. Xu, D. and Y. Zhang. 2013. Toward optimal fragment generations for ab initio protein structure assembly. Proteins. 81:229-239, doi: 10.1002/prot.24179.
87. Šali, A., & Blundell, T. L. (1993). Comparative protein modelling by satisfaction of spatial restraints. Journal of Molecular Biology, 234(3), 779–815. https://doi.org/10.1006/jmbi.1993.1626
88. Schneidman-Duhovny, D., Hammel, M., Tainer, J. A., & Sali, A. (2013). Accurate SAXS profile computation and its assessment by contrast variation experiments. Biophysical Journal, 105(4), 962–974. https://doi.org/10.1016/j.bpj.2013.07.020
89. Schneidman-Duhovny, D., Hammel, M., Tainer, J. A., & Sali, A. (2016). Foxs, FoXSDock and multifoxs: Single-state and multi-state structural modeling of proteins and their complexes based on SAXS Profiles. Nucleic Acids Research, 44(W1). https://doi.org/10.1093/nar/gkw389
90. Grant, T. D. (2018). Ab initio electron density determination directly from solution scattering data. Nature Methods, 15(3), 191–193. https://doi.org/10.1038/nmeth.4581
91. Xia, Y., Rossi, P., Subrahmanian, M. V., Huang, C., Saleh, T., Olivieri, C., Kalodimos, C. G., & Veglia, G. (2017). Enhancing the sensitivity of multidimensional NMR experiments by using triply-compensated π pulses. Journal of Biomolecular NMR, 69(4), 237–243. https://doi.org/10.1007/s10858-017-0153-2

92. Díaz Casas, A., G. Casanova Sepúlveda, O. Sánchez Negrón, A.P. Caro Muñiz, S.R. Malavé Ramos, A.P. Cebollero López, and B. Pastrana-Rios. 2018. Molecular biophysical characterization of the third FF domain of Homo sapiens Prp40 homolog A. J. Mol. Struct. 1167. 174-179, doi: 10.1016/j.molstruc.2018.04.059.
93. Bonet, R., L. Ruiz, E. Aragón, P. Martín-Malpartida, and M.J. Macias. 2009. NMR structural studies on human p190-A RhoGAPFF1 revealed that domain phosphorylation by the PDGF-Receptor α requires its previous unfolding. J. Mol. Biol. 389: 230-237, doi: 10.1016/j.jmb.2009.04.035.
94. Sánchez-Hernández, N., L. Ruiz, M. Sánchez-Álvarez, M. Montes, M.J. Macias, C. Hernández-Munain, and C. Suñé. 2012. The FF4 and FF5 domains of transcription elongation regulator 1 (TCERG1) target proteins to the periphery of speckles. J. Biol. Chem. 287: 17789-17800, doi: 10.1074/jbc.M111.304782.
95. Bonet, R., Ruiz, L., Morales, B., & Macias, M. J. (2009). Solution structure of the fourth ff domain of yeast PRP40 splicing factor. Proteins: Structure, Function, and Bioinformatics, 77(4), 1000–1003. https://doi.org/10.1002/prot.22547
96. Allen, M., A. Friedler, O. Schon, and M. Bycroft. 2002. The structure of a FF domain from human HYPA/FBP11. J. Mol. Biol. 323: 411-416, doi: 10.1016/S0022-2836(02)00968-3.
97. Bonet, R., X. Ramirez-Espain, and M.J. Macias. 2008. Solution structure of the yeast URN1 splicing factor FF domain: Comparative analysis of charge distributions in FF domain structures-FFs and SURPs, two domains with a similar fold. Proteins. 73: 1001-1009, doi: 10.1002/prot.22127.
98. Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., Thompson, J. D., & Higgins, D. G. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Molecular Systems Biology, 7(1), 539. https://doi.org/10.1038/msb.2011.75
99. Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., & Ferrin, T. E. (2004). UCSF chimera?a visualization system for exploratory research and analysis. Journal of Computational Chemistry, 25(13), 1605–1612. https://doi.org/10.1002/jcc.20084
100. Topolska-Woś, A.M., N. Sugitani, J.J. Cordoba, K.V. Le Meur, R.A. Le Meur, H.S. Kim, J-E. Yeo, D. Rosenberg, M. Hammel, O.D. Schärer, and W.J. Chazin. 2020. A key interaction with RPA orients XPA in NER complexes. Nucleic Acids Res. 48: 2173-2188, doi: 10.1093/nar/gkz1231.
101. Johnson, C.N., F. Potet, M.K. Thompson, B.M. Kroncke, A.M. Glazer, M.W. Voehler, B.J. Knollmann, A.L. George Jr., and W.J. Chazin. 2018. A mechanism of calmodulin modulation of the human cardiac sodium channel. Structure. 26: 683-694, doi: 10.1016/j.str.2018.03.005.
102. Chow, J. Y. H., Jeffries, C. M., Kwan, A. H., Guss, J. M., & Trewhella, J. (2010). Calmodulin disrupts the structure of the HIV-1 MA protein. Journal of Molecular Biology, 400(4), 702–714. https://doi.org/10.1016/j.jmb.2010.05.022
103. Persechini A, Kretsinger RH. The central helix of calmodulin functions as a flexible tether. J Biol Chem. 1988 Sep 5;263(25):12175-8.
104. Matsushima, N., N. Hayashi, Y. Jinbo, and Y. Izumi. 2000. Ca2+-bound calmodulin forms a compact globular structure on binding four trifluoperazine molecules in solution. Biochem. J. 347. 211-215, doi: 10.1042/bj3470211.

105. Mazzorana, M. and T. L-M. Sørensen. 2019. Calcium-induced protein folding in calumenin and calmodulin. Methods Mol. Biol. 1929: 517-537, doi: 10.1007/978-1-4939-9030-6_32.
106. Chattopadhyaya, R., W.E. Meador, A.R. Means, and F.A. Quiocho. 1992. Calmodulin structure refined at 1.7 Å resolution. J. Mol. Biol. 228: 1177-1192, doi: 10.1016/0022-2836(92)90324-d.
107. Chow, J. Y. H., Jeffries, C. M., Kwan, A. H., Guss, J. M., & Trewhella, J. (2010). Calmodulin disrupts the structure of the HIV-1 MA protein. Journal of Molecular Biology, 400(4), 702–714. https://doi.org/10.1016/j.jmb.2010.05.022
108. Köster, S., T. Pavkov-Keller, W. Kühlbrandt, and Ö. Yildiz. 2011. Structure of human Na+/H+ exchanger 1 NHE1 regulatory region in complex with calmodulin and Ca2+. J. Biol. Chem. 286: 40954-40961, doi: 10.1074/jbc.M111.286906.
109. Ikura M, Clore GM, Gronenborn AM, Zhu G, Klee CB, Bax A. Solution structure of a calmodulin-target peptide complex by multidimensional NMR. Science. 1992 May 1;256(5057):632-8.
110. Shah VN, Wingo TL, Weiss KL, Williams CK, Balser JR, Chazin WJ. Calcium-dependent regulation of the voltage-gated sodium channel hH1: intrinsic and extrinsic sensors use a common molecular switch. Proc Natl Acad Sci U S A. 2006 Mar 7;103(10):3592-7. doi: 10.1073/pnas.0507397103
111. Qi, L. S., Larson, M. H., Gilbert, L. A., Doudna, J. A., Weissman, J. S., Arkin, A. P., & Lim, W. A. (2013). Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. Cell, 152(5), 1173–1183. https://doi.org/10.1016/j.cell.2013.02.022
112. Dunn, K. W., Kamocka, M. M., & McDonald, J. H. (2011). A practical guide to evaluating colocalization in Biological Microscopy. American Journal of Physiology-Cell Physiology, 300(4). https://doi.org/10.1152/ajpcell.00462.2010
113. Walker, J. M. (2005). The proteomics protocols handbook. Humana Press.