Investigating the Dynamics of LARP1 DM15: A Stalwart of Ribosome Biogenesis Regulation

by

Kevin C. Cassidy

B.S in Biology, The Pennsylvania State University, 2016

Submitted to the Graduate Faculty of the

Dietrich School of Arts and Sciences in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2021

UNIVERSITY OF PITTSBURGH

DIETRICH SCHOOL OF ARTS AND SCIENCES

This dissertation was presented

by

Kevin C. Cassidy

It was defended on

June 22, 2021

and approved by

Karen M. Arndt, PhD, Professor

Donald Hamelberg, PhD, Professor

Graham F. Hatfull, PhD, Professor

Justin Kitzes, PhD, Assistant Professor

Thesis Advisor/Dissertation Director: Andrea J. Berman, PhD, Associate Professor

Thesis Advisor/Dissertation Director: Jacob D. Durrant, PhD, Assistant Professor

Investigating the Dynamics of LARP1 DM15: A Stalwart of Ribosome Biogenesis Regulation

Kevin C. Cassidy, PhD

University of Pittsburgh, 2021

Eukaryotic cells regulate the synthesis of proteins based on environmental conditions such as nutrient abundance or stress. Ribosome biogenesis is a key node in this regulation. La-related protein 1 (LARP1) has recently been identified as a downstream target of mTORC1 that regulates the translation of terminal oligopyrimidine (TOP) mRNAs, which encode all ribosomal proteins. LARP1 phosphorylation by mTORC1 modulates the association of LARP1 with TOP mRNA. Previous studies support both a repressive and stimulatory role for LARP1 in TOP mRNA translation, and its exact role is still unknown; LARP1 has also been shown to stabilize TOP mRNAs. Fully characterizing the role of LARP1 in translation regulation is vitally important because it is upregulated in several carcinomas and is an attractive pharmacological target. The best strategy to pharmacologically target LARP1 depends on the knowledge of how it affects TOP translation, and consequently, how best to modulate its role in that regulation.

Here we used molecular dynamics simulations to investigate the dynamics of the LARP1 DM15 region and provide insight into the dynamics that govern its role in translation regulation. We found that the open or closed state of the cap-binding pocket is correlated with the secondary structure of an adjacent stretch of residues. Additional LARP1 DM15 pockets were identified that have putative allosteric relationships with either the cap- or +1- binding pockets. Co-evolution analysis reveals a potential autoregulatory mechanism via an interdomain interaction with two of these pockets. A breast cancer-associated LARP1 mutation was introduced into one of the pockets

in silico, and molecular dynamics simulations revealed the mutation alters the dynamics of both the cap- and +1-binding pockets, potentially altering the ability to bind mRNAs. Lastly, a combination of *in silico* and *in vitro* techniques were used to identify compounds that putatively bind LARP1 DM15. Future experiments could utilize the compound hits to investigate the role of LARP1 in TOP mRNA translation, and optimize them for therapeutic use.

Table of Contents

List of Tables

List of Figures

Preface

I would like to thank my family, friends, and coworkers for their support during this journey.

1.0 Introduction

All life hinges upon the orchestration of dynamic processes that use and/or provide energy. For eukaryotic organisms, one such essential process is the creation of the ribosome[1]. The ribosome is the key part of the protein producing machinery at the heart of translation, the most downstream process in the central dogma of molecular biology: DNA↔RNA→Protein. The resulting proteins then help tip the energetic balance in favor of sustaining life; examples include transporting essential ions into cells, maintaining cell structure, and accelerating chemical reactions. Although proteins are essential to regulating the use and production of energy, ribosome biogenesis itself is an energetically costly process[1]. This means that the cell must carefully regulate ribosome biogenesis and sense when increased protein production is needed[1-3].

1.1 Eukaryotic Translation

Eukaryotic translation is broken into three main steps: initiation, elongation, and termination[4-8]. Translation initiation is regulated at several points and involves the recruitment and assembly of the 40S and 60S ribosomal subunits on the mRNA[4, 5, 7, 8]. The appropriate amino acids are added to the polypeptide being synthesized from the mRNA in the elongation step; during this step several ribosomes may be associated with same mRNA forming a polysome and thereby increasing the amount of the encoded protein produced[6-9]. The synthesized protein is then released in the termination step[6-8]. Regulating the initiation step is a key process that allows the cell to

respond to environmental cues, such as nutrient abundance which indicates whether the cell should produce more protein or not[3-5, 10, 11].

1.1.1 The 5' Cap and Poly(A) Tail

A crucial cis-acting translation regulatory factor is the 7-methylguanosine cap at the 5' end of mRNAs[12]. This 5' cap also stabilizes the mRNA by helping to prevent degradation[13]. The 5'cap precedes the 5' untranslated region (5'UTR), which itself precedes the coding sequence. Capping takes place co-transcriptionally after the first twenty-five to thirty nucleotides have been added[7, 8, 14]. The incorporation of the 5' cap is the first modification made to RNA transcribed by RNA polymerase II. First, the γ-phosphate of the most 5' nucleotide is removed by RNA triphosphatase, resulting in a 5' diphosphate. Second, a GMP group from GTP is transferred to the 5' diphosphate by RNA guanylyltransferase. Lastly, guanine-N7 methyltransferase modifies the N7 amine of the guanine cap via the addition of a methyl group[7, 8, 12, 14]. This results in the 5' 7-methylguanosine cap which directly precedes the +1-nucleotide position of the RNA, also known as cap 0 (Figure 1).
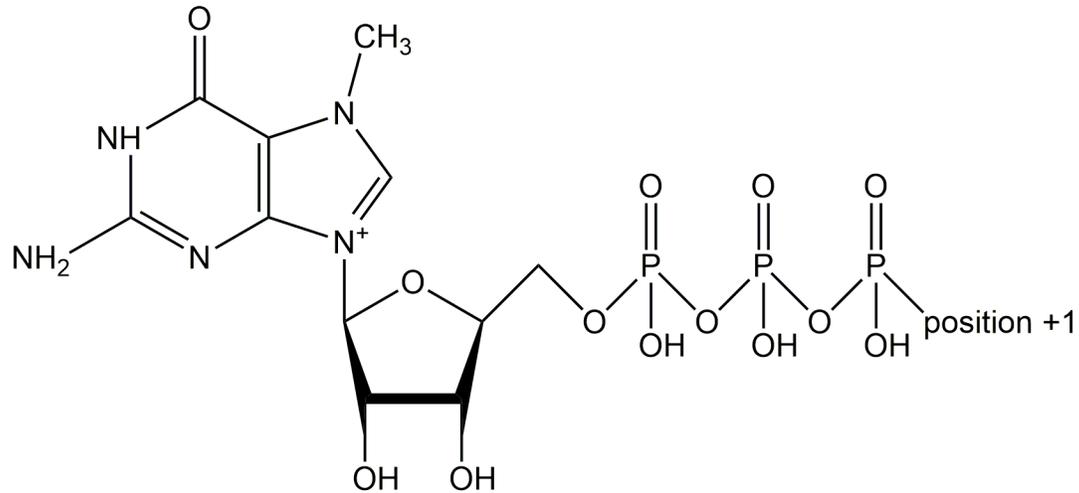
Figure 1. 2D structure of the 7-methylguanosine cap which preceeds the +1 nucleotide[12, 14]

The 3' poly(A) tail, an example of another cis-acting translation regulatory feature, is usually two hundred to two hundred and fifty nucleotides long[7, 8, 15, 16]. The poly(A) tail is preceded by the 3'untranslated region, which is preceded by the coding sequence. There are three main sequence components that signal for polyadenylation in mammals: 1) a conserved hexanucleotide sequence (AAUAAA) exists ten to thirty nucleotides upstream of the cleavage site; 2) a less conserved and variable U or UG containing sequence is downstream of the cleavage site; 3) lastly, there is the cleavage site[7, 8, 15, 16]. Additionally, other sequence elements can affect the polyadenylation, although they are not discussed here[15, 16]. Protein machinery is also required for this 3' processing and addition of the poly(A) tail modification. The RNA is cleaved by the cleavage and polyadenylation specificity factor (CPSF) complex endonuclease; CPSF recognizes the AAUAAA motif. In the second step, polyadenylate polymerase (PAP) catalyzes untemplated polyadenylation[7, 8, 15-17].

1.1.2 5' Cap-Dependent Translation Initiation

The 5' cap and poly(A) tail recruit crucial translation initiation factors[4, 5, 7, 8]. Eukaryotic translation initiation factor 4E (eIF4E) recognizes and binds the 5' cap, followed by eukaryotic translation initiation factor 4G (eIF4G) binding to eIF4E[4, 5, 7, 8, 18]. eIF4G is a scaffolding protein, and the RNA helicase eukaryotic translation initiation factor 4A (eIF4A) binds to it before unwinding 5' proximal regions of the mRNA. eIF4E, eIF4G, and eIF4A make up the eukaryotic translation initiation factor 4F (eIF4F) complex[4, 5, 7, 8, 19]. Poly(A)-binding protein (PABP) associates with the poly(A) tail of the mRNA, and subsequently binds to eIF4G; this aids in circularizing the mRNA (Figure 2) [4, 5, 8, 18-20]. The circularization of mRNA may aid in the translation efficiency of the mRNA by positioning the sites of 80S ribosome formation and translation termination closer together; ribosomes may be more quickly and efficiently recycled back onto the translated mRNA[8, 18, 21].
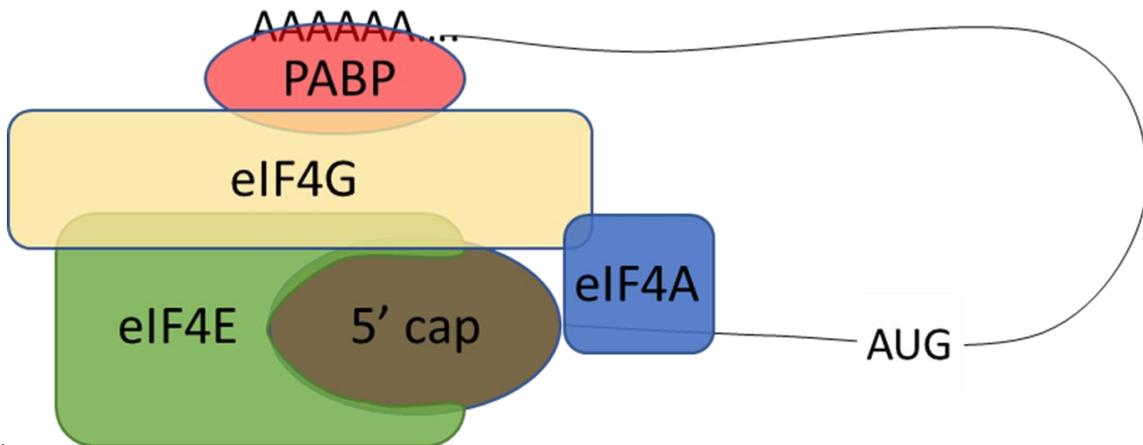
Figure 2. Model of the eIF4F complex and PABP assembled on mRNA (not to scale). The eIF4F complex is made up of eIF4E, eIF4G, and eIF4A. eIF4E binds the 5' cap of mRNA, and the scaffolding protien eIF4G binds to eIF4E. eIF4G binds to the RNA helicase eIF4A, as well as PABP (which binds the poly(A) tail). The

binding of eIF4G to PABP circularizes the mRNA, and eIF4A unwinds the 5' proximal region of the mRNA[4, 5, 7, 8].

The 43S pre-initiation complex (PIC) is then recruited to the mRNA via interactions with eIF4F. Initiator methionyl-tRNA (Met-tRNA$_i$), GTP-bound eukaryotic initiation factor 2, and the 40S ribosomal subunit comprise the 43S PIC[1, 4-8, 22, 23]. The 43S PIC scans the mRNA for the start (AUG) codon[7, 8, 22, 24]. The 60S ribosomal subunit is then recruited, following the release of several factors within the PIC. This results in the formation of the 80S ribosome[1, 4, 5, 7, 8, 22].

The 80S ribosome is formed and the Met-tRNA$_i$ is paired with the AUG codon in the peptidyl site (P-site), now the elongation step can commence[4, 5, 7, 8, 22]. Aminoacyl-tRNA is loaded into the aminoacyl site (A-site), and the nascent peptide is elongated[6, 8]. The tRNAs are cycled through the A-, P-, and exit (E-) sites as the polypeptide chain grows[6, 8]. The tRNAs are released at the E-site[6, 8].

Termination takes place once a stop codon (UAG, UAA, UGA) is recognized in the A-site of the ribosome[6, 8, 25]. Eukaryotic translation termination factors 1 and 3 (eRF1, eRF3) facilitate the hydrolytic release of the nascent peptide from the P-site[6, 8, 25]. The 80S ribosome is then recycled, starting with the separation of the 60S and 40S subunits mediated by ABCE1[6, 25]. The deacylated tRNA and mRNA can then be released from the remaining 40S subunit[6, 25].

1.1.3 The Ribosome

The 80S ribosome consists of both ribosomal proteins and ribosomal RNA (rRNA)[1]. There are forty-seven ribosomal proteins and three rRNAs (5S, 5.8S, and 25S) in the 60S subunit[1]. The 40S subunit is made up of thirty-three ribosomal proteins and the 18S rRNA. Assembly of the ribosome is an energetically intensive process and there about 200 ribosomal assembly factors.

This process is regulated and tuned depending on a cell's needs dependent on environmental cues. Studies support that upregulation of ribosome assembly is crucial for cancer cells which divide at a high rate, and factors that directly or indirectly regulate ribosome biogenesis have been targeted in cancer treatment development[26, 27].

1.2 5'Terminal Oligopyrimidine Motif mRNAs

The transcripts that encode all seventy-nine ribosomal proteins have a cis-regulatory element in their sequence termed the 5'terminal oligopyrimidine (5'TOP) motif[2, 28, 29]. This motif is characterized by an invariant +1C just downstream of the 5'cap and is followed by a series of about 4-15 pyrimidines[28, 29]. This can then be followed by a G-rich region. In addition to ribosomal proteins, several translation factors contain the 5'TOP motif such as eukaryotic translation initiation factor 3 subunit A (eIF3A), and eukaryotic translation elongation factor 1A (eEF1A)[28]. The 5'TOP motif is conserved among all vertebrates, as well as in *Drosophila melanogaster* ribosomal protein transcripts[28]. The invariant +1C is somewhat unique to 5'TOP mRNAs because eukaryotic transcripts generally start with a purine (A, G). This +1C is also critical to its cis-regulatory function, as a switch to another nucleotide results in loss of 5'TOP regulatory influence[28].

1.2.1 TOP mRNA Translation Regulation

The 5'TOP motif has evolved to respond to signals in the cell to regulate the translation of a translation machinery class of mRNA[2, 28, 30]. In oxygen rich and nutrient abundant conditions,

translation of TOP mRNAs is increased[2, 28, 30]. One source of TOP mRNA translation regulation is tied to the cell cycle[28]. TOP mRNA translational repression occurs when the cell is in a quiescent state during the G0 or resting phase. When the cell is not dividing, the need for translational machinery and ribosomes decreases substantially. Conversely, TOP mRNA translation is upregulated by growth signals, such as amino acid abundance; the mammalian target of rapamycin complex 1 (mTORC1) pathway plays a central role in this growth signal dependent regulation of TOP mRNAs[10, 28, 30-40].

1.3 Mammalian Target of Rapamycin Complex 1 (mTORC1) Signaling

Named after the drug it is targeted by, the mammalian target of rapamycin (mTOR) is at the center of the cellular response to nutrients and growth conditions[10, 31-34, 36, 37, 41-43]. mTOR is a serine/threonine protein kinase and is the catalytic subunit of both mTORC1 and mTORC2. mTORC1 in particular regulates several aspects of cell growth and proliferation such as protein synthesis, autophagy, and lysosome biogenesis[10, 28, 30-34, 36, 37, 42, 43]. The translation of TOP mRNAs is regulated by mTORC1 and is activated in the presence of nutrient abundance and oxygen, while repressed during stress and a lack of available nutrients[2, 28, 30]. mTORC1 acts through the coordination of its three essential components: mTOR (the serine/threonine kinase and catalytic subunit), regulatory protein associated with mTORC (RAPTOR, the scaffolding protein and substrate recognition module), and mammalian lethal with Sec13 protein 8 (mLST8, required for the activation of mTOR)[10, 31-34, 36, 43]. In addition, there are two subunits with inhibitory roles: DEP domain containing mTOR interacting protein (DEPTOR), and proline-rich Akt substrate of 40 kDa (PRAS40)[10, 31-34, 36, 41].

7

1.3.1 mTORC1 Signal Pathway
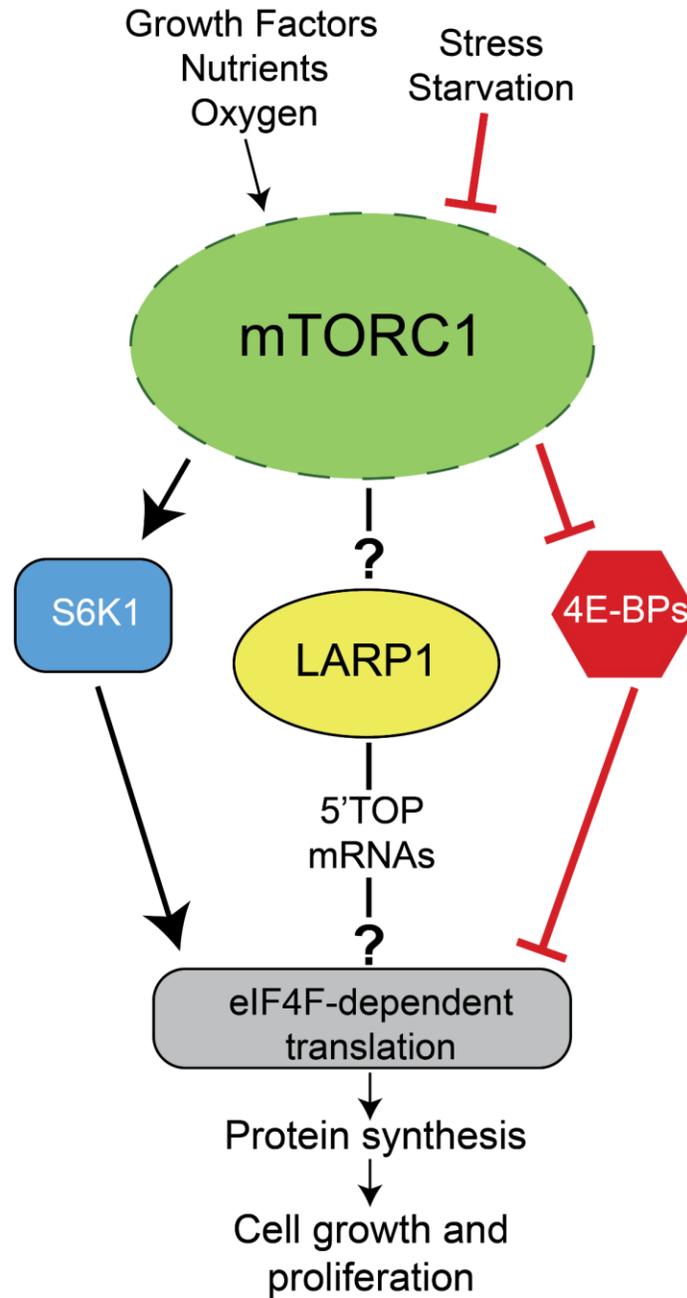
Figure 3. mTORC1 regulates global and TOP mRNA translation. La-related protein 1 (LARP1), ribosomal S6 kinase 1 (S6K1), and the eIF4E-binding proteins (4E-BPs) are three main downstream targets of mTORC1. mTORC1 phosphorylates all three proteins under nutrient and growth factor abundant conditions. Phosphorylated S6K1 activates several downstream factors that promote translation initation.

Phosphorylated 4E-BP dissociates from eIF4E, allowing for translation initiation. The effect of LARP1 phosphorylation by mTORC1 is not fully characterized, and the overall role LARP1 plays in TOP mRNA translation is not known.

Several factors upstream of mTORC1 react to environmental cues, and subsequently regulate mTORC1 activity[10, 32, 35-37, 39, 41]. mTORC1 is recruited to lysosomes when nutrients are abundant; there the small GTPase Ras homolog enriched in brain (Rheb) activates mTORC1 by binding and inducing an allosteric structural alteration[10, 43]. The tuberous sclerosis complex (TSC) suppresses mTORC1 activity via inhibiting Rheb[10, 31-34, 43, 44]. TSC itself is inhibited via phosphorylation by protein kinase B (PKB/Akt) when growth factors are present[10, 44, 45]. PKB/Akt also phosphorylates PRAS40 and prevents its inhibition of mTORC1[10, 32, 43]. Upstream of PKB/Akt is the lipid kinase, class I phosphoinositide 3-kinase (PI3K), a crucial factor in cell growth, metabolism, motility and proliferation[10, 32, 39].

Three important downstream targets of mTORC1 are the eIF4E-binding proteins (4E-BPs), ribosomal S6 kinase 1 (S6K1), and La-related protein 1 (LARP1) (Figure 3)[30, 35, 46-48]. In environmental conditions where nutrients and growth factors are prevalent, mTORC1 phosphorylates the 4E-BPs causing them to dissociate from eIF4E; eIF4E is then free to bind the 5'cap and initiate translation[10, 32-34, 48]. Alternatively, environmental stress causes inactivation of mTORC1 leaving 4E-BPs bound to eIF4E, inhibiting translation initiation. Phosphorylated S6K1 activates several downstream factors that then promote translation[10, 31, 32, 47]. mTORC1 is also known to specifically regulate the translation of TOP mRNAs, and recently it has been shown to be through an interaction with La-related protein 1 (LARP1)[2, 28, 49]. However, the exact effect mTORC1 imparts on LARP1 and the overall effect LARP1 has on TOP mRNA translation is not known.

1.3.2 mTORC1 in Cancer

Abnormal activation of mTORC1 has been observed in several cancers. Gastric, lung, breast, and renal cancers are just some of the mTORC1 associated cancers[10, 31-33, 50-53]. Cancer contributing activation of mTORC1 can occur by mutations in upstream factors, or within mTORC1 itself[10, 31-33, 50, 54]. Mutations within or upregulation of PKB/Akt and growth factor receptors (such as insulin growth factor receptor or epidermal growth factor receptor) can lead to abnormal activation of mTORC1[54-56]. Additionally, thirty-three mutations within mTOR itself were found to cause hyperactivity and discovered within tumors of various cancers[56].

These types of genetic alterations have been identified within specific cancer types. In an Eastern Chinese population, mTORC1 polymorphisms were found to increase the risk of gastric cancer[51]. Renal cancer has associated genetic alterations within the PKB/Akt signaling pathway[57]. Mutations within mTOR have not only been recorded within many cancer subtypes, but also can predict the sensitivity to the drug rapamycin[58].

1.3.3 mTORC1 Inhibitors

Rapamycin was first discovered as an immunosuppressive and antifungal agent; however, it was later found to have anti-tumor properties as well[10, 31-33, 41, 50, 53, 59]. The search for rapamycin's target was first pioneered in yeast (TOR1/2), and mTOR was later identified as the mammalian target[59]. Rapamycin first binds to FK506-binding protein (FKBP12) and the two then bind mTORC1 and cause partial inhibition[59, 60]. This partial inhibition limits the effectiveness of rapamycin as a cancer therapeutic and the predominant effect is on weak substrates of mTORC1[59]. Subsequently, derivatives of rapamycin were developed for cancer treatment and were the first

mTOR inhibitors used[31]. The rapamycin analogs (also called rapalogs) everolimus and temsirolimus have both been used in the treatment of renal cell carcinoma; everolimus is also used to treat advanced breast cancer[31, 50, 54, 56, 61-65]. However, treatment with rapamycin analogs alone has been shown to lead to only modest results likely due to incomplete inhibition[56].

Competitive inhibitors that target the ATP binding site of mTORC1 have also been developed as a way to pursue complete as opposed to partial inhibition; these ATP-competitive agents are also known as second generation mTOR inhibitors[31, 41, 54, 56, 63-66]. The mTOR inhibitor PP242 is able to effectively block mTORC1 phosphorylation of 4E-BP1, unlike the rapamycin derivatives[56, 63, 67]. These generation 2 mTOR inhibitors have entered clinical trials and have shown promising anti-cancer activity in several cancer types; although, some like vistusertib have had their use discontinued[31, 50, 53, 54, 56, 63, 64, 67, 68].

Inhibition of mTORC1 can actually augment the PI3K signaling pathway, leading to the development of dual PI3K/mTOR inhibitors[54, 56, 59, 63]. Several dual PI3K/mTOR inhibitors have shown potent anti-cancer activity and have been investigated in clinical trials.[56, 59, 63] Notable examples include voxtalisib, dactolisib, and gedatolisib, the latter of which inhibits tumor growth in colon, lung, breast, and glioma xenograft models[59].

An obstacle to using mTORC1 inhibitors is that they can cause several side effects[64, 65]. This is unsurprising, as mTORC1 has many downstream targets and affects a wide range of cellular processes. Rapamycin analogs can produce side effects that include, hematological, respiratory, renal, metabolic, and dermatological toxicities[64, 65]. Side effects can be severe and debilitating to the point of contributing to drug discontinuation[64, 65]. The second generation mTOR inhibitors could have more side effects due to their ability to more completely inhibit mTORC1[63]. Dual PI3K/mTOR inhibitors have been shown to cause serious side effects in clinical trials[59, 69]. Clinical

use of these drugs is also hampered by the development of drug resistance, such as that caused by mutations leading to a more catalytically active mTOR[31, 54, 56, 59, 63, 64]. These drawbacks and weaknesses of mTORC1 inhibitors could be overcome by targeting a less ubiquitous node downstream of mTORC1, such as LARP1. LARP1 could be targeted alone or in combination with mTORC1 for the desired anti-cancer effect.

1.4 La-related Protein 1

LARP1 is a member of the La-related protein (LARP) family; proteins in the LARP family contain a characteristic La-motif[70-72]. This family of proteins is well conserved throughout eukaryotes and are divided into seven LARP subfamilies: LARP1, LARP2, LARP3 or Genuine La, LARP4, LARP5, LARP6, and LARP7[8, 70, 72]. Each subfamily binds a specific subclass or subclasses of RNA. Genuine La or LARP3 is mostly localized to the nucleus and aids in the folding of premature tRNAs and binds the 3' UUU-OH[8, 70, 72, 73]. LARP4 binds poly(A) RNA and stimulates mRNA translation[8, 70, 72, 74]. LARP5 binds AU-rich sequences and also stimulates mRNA translation[8, 70, 72, 75]. LARP6 binds the 5' stem-loop of collagen mRNAs and regulates their localized translation[8, 70, 72, 76, 77]. LARP7 binds the 3'UUU-OH of the 7SK small non-coding RNA in animals and aids circularization and stabilization, and it functions to indirectly regulate transcription[8, 70, 72, 78]. There is very little known about LARP2, although it does share a similar C-terminal RNA binding domain with LARP1[70, 72]. Here I will focus on LARP1, which is primarily cytoplasmic, binds TOP mRNAs, and regulates their translation; I will be using the isoform 2 numbering, which is standard within the literature[2, 72, 79, 80].

LARP1 has recently been discovered as a TOP mRNA translation regulatory node downstream of mTORC1[2, 79-82]. Several studies have investigated the role of LARP1 in TOP mRNA translation, however there is still controversy surrounding its overall impact. Data support both a stimulatory and repressive LARP1 role in TOP mRNA translation[2, 81-83]. Separate studies have provided seemingly contradictory data; some show that LARP1 knockdown results in a decrease in polysome loading of TOP mRNAs (indicating a stimulatory role for LARP1), and in others LARP1 knockdown results in an increase in polysome loading of TOP mRNAs (indicating a repressive role for LARP1)[2, 49, 81-84]. Long term depletion of LARP1 has also been shown to decrease the abundance of proteins encoded by TOP mRNAs[82]. Another piece of the puzzle is the fact that LARP1 has also been shown to stabilize TOP mRNAs[2, 81, 83]. Allosteric regulation of LARP1 could potentially reconcile the seemingly contradictory data and LARP1 phosphorylation by mTORC1 has been shown to alter this role in translation regulation[81]. There are at least twenty-six mTORC1 phosphorylation sites within LARP1[35]. The exact role of LARP1 in TOP mRNA translation is also disease relevant due to its upregulation in several carcinomas[71, 85-89]. Any cancer treatment development that targets LARP1 needs to be informed by how best to tune its role in TOP mRNA translation regulation, in order to achieve the desired effect.

This role of LARP1 in translation regulation is defined by the functions of its RNA binding domains. Towards the N-terminus, the La-Module is comprised of the La motif, the PABP-interacting motif 2 (PAM2), and the RNA recognition motif like 5 (RRM-L5) (Figure 4)[2, 90]. At the C-terminus there is the DM15 region which is unique to LARP1 and LARP2 (Figure 4)[79, 80]. The La-Module has recently been shown to bind both TOP motifs, and poly(A) tails, and in fact can do so simultaneously; another function attributed to the La-Module is stabilization of mRNAs[90]. The current model suggests the ability of the La-Module to aid in the circularization of

13

TOP mRNAs, thereby increasing the efficiency of translation[90]. Here we focus on the DM15 region of LARP1, which is also able to stabilize TOP mRNA transcripts, and has been shown to be necessary and sufficient to convey a repressive role in TOP mRNA translation[81].

Figure 4. LARP1 domain organization. The N-terminal La-Module is comprised of two RNA binding motifs: the La motif (LAM, pink), and the RNA recognition motif like 5 (RRM-L5, orange). The PABP-interacting motif 2 (PAM2, red) is situated between the LAM and the RRM-L5. The DM15 region (yellow) is the C-terminal RNA binding domain, and is unique to LARP1 and LARP2[2].

1.4.1 LARP1/LARP2 Unique DM15 Region

The DM15 region of LARP1 stabilizes and regulates the translation of TOP mRNAs via recognizing the 5'cap and +1 C of the TOP motif[2, 79-82, 84, 91]. Several crystal structures of human LARP1 DM15 have been solved, including co-crystal structures that illustrate how it engages the 5' cap and +1C (Figure 5)[79, 80]. Key residues within the cap-binding pocket of LARP1 DM15 include Y922, E886, and Y883. Y922 and Y883 stack with the 5'cap, and E886 hydrogen bonds with the Watson-Crick face. The +1-binding pocket π- π stacks with the +1C via Y883 and F844; R847 hydrogen bonds with the Watson-Crick face of the +1C[79, 80].

LARP1 DM15 can outcompete eIF4E for capped TOP transcripts[79, 80]. This suggests a repressive role of LARP1 in translation regulation. However, the DM15 region also stabilizes transcripts and could be doing so by protecting from exonucleases[2, 81]. mTORC1 phosphorylation

14

of LARP1 has been shown to cause the DM15 region to disassociate with TOP mRNAs[2, 35, 81]. The DM15 region alone has been shown to be capable of repressing TOP mRNA translation, however it is insensitive to this mTORC1 regulation[81]. A construct containing an adjacent N-terminal region in addition to the DM15 region, was found to regain sensitivity to mTORC1 regulation[81]. Further, specific mTORC1 phosphorylation sites on LARP1 have been shown to increase and decrease the affinity of LARP1 DM15 for TOP mRNAs[35]. This allosteric regulation is likely the key to discovering the true role of LARP1 in TOP mRNA translation. It is possible that LARP1 DM15 binds to the cap and +1 of TOP mRNAs and represses translation in the short term, while also stabilizing the transcript. Then when growth factors are present mTORC1 phosphorylates LARP1 causing DM15 to dissociate from the TOP mRNAs, and allowing for cap-dependent translation initiation of the stabilized pool of TOP transcripts.
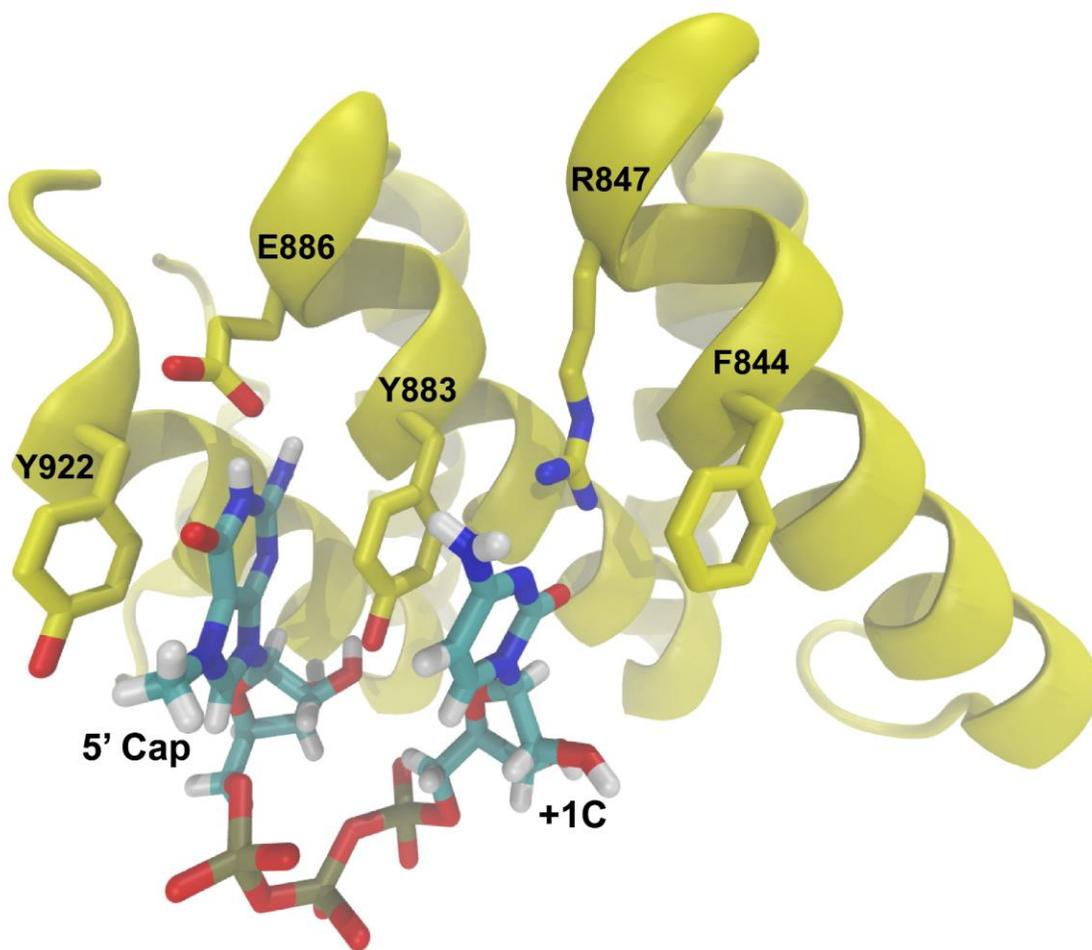
Figure 5. The cap- and +1-binding pockets of LARP1 DM15. LARP1 DM15 bound to m^{7}GpppC (PDB: 5V87:B)[79]. Key residues in the cap-binding pocket are Y922, E886, and Y883. Y922 and Y883 stack with the cap nucelobase, and E886 hydrogen bonds with the Watson-Crick face. Key residues in the +1-binding pocket are Y883, R847, and F844. Y883 and F844 π- π stack with the +1C nucleobase, and R847 hydrogen bonds with the Watson-Crick face.

1.4.2 LARP1 in Cancer and COVID-19

LARP1 regulates the translation of all ribosomal proteins, and so it is unsurprising that it is associated with several carcinomas. High levels of LARP1 have been found in hepatocellular,

16

prostate, breast, colorectal, lung, cervical, and ovarian cancer[71, 85-89, 92, 93]. The level of LARP1 expression has been found to correlate with poor prognosis and disease progression in several of these epithelial cancers[71, 85-87]. Substantially increased levels of LARP1 were found with increasing stage of invasive cervical cancer[87]. Other effects of LARP1 upregulation in cancer include positively affecting the epithelial-mesenchymal transition, tumorigenesis, cell migration, and invasion[71, 85-89, 93]. Additionally, LARP1 has been found to be the target of several miRNAs, and overexpression of miR-374a counteracted the positive effect LARP1 had on proliferation and metastasis in lung cancer; miRNAs mi-26a and b have been shown to negatively regulate LARP1 expression in prostate cancer[71, 94, 95]. Targeting LARP1 instead of mTORC1 pharmacologically could reduce side effects of treatment because LARP1 is downstream and is not as connected to other essential cell processes.

Recently, LARP1 was found to associate with the nucleocapsid protein of SARS-CoV-2, the virus which causes COVID-19[96]. This was followed up with work that showed LARP1 bound to both genomic and subgenomic SARS-CoV-2 RNAs; it is possible the nucleocapsid association is RNA dependent[97]. LARP1 was found to bind SARS-CoV-2 5'UTRs, and specifically the TOP-like motif sequence, which is in all viral subgenomic mRNAs[97]. This study also found that LARP1 represses SARS-CoV-2 replication in infected cells[97]. One opportunity for therapeutic development could be targeting LARP1 allosterically with a small molecule to enhance the interaction between LARP1 and SARS-CoV-2 RNAs thereby amplifying its repressive effect on viral replication.

1.5 Goals and Discoveries

Here I investigate the dynamics of LARP1 DM15, and provide insight into the mechanism behind its binding of TOP mRNAs and its role in translation regulation. We show through the use of molecular dynamics simulations, that the cap-binding pocket fluctuates between a "ligand-ready" and a closed state that would prevent TOP mRNA binding; we found this state was tied to the secondary structure of an adjacent bridge of residues[98]. Additionally, we characterized three other pockets, and found two of them were putatively allosterically tied to either the cap- or +1-binding pocket. Co-evolution analysis also revealed a potential interdomain interaction between the La-Module and DM15, which could transduce an allosteric signal to the cap- or +1-binding pockets. A LARP1 DM15 breast cancer-associated mutation was also introduced in silico, and a molecular dynamics simulation revealed the mutation altered the dynamics of the cap- and +1-binding pocket, potentially altering the ability to bind TOP mRNAs. Lastly, since LARP1 is an attractive pharmacological target in several cancers and in COVID-19, we utilized a combination of *in silico* and *in vitro* techniques to identify five compounds that putatively bind LARP1 DM15.

2.0 Capturing the Mechanism Underlying TOP mRNA Binding to LARP1

This chapter is presented as published. Detailed author contributions are as follows: J.D.D. performed the 4ZC4:B and 4ZC4:C simulations. K.C.C. and J.D.D. analyzed the simulation data. K.C.C. performed the site-directed mutagenesis for LARP1 DM15 F844Y/R847E, K921A, K924A, K924D, and K815A. K.C.C. and R.M.L. performed the LARP1 DM15 WT, F844YR847E , K921A, K924A, K924D, and K815A protein purifications. K.C.C. performed the protein melting temperature assays. R.M.L. performed the EMSAs, and crystallization. A.J.B. and R.M.L. performed the structure solution. K.C.C. created figures 1, 2, 3, 5, S1, S2, S3, S4, S5. R.M.L. created figure 4. K.C.C., R.M.L., A.J.B., and J.D.D. edited the figures. K.C.C., R.M.L., J.C.K., B.D.F., A.J.B., and J.D.D. contributed to writing. S.M. and S.R.D. provided reagents. J.C.K. created custom Python code for use in analyses.

Please note that all figures marked supplemental can be found in this chapter.

2.1 Summary

The RNA-binding protein La-related protein 1 (LARP1) plays a central role in ribosome biosynthesis. Its C-terminal DM15 region binds the 7-methylguanosine (m^7G) cap and 5' terminal oligopyrimidine (TOP) motif characteristic of transcripts encoding ribosomal proteins and

translation factors. Under the control of mammalian target of rapamycin complex 1 (mTORC1), LARP1 regulates translation of these transcripts. Characterizing the dynamics of DM15-TOP recognition is essential to understanding this fundamental biological process. We use molecular dynamics simulations, biophysical assays, and X-ray crystallography to reveal the mechanism of DM15 binding to TOP transcripts. Residues C-terminal to the m^7G-binding site play important roles in cap recognition. Furthermore, we show that the unusually static pocket that recognizes the +1 cytosine characteristic of TOP transcripts drives binding specificity. Finally, we demonstrate that the DM15 pockets involved in TOP-specific m^7GpppC-motif recognition are likely druggable. Collectively, these studies suggest unique opportunities for further pharmacological development.

2.2 Introduction

The ribosome and associated translation machinery lie at the heart of gene expression. Their biosynthesis is controlled at multiple levels (Fonseca et al., 2014)[30]. Under energy or nutritional stress, cells reduce processes associated with high energy expenditure and direct resources toward survival. Nutrient and energy deprivation are invariably accompanied by reduced ribosome production, an energetically demanding cellular process. In contrast, cells replete with nutrients and growth factors engage ribosome biosynthesis to ensure maximal growth and proliferation (Fonseca et al., 2018; Warner, 1999)[2, 3]. Cell propagation is fueled by amplified protein synthesis, which requires the *de novo* synthesis and assembly of new ribosomes (Pelletier et al., 2018)[11]. The signals that control protein synthesis, protein turnover (macroautophagy), and cell propagation converge on the mammalian target of rapamycin complex 1 (mTORC1) (Saxton and Sabatini, 2017)[10].

mTORC1 regulates protein synthesis by phosphorylating proteins that play key roles in translation control (Fonseca et al., 2014)[30]. For instance, mTORC1 stimulates cap-dependent translation initiation by phosphorylating eukaryotic initiation factor 4E (eIF4E)-binding proteins (4E-BPs) (Gingras et al., 2001)[99]. Phosphorylation causes 4E-BPs to dissociate from eIF4E (Brunn et al., 1997; Burnett et al., 1998)[100, 101], a 7-methylguanosine (m^7G) cap-binding protein (Filipowicz et al., 1976; Sonenberg et al., 1978)[102, 103]. When bound to unphosphorylated 4E-BPs, eIF4E cannot associate with the m^7G cap or eIF4G, as required to recruit the 40S ribosome subunit to the 5' UTR of the mRNA. 4E-BP- bound eIF4E prohibits cap-dependent translation initiation by inhibiting the formation of the eIF4F translation initiation complex (Sonenberg and Hinnebusch, 2009)[104]. mTORC1 stimulates cap-dependent translation initiation by relieving eIF4E inhibition via 4E-BP phosphorylation (Sonenberg and Hinne- busch, 2009)[104].

In addition to regulating 4E-BP phosphorylation and eIF4E binding, mTORC1 also regulates the phosphorylation of another recently identified substrate: La-related protein 1 (LARP1) (Hong et al., 2017; Kang et al., 2013)[34, 84]. mTORC1 phosphorylates LARP1, causing it to release the 5' UTRs of ribosomal protein transcripts (e.g., the RPS6 transcript, per our unpublished data). In doing so, mTORC1 orchestrates the translation of transcripts that encode ribosomal proteins and some translation-associated factors (Fonseca et al., 2015)[83]. These transcripts share a motif in their 5' UTRs termed the 5' terminal oligopyrimidine (5'TOP) motif that is required for the coordinated regulation of TOP mRNA translation (Fonseca et al., 2014; Meyuhas and Kahan, 2015)[28, 30]. The 5'TOP motif has an invariant C in the +1 position, followed by a tract of cytidines and uridines (Levy et al., 1991)[29].

LARP1 is a translational regulator of TOP transcripts that functions downstream of mTORC1 (Fonseca et al., 2015)[83]. LARP1 is hypothesized to repress TOP translation (Aoki et al.,

2013; Fonseca et al., 2015)[83, 105] by sequestering key mRNA transcripts from translation initiation factors (Lahr et al., 2017)[79] and anchoring them to stress granules (SGs) (Wilbertz et al., 2019)[106]. LARP1 appears to elicit its inhibitory effects on TOP mRNA translation via its C-terminal DM15 region (Philippe et al., 2017)[81]. Specifically, the C-terminal LARP1 DM15 region, comprised of three HEAT-like repeats, directly engages the 5'TOP sequence and the m^7G cap via a conserved and positively charged surface (Lahr et al., 2015, 2017)[79, 80]. By characterizing the molecular interactions between the DM15 region, the m^7G cap, and the 5'TOP motif, we provided a molecular mechanism linking LARP1 to TOP mRNA translation repression. Recent work has demonstrated that this repressive role is mTORC1 dependent (Philippe et al., 2017)[81].

The precise biological role of LARP1 is still debated. In addition to observations that LARP1 represses TOP mRNA translation (Fonseca et al., 2015; Lahr et al., 2017; Philippe et al., 2017)[79, 81, 83], several studies have also noted that LARP1 controls mRNA decay in both plants (Merret et al., 2013)[107] and animals (Aoki et al., 2013; Fonseca et al., 2015; Gentilella et al., 2017)[83, 105, 108]. In mammals, LARP1 binds and stabilizes TOP transcripts (Aoki et al., 2013; Fonseca et al., 2015; Gentilella et al., 2017)[83, 105, 108], possibly helping to sustain translation. Nutritional (Damgaard and Lykke-Andersen, 2011)[109] and oxidative stress (Wilbertz et al., 2019)[106] halt TOP mRNA translation and lead to an accumulation of TOP transcripts in SGs. Recent data (Wilbertz et al., 2019)[106] indicate that LARP1 plays a seminal role in anchoring TOP transcripts to SGs, thereby averting their rapid decay. Storage of TOP transcripts in SGs by LARP1 likely allows for engagement of translation without the requirement for *de novo* transcription. Consistent with an important role for LARP1 in maintaining mRNA stability, LARP1 depletion has been observed to reduce the levels of some TOP-encoded proteins (Tcherkezian et al., 2014)[82].

A number of studies suggest that, analogous to other downstream mTORC1 targets, LARP1 instigates cell proliferation (Tcherkezian et al., 2014)[82] and thus contributes to the onset and maintenance of several cancers (Hopkins et al., 2016; Mura et al., 2015; Selcuklu et al., 2012; Xie et al., 2013)[86-88, 93]. Given its potential role as both a repressor of translation and stabilizer of mRNA transcripts, some have argued that LARP1 functions as a molecular switch in cancer biology (Hong et al., 2017)[84]. Small-molecule modulators of the mTORC1-LARP1 axis thus have therapeutic potential. One approach to altering the activity of this axis is to inhibit mTORC1 itself, and potent and specific mTORC1 inhibitors exist (Feldman et al., 2009; Thoreen et al., 2009; Yu et al., 2010)[41, 50, 66]. But mTORC1 inhibition affects many critical pathways, leading to side effects that complicate the long-term therapeutic use of mTORC1 inhibitors (Creel, 2009)[65]. Targeting the downstream LARP1 protein could bypass these broad effects by limiting inhibition to the mTORC1-LARP1 axis. Herein, we present a molecular dynamics (MD) simulation of the LARP1 DM15 region. The flexibility of this cancer-associated protein, here revealed in atomic detail, suggests a molecular mechanism whereby LARP1 structural features mediate TOP mRNA binding. The simulation reveals interactions not observed in any crystal structure that may contribute to the binding mechanism. We present biochemical and biophysical evidence to support the proposed interactions. We also identify DM15 druggable hotspots, taking into account the pocket dynamics that the simulations reveal. Some of these hotspots are not evident in any crystal structure. They provide unique drug discovery opportunities to specifically treat a variety of LARP1-implicated cancers while minimizing off-target effects.

2.3 Results

The 4ZC4 crystallographic dimer captures the flexible region connecting the α7 and α8 helices (S923-L928) in two distinct conformations (Lahr et al., 2015)[80]. This α7–α8 bridge is a 3_{10} helix in chain B, but a disordered loop in chain C. Given that the bridge abuts the m⁷G- binding pocket (Lahr et al., 2017; Lahr et al., 2015)[79, 80], we hypothesize that its dynamics contribute to the mechanism of cap recognition. Though the DM15 region often crystalizes as a dimer, we have shown that it is a monomer in solution (Lahr et al., 2015)[80]. We therefore used two molecular dynamics (MD) simulations of LARP1 DM15 monomers, based on the 4ZC4:B and 4ZC4:C structures, to test this hypothesis. These monomeric simulations suggest a molecular mechanism that may arbitrate LARP1/5'TOP-motif binding *in vivo*.

2.3.1 Simulation Equilibration

We first verified that each DM15 simulation had properly equilibrated by calculating the C_α root-mean-square deviation (RMSD) between each simulation frame and the respective first frame. This analysis revealed that the RMSD values had not stabilized by the beginning of the production simulations (Figure 7, blue line). As the 4ZC4 protein construct had been truncated after D946 to facilitate crystallization (Lahr et al., 2015)[80], we recognized that the dynamics of the most C-terminal residues might be artefactual. Therefore, we recalculated the RMSD of all simulated C_α up to G939, ignoring subsequent C-terminal residues. The revised RMSD plot was much flatter (Figure 6A), confirming that the biologically relevant portion of the DM15 region had equilibrated. Out of an abundance of caution, we discarded the first 60 and 50 ns of the 4ZC4:B

and 4ZC4:C simulations, respectively (Figure 6A and Figure 7, dotted lines). Subsequent analyses focused only on the remaining portions of each simulation.

2.3.2 The α7–α8 Bridge is Highly Dynamic

To investigate the flexibility of the α7–α8 bridge, we calculated the root mean square fluctuation (RMSF) of the center of geometry of each residue. As expected, the C-terminal residues of the 4ZC4:B simulation (Figure 6B, in blue) had RMSF values that were substantially higher than those of the remaining construct (maximum: 8.26 Å), likely an artefact of protein construct truncation. For reference, the average RMSF of the remaining residues was only 1.09 Å.

Other, non-terminal residues with notably high RMSF values were positioned near the m^7G pocket, suggesting that this portion of the DM15 region is particularly flexible (Figure 6B inset, in blue). These residues had RMSF values greater than 2.0 Å, with the exception of K921 (1.73 Å) and S923 (1.47 Å). K921 and Y922 are located on the α7 helix, near the m^7G pocket (Figure 8A). S923, K924, A925, K926, N927, and L928 belong to the α7–α8 bridge (Figure 8A).

The 4ZC4:C RMSF values were generally similar to those of the 4ZC4:B simulation, except for the same key residues of the α7 helix and α7–α8 bridge (Figure 6B inset, in red).
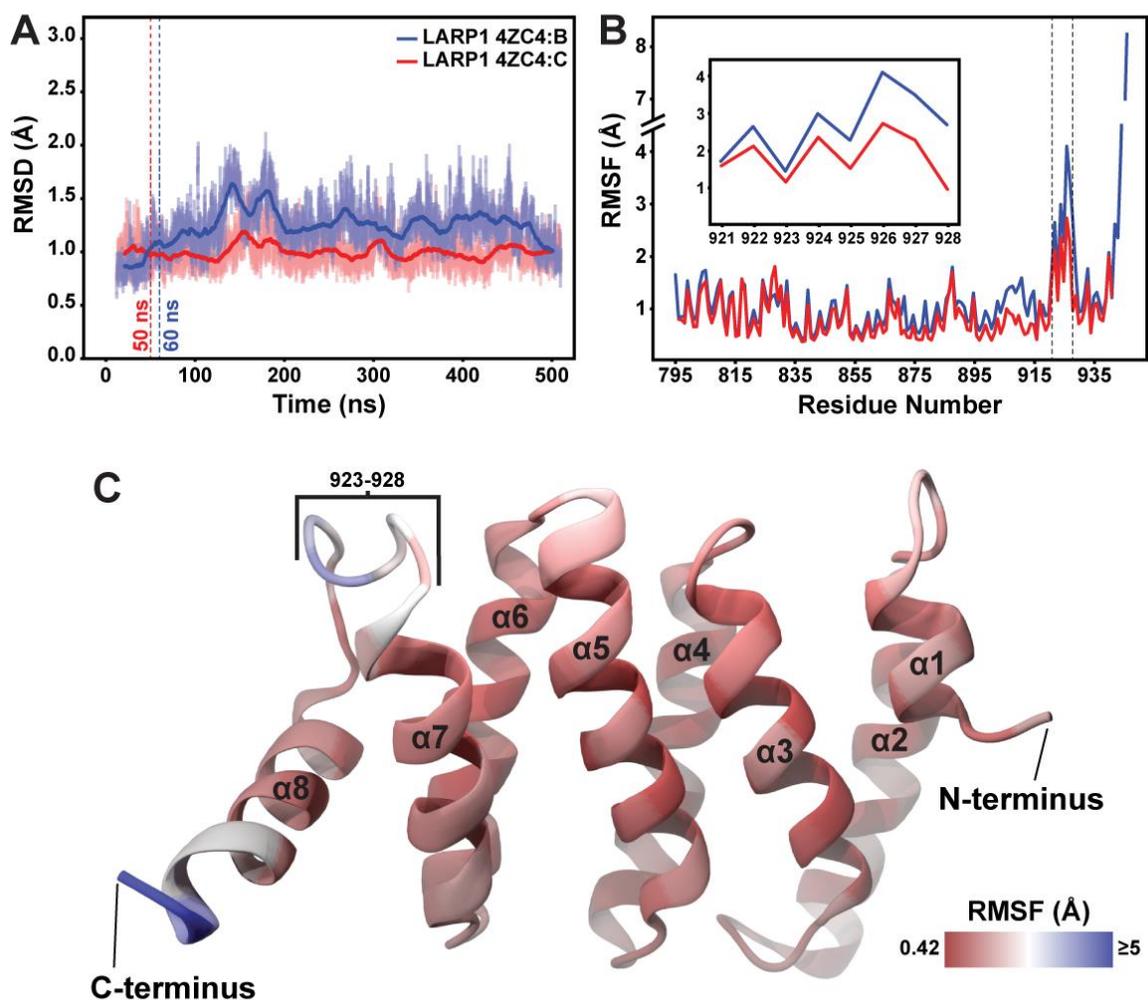
Figure 6. RMSD and RMSF analyses of the 4ZC4:B and 4ZC4:C simulations. The Cα RMSD between each frame and the first, excluding residues that were C-terminal to G939. Running averages are shown in darker colors, against the lighter-colored raw data. The frames preceding the dashed lines were not used in subsequent analyses. (B) The center-of-geometry, per-residue RMSF values of the 4ZC4:B and 4ZC4:C simulations, in blue and red respectively. The inset highlights the RMSF values of the α7–α8 bridge residues. (C) The RMSF values of the 4ZC4:B simulation projected onto a ribbon representation of the LARP1 DM15 region. Helices are labelled as in ref. (Lahr et al., 2015)[80]. See also Figure 7.
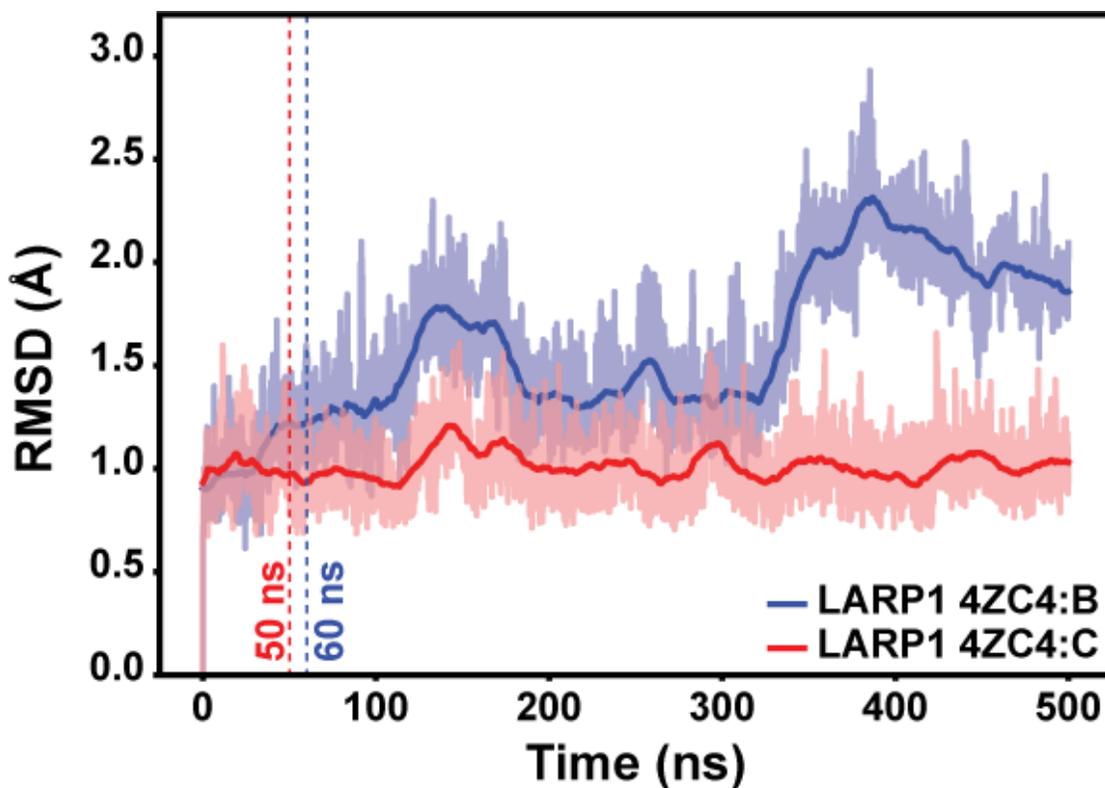
Figure 7. RMSD of the entire 4ZC4:B chain. The Cα RMSD between each frame and the first, when all residues of the 4ZC4:B chain were included. Running averages are shown in darker colors, against the lighter-colored raw data. Related to Figure 6.

Though these residues were notably flexible in the 4ZC4:C simulation as well, their motions were less dynamic than those observed in the 4ZC4:B simulation. Only Y922, K924, K926, and N927 had RMSF values greater than 2.0 Å. K926, the most flexible non-terminal residue, had an RMSF value of 4.12 Å and 2.76 Å in the 4ZC4:B and 4ZC4:C simulations, respectively. We chose to focus on the more dynamic 4ZC4:B simulation. The 4ZC4:C simulation is further described in the Supporting Information.

The instability of the α7–α8 bridge was surprising given that the 4ZC4:B bridge was helical in the starting conformation. Visual inspection of the trajectory revealed that the helix unfolds and refolds over the course of the simulation (Video S1). To quantify this event, we calculated the hydrogen-bond lengths between the backbone atoms of S923 and K926, K924 and N927, and

27

A925 and L928, per the *i* to *i*+3 pattern characteristic of a 3_{10} helix. A given α7–α8 bridge conformation was assumed to be helical if none of these three distances exceeded 4 Å. By this metric, the bridge was helical 11.7% of the time over the course of the production simulation.

2.3.3 m^7G-Pocket Dynamics

The secondary structure of the α7–α8 bridge determines the volume of the neighboring m^7G pocket. We used the POVME 2.0 algorithm (Durrant et al., 2014)[110] to calculate the m^7G-pocket volumes of 1,000 equally spaces frames taken from the 4ZC4:B production simulation. The pocket volumes ranged from 0 to 393.375 $Å^3$. We separated these volumes into two populations: those associated with a helical α7–α8 bridge, and those associated with a disordered-loop bridge (Figure 8B). A probability distribution of the pocket volumes associated with each bridge conformation shows that the volume tends to be larger when the α7–α8 bridge is helical (Figure 8C).

A transient interaction between K924 and E886 appears to influence m^7G-pocket volume by sequestering E886 away from the pocket. To investigate if this interaction correlated with the conformation of the α7–α8 bridge, we calculated the distances between the K924 sidechain nitrogen atom and the E886 side-chain carboxylate carbon atom over the course of the simulation. We divided the distances into bridge-helical and bridge-loop populations. The probability distribution of the K924-E886 distances shows that when the α7–α8 bridge is a disordered loop, these two residues participate in a tight electrostatic interaction. The distance between the K924 side-chain nitrogen atom and the E886 side-chain carboxyl carbon atom is less than 4 Å roughly 20% of the time (Figure 8D). This same interaction is present only 4.3% of the time when the bridge is a 3_{10} helix.

The conformation of the α7–α8 bridge also appears to influence an interaction between K921 and Y922 that stabilizes the m^{7}G pocket in a state preferential for ligand binding. We again divided the K921-sidechain-nitrogen/Y922-sidechain-oxygen distances into bridge-helical and bridge-loop populations. The probability distribution of the K921-Y922 distances shows that the K921–Y922 interaction-which prevents Y922 from occluding the m^{7}G pocket-is much more frequent when the α7–α8 bridge is a 3$_{10}$ helix than when it is not (28.2% vs. 12.7% of the time, respectively; Figure 8E).

The Supporting Information describes a possible bridge-dependent interaction between S923 and E886 that may also displace E886, just as the K924–E886 interaction does (Figure 9). The S923-E886 interaction has never been captured in any crystal structure. The Supporting Information also describes the interaction between R817 and E857, two residues that are distant from the m^{7}G pocket. This interaction is not correlated with the conformation of the α7–α8 bridge and so serves as a negative control (Figure 10).

2.3.4 Mutagenesis Studies Support the MD-Predicted Interactions
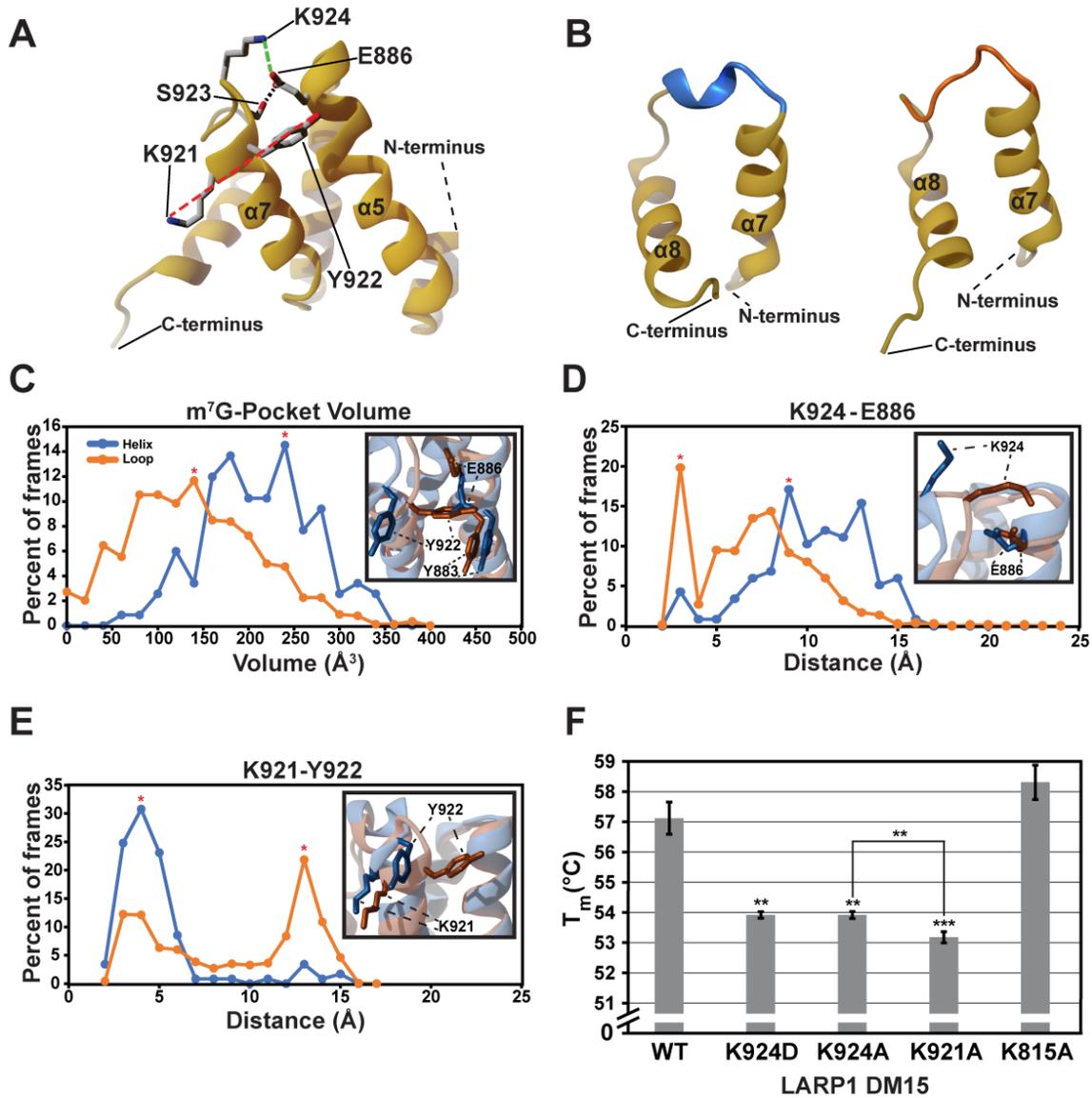
Figure 8. Specific residue interactions contribute to m7G-pocket dynamics. A frame from the MD simulation that illustrates key interactions. (B) Representative conformations of the α7-α8-bridge helix and loop microstates colored by secondary structure (blue: helical; orange: loop), extracted from the simulation. (C) The probability distributions of the m7G-pocket volumes (helical bridge, n = 117; disordered bridge, n = 883). The two conformational populations are statistically different (Levene's Test, p = 0.021; t-test, p = 9.6 × 10−25). (D) The probability distributions of the distance between the K924 amino and E886 carboxyl groups.

The two populations are statistically different (Levene's Test, p = 0.65; t-test, p = 2.0 × 10−25). (E) The

probability distributions of the distance between the K921 amino and Y922 hydroxyl groups. The two

populations are statistically different (Levene's Test, p = 2.7 × 10−37; t-test, p = 8.3 × 10−47). In all cases, the

insets show representative helical- and disordered-bridge conformations from the simulation, taken from the

labelled bins (asterisks). (F) Protein melt-curve data (n = 3). Error bars are standard deviations. **, p<0.01;

*** p<0.001. See also Figure 9, Figure 10, and Video S1.
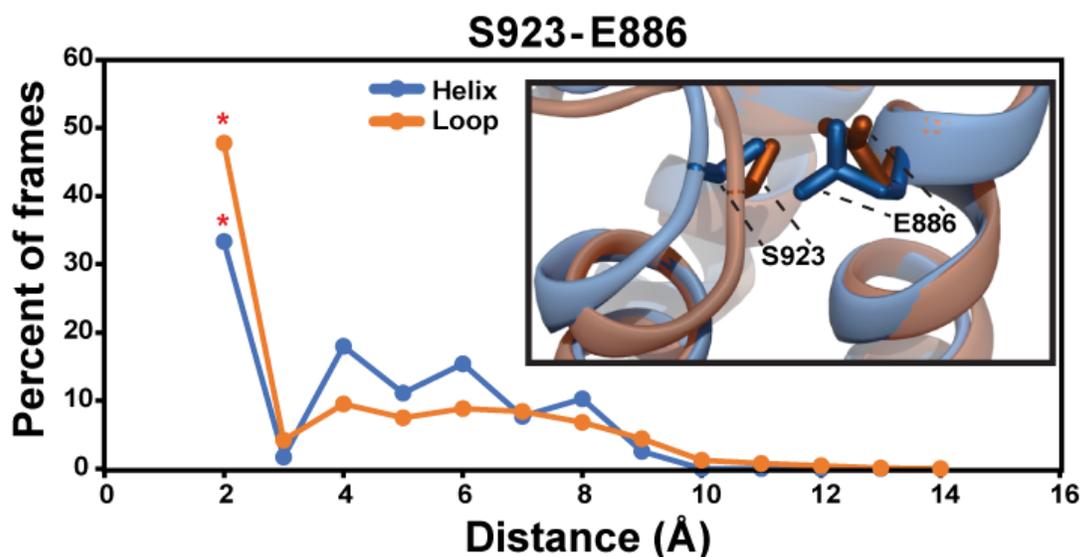
Figure 9. The dynamics of the S923-E886 interaction. The difference between the α7-α8 bridge and helical

populations was not quite statistically significant (helix n = 117, loop n = 883). Levene's Test for Equality of

Variances led us to reject the null hypothesis that the two populations have equal variances (p = 0.019). A

two-tailed t-test (unequal variances assumed) led us to fail to reject the null hypothesis that the two

populations have equal means (p = 0.092). Like the K924-E886 interaction, the S923-E886 interaction forms

and breaks multiple times over the course of the simulation. This interaction may be affected by the

formation of the α7-α8 bridge. Related to Figure 8.
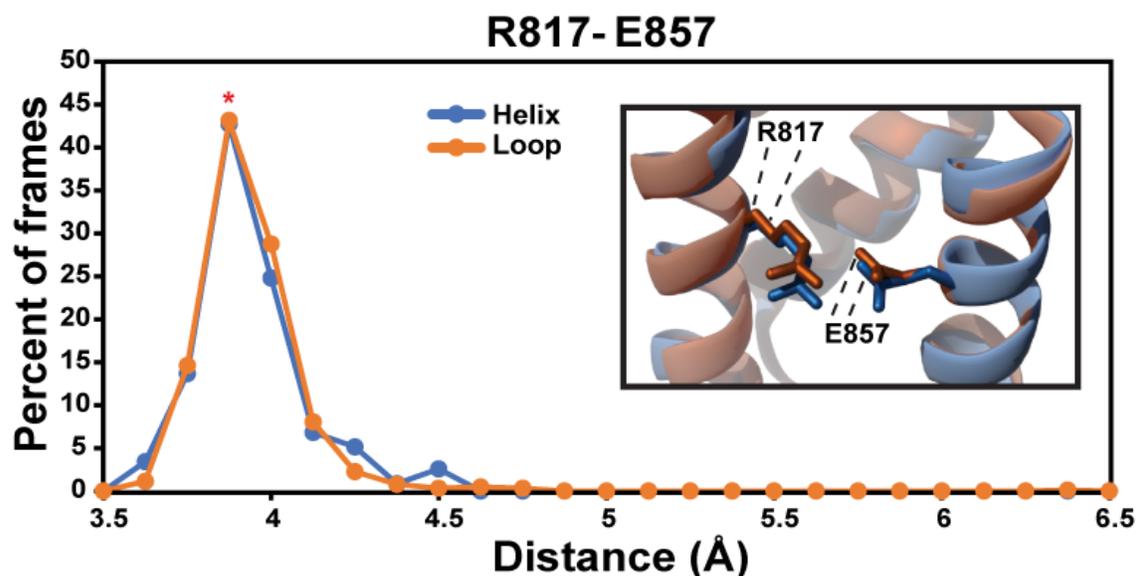
Figure 10. The dynamics of the R817-E857 interaction. This interaction is distant from and unaffected by the α7-α8 bridge conformation. Levene's Test for Equality of Variances led us to fail to reject the null hypothesis that the two populations have equal variances (helix n = 117, loop n = 883, p = 0.20). A two-tailed t-test (equal variances assumed) led us to fail to reject the null hypothesis that the two populations have equal means (p = 0.875). Related to Figure 8.

If K924 and K921 do in fact participate in the key interactions observed in the *apo* simulations, these residues should contribute to the stability of the *apo* DM15 region. To test this hypothesis, we generated K924A and K924D mutants to disrupt the MD-predicted K924–E886 interaction, and a K921A mutant to disrupt the K921–Y922 interaction. The stability of each recombinantly expressed and purified mutant was assayed using melting temperature (T_m) as a readout, where a decrease in melting temperature corresponds to a decrease in thermal stability. Compared to wild-type (WT) DM15, these mutants exhibited a significant average decrease in T_m of at least 3.2°C (Figure 8F). The K924A and K924D mutants both resulted in an equal loss of protein stability (p = 7.0×10^{-3} and p = 7.1×10^{-3}, respectively). The T_m of the K921A mutant

was also significantly lower than that of WT DM15 and K924A ($p = 2.6 \times 10^{-4}$ and $p = 4.1 \times 10^{-3}$, respectively). In contrast, the melting temperature of the K815A control mutant was not significantly different from WT (Figure 8F); K815 is distant from the m^7G pocket and does not form substantial interactions with other residues, per our simulation.

2.3.5 Key Interactions Gate TOP-mRNA Access

To better understand how the conformation of the α7–α8 bridge impacts pocket dynamics, we identified maximally open (Figure 11A and 11B, in green) and collapsed (Figure 11B, in purple) m^7G-pocket conformations. As expected, in the selected collapsed state the α7–α8 bridge is a disordered loop. This conformation allows amino acids such as Y922 and K924 to occlude mRNA binding (Figure 11). In contrast, the maximally open conformation approaches what we call the "ligand-ready state" (Figure 11, in green). In this state, the α7–α8 bridge is helical, as seen in the 4ZC4:B crystal structure (Lahr et al., 2015)[80]. The simulations show that K924 is a key pocket-occluding residue. In the collapsed state, it participates in a transient electrostatic interaction with E886, the residue primarily responsible for recognizing the Watson-Crick face of the m^7G moiety (Figure 11B, purple). This interaction displaces E886 from the mRNA-bound orientation observed crystallographically (Lahr et al., 2017; Lahr et al., 2015)[79, 80]. In the ligand-ready state, K924 is more distant from E886. E886 is thus more prone to adopt a position that can accommodate cap binding, as seen in the 4ZC4:B crystal structure and some of our simulation frames.

Y922 also plays a prominent role in determining the m^7G-pocket volume. Our simulations show that in the collapsed state, the Y922 sidechain flips in and out of the m^7G pocket, potentially occluding mRNA binding. In contrast, in the ligand-ready state, the helical α7–α8 bridge pulls

Y922 away from the binding pocket (Figure 11B, in green). K921 appears to help maintain Y922 in this mRNA-compatible conformation. A single bridging water molecule forms hydrogen bonds with both K921 and Y922, helping to mediate the K921– Y922 interaction. The 4ZC4:B structure includes this water molecule, and the 4ZC4:A electron density also suggests its presence. Using in-house scripts, we verified that the same location (relative to K921 and Y922) is often water occupied over the course of our MD simulations (Figure 13).

2.3.6 Ligand Specificity by Exploiting the TSS Pocket
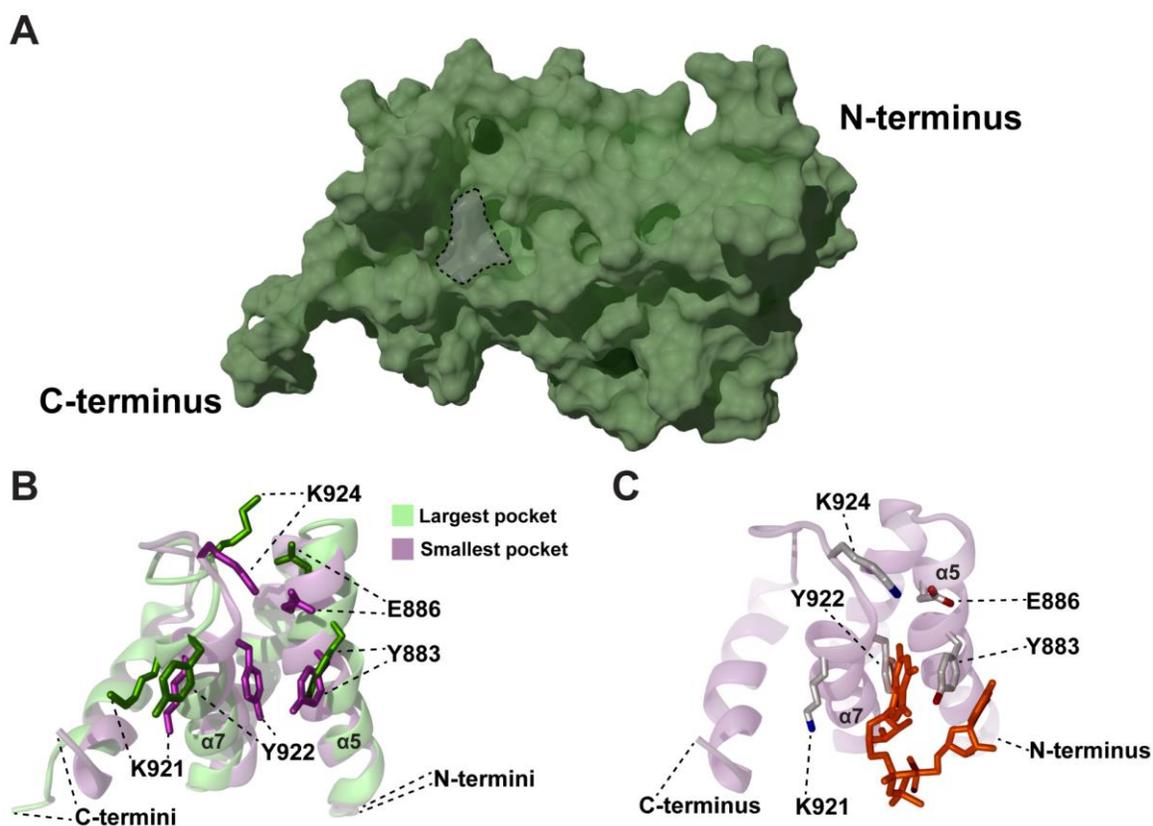
Figure 11. The largest and smallest m^{7}G-pocket conformations, by volume. (A) A surface representation of the frame with the largest pocket volume. (B) Large and small pocket conformations, superimposed. (C) An

example of a small pocket conformation. In orange, the m7GpppC ligand from PDB ID:5V87 (Lahr et al., 2017)[79] is superimposed for reference. See also Figure 12 and Figure 13.
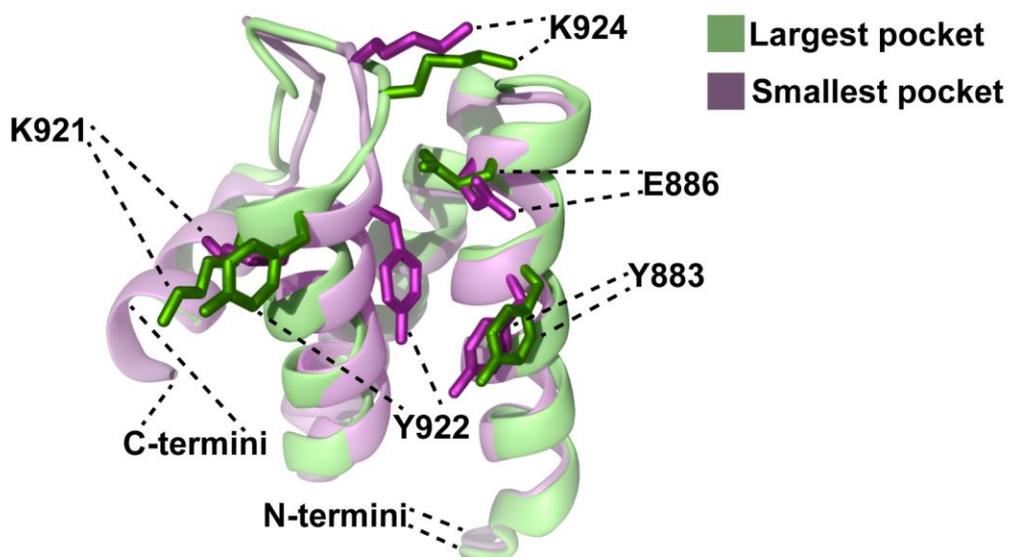
Figure 12. LARP1 4ZC4:C m7G-pocket POVME analysis. The 4ZC4:C simulation captured changes in residue positions similar to those seen in the 4ZC4:B simulation. In the smallest m7G pocket conformation, Y922 juts into the pocket, possibly occluding mRNA binding. Related to Figure 11.
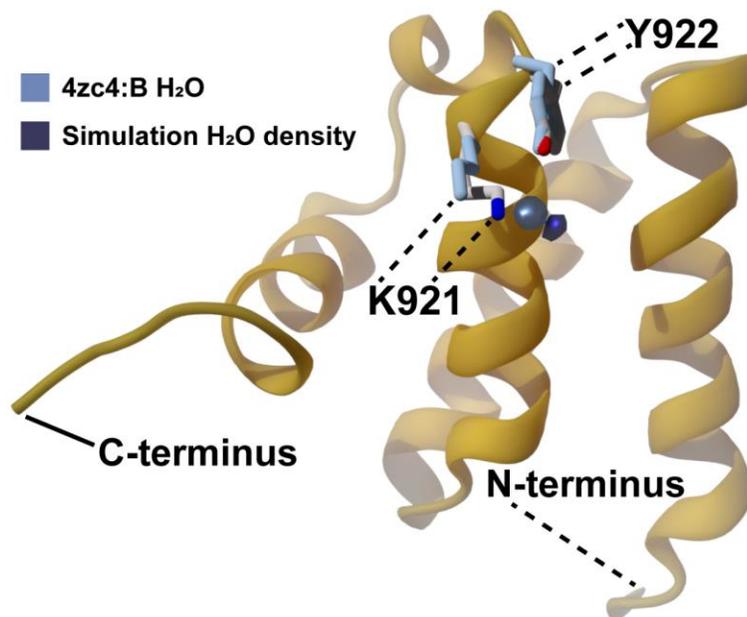
Figure 13. A transient K921-Y922 interaction is mediated by an adjacent, bridging water molecule. A water molecule from the 4ZC4:B crystal structure (light blue) coordinates the K921-Y922 interaction. An analysis of the water density surrounding K921 and Y922 over the course of the 4ZC4:B simulation reveals a region of high water occupancy at a similar location (dark blue). These findings support the hypothesis that this transient interaction is coordinated by a water molecule in this position. Much smaller non-representative water densities were removed to facilitate visualization. Related to Figure 11.

To further future drug-discovery efforts, we also studied the TSS (transcription start site) pocket, which binds the cap-adjacent cytosine that is characteristic of 5'TOP-motif-containing transcripts. Based on published crystal structures and our simulations, we hypothesized that this pocket determines TOP transcript specificity. If true, better characterizing the TSS pocket may suggest new drug-discovery approaches that achieve small-molecule/DM15 specificity through a similar mechanism.

We predicted that R847 is primarily responsible for +1 nucleotide recognition and that it determines LARP1–5'TOP-motif specificity through Watson-crick recognition of the +1 nucleotide. We further surmised that the aromatic residues stacking the +1 nucleotide, Y883 and

36

F844, impart stability and align the nucleotide in the TSS pocket via π-stacking interactions. Our simulations confirm that these residues are notably stable. Even in the *apo* 4ZC4:B simulation, Y883 and F844 have RMSF values of only 1.12 and 0.95 (Figure 6B).

To test these hypotheses, we used mutagenesis to alter the amino acids in the cytosine-specific TSS pocket. We tested R847E, F844Y, and F844W mutants (individually and in combination) for RNA binding using electrophoretic mobility shift assays (EMSA). We predicted that R847E would change sequence recognition from m^7GpppC to m^7GpppG. We further predicted that changing F844 to a larger aromatic residue would increase the stability of RNA binding. As expected, the F844Y/R847E double mutant (which we call FYRE) had a higher affinity for m^7GpppG (non-TOP) than m^7GpppC (TOP), indicating a shift of RNA substrate. It is important to note that the affinities reported here are weaker than those reported previously (Lahr et al., 2017)[79] due to substrate length; here, we tested the binding of DM15 to a 20-mer representing RPS6 rather than the 42-mer RNA sequence used previously (Lahr et al., 2017)[79].

Notably, the FYRE mutant did not merely weaken specificity for m^7GpppC relative to WT DM15; rather, it switched specificity to m^7GpppG. The FYRE variant had a single discrete shift on the EMSA even at higher protein concentrations. FYRE apparently establishes a single register of binding, similar to the single resister of WT DM15 binding to m^7GpppC (Lahr et al., 2017)[79] (Figure 14A, B).

Changing F844 in isolation did not alter the affinity of the mutant for m^7GpppG RNA (data not shown). This suggests that hydrogen bonding with R847 is important for recognition of the +1C nucleotide. It also implies that the multiple shifted bands that we previously observed with uncapped or capped +1G RNA are most likely nonspecific binding events, with each band representing increasing stoichiometries of protein:RNA (Lahr et al., 2017)[79]. Locking in the first

and second positions of the RNA restricts these non-specific binding events observed at high protein concentrations via EMSA.

To visualize the specificity switch, we resolved a cocrystal structure of m^{7}GpppG with the FYRE mutant to 2.34 Å resolution (Table 1). The structure reveals that E847 recognizes the +1G of the co-crystallized dinucleotide (Figure 14C), as predicted. Though LARP1 is a monomer in solution (Lahr et al., 2015)[80], the WT DM15 region bound to m^{7}GpppC most often crystallizes with a two-fold non-crystallographic symmetry (NCS) axis orthogonal to the phosphate linkage between the bases (Lahr et al., 2017)[79]. The crystal form of FYRE bound to m^{7}GpppG is similar. The occupancy of the dinucleotide is approximately evenly divided between the conformation shown in Figure 4C and a second conformation splayed across the NCS (Lahr et al., 2017)[79].

2.3.7 Druggability Assessment
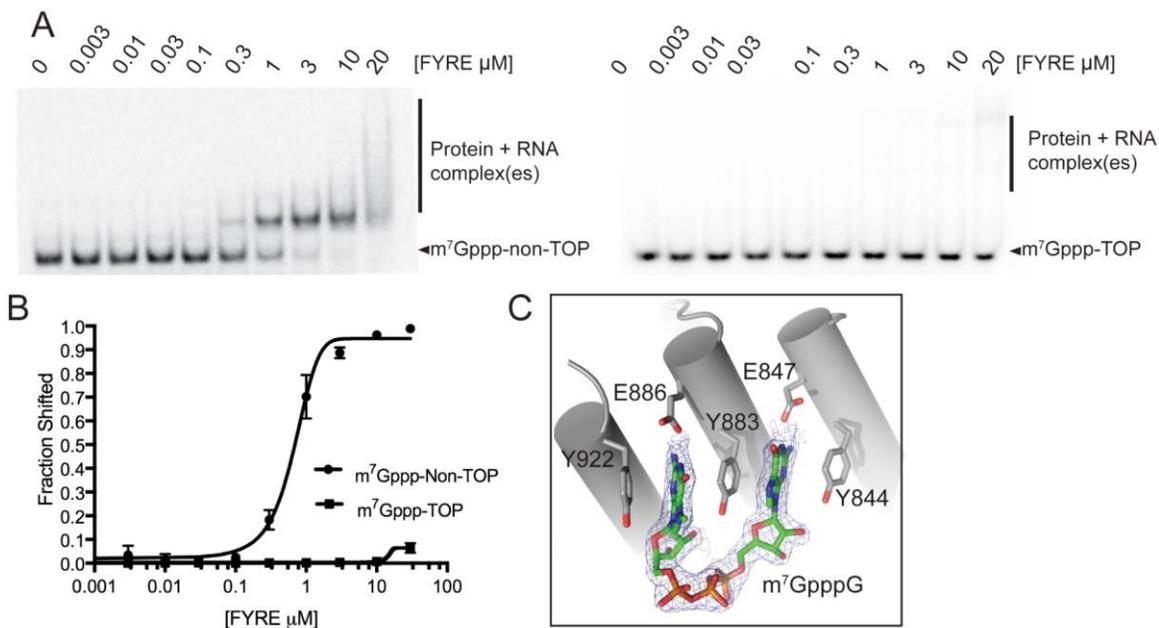
Figure 14. Mutations that affect the TSS pocket alter TOP-motif recognition. (A) A representative EMSA of FYRE DM15 binding to m7Gppp-TOP RNA and m7Gppp- non-TOP RNA, with the indicated protein titrations. (B) Quantification of triplicate binding assays of the FYRE mutant with the indicated RNAs, analyzed by EMSA. Error bars represent one standard deviation (n = 4). (C) Co-crystal structure of the FYRE mutant bound to m7GpppG (in green) resolved to 2.34 Å resolution. α-helices are represented as cylinders, and amino acids involved in dinucleotide recognition are represented as sticks. The electron density of the composite omit map contoured at 2.0 sigma, carved around the m7GpppG dinucleotide, is represented in blue mesh.

Given the potential benefits of pharmaceutically targeting the LARP1 m^{7}G and TSS pockets, we assessed the DM15 region for druggability. We used affinity propagation clustering (Frey and Dueck, 2007)[111] to extract representative ensembles of distinct LARP1-DM15 microstates from the simulations. The 4ZC4:B simulation yielded an ensemble of 12 representative conformations. We next used the FTMap server (Brenke et al., 2009)[112] to flood the surface of

39

each ensemble conformation with virtual organic probes. A custom Python script identified consistent druggable hotspots (across all ensemble members) where probes tended to congregate (Figure 15A). Our ensemble-based druggability analysis of the 4ZC4:B simulation confirmed that the m^{7}G and TSS pockets are druggable (Figure 15A, hotspots "a" and "b" in red). An unexpected druggable hotspot was also identified beneath the α7–α8 bridge (Figure 15A, hotspot "c").

To further study the druggability of the m^{7}G pocket, we repeated the FTMap analysis using the largest m^{7}G-pocket conformation (by volume) sampled over the course of the 4ZC4:B simulation. Though the TSS pocket is collapsed in this conformation, the m^{7}G pocket is much expanded. The m^{7}G-pocket hotspot (Figure 15B, hotspot "b") is contiguous with two adjacent druggable hotspots (Figure 15B, "a" and "c") that could also be pharmacologically exploited.

2.4 Discussion

LARP1 occupies a central node in mTORC1 signaling (Fonseca et al., 2015; Hong et al., 2017; Philippe et al., 2017; Tcherkezian et al., 2014)[81-84], and LARP1 protein levels are altered in several cancers (Hopkins et al., 2016; Mura et al., 2015; Stavraka and Blagden, 2015; Ye et al., 2016)[71, 86, 87, 89]. Small molecules that alter LARP1 activity could thus be developed into novel anti-cancer therapies. Crystal structures of the LARP1 DM15 region bound to TOP mRNA and m^{7}GpppC reveal two potentially druggable pockets that bind the m^{7}G cap and first nucleotide, respectively (Lahr et al., 2017; Philippe et al., 2017)[79, 81]. The m^{7}G pocket of these structures strongly resembles that of other cap-binding proteins. However, unlike other cap-binding proteins, LARP1 preferentially binds TOP transcripts with m^{7}GpppC motifs at the 5' end. Understanding

the structural, dynamic, and biochemical factors that enable this selectivity may be key to designing new small-molecule therapeutics that bind specifically to the LARP1 DM15.

2.4.1 The LARP1-DM15 m^{7}G Pocket

Remarkable Pocket Flexibility

Published crystal structures provide some evidence that the cap-binding region of LARP1 DM15 is highly flexible. The pocket-adjacent linker that connects the α7 and α8 helices (the α7–α8 bridge) has been captured in several distinct conformations (Lahr et al., 2015)[80], including a helical conformation (4ZC4:B), a loop conformation (4ZC4:C), and a loop conformation so disordered that it could not be resolved crystallographically (4ZC4:D).

The computational and experimental analyses described here build on this previous work. Our simulations suggest that the secondary structure of the α7–α8 bridge may play a critical role in promoting or discouraging TOP mRNA binding. By capturing the transitions between helical- and disordered-bridge conformations, the simulations reveal m^{7}G-pocket flexibility beyond what has been observed crystallographically. These transitions alter the volume of the m^{7}G pocket and so likely impact TOP mRNA binding.

The Collapsed State

Some of our simulation frames captured the m^{7}G pocket in a transient collapsed state that has not been seen in any crystal structure. TOP mRNA binding is unlikely in this state for three reasons. First, the disordered α7–α8 bridge allows the Y922 side chain to flip in and out of m^{7}G pocket, sterically hindering TOP mRNA binding.

Second, the bridge residue K924 transiently interacts with and displaces E886, a key residue that directly mediates TOP mRNA binding via hydrogen bonds with the m^{7}G-cap Watson-

Crick face (Lahr et al., 2017)[79]. This interaction may sequester E886 away from the optimal m⁷G-binding conformation.

Table 1. Data collection and refinement statistics for the FYRE DM15 (+1C pocket) mutant.

Data collection	
Space group	P21
a, b, c (Å)	58.58, 87.36, 72.89
β (°)	93.37
Resolution (Å)	29.24–2.34
R_{merge} (%)	0.05 (0.586)
I/ σ (I)	12.6 (1.6)
Completeness (%)	99.7 (98.1)
Redundancy	3.9 (3.7)
Refinement	
Resolution (Å)	29.24–2.34
No. of unique reflections	29,243
R_{work}/R_{free}	0.220/0.267
RMSD bond angle (°)	0.6
RMSD bond length (Å)	0.002
Average B factor	74.0
PDB	6PW3

Third, the K924–E886 interaction may also contribute to the free energy of binding in ways that indirectly promote LARP1/mRNA dissociation. Binding affinity is determined by the difference in molar Gibbs free energy between the ligand-bound and unbound states. Put another way, ligand affinity is proportional to the extent to which the binding event increases the stability of the system. When TOP mRNA is bound, the positively-charged m⁷G guanine moiety forms energetically favorable interactions with the pocket (e.g., cation-π interactions with both Y922 and Y883), enhancing the stability of the complex (Lahr et al., 2017)[79]. But a similarly favorable K924–E886 interaction occurs exclusively in the absence of bound mRNA, reducing the energetic/stability difference between the mRNA-bound and collapsed states.

The K924–E886 interaction forms and breaks multiple times over the course of the MD simulation. Although some *apo* crystal structures hint at this interaction (e.g., 4ZC4:C), the simulations suggest it is far more prominent. To confirm the role that K924 plays in stabilizing the mRNA-unbound (*apo*) protein, we used a protein melting temperature assay to evaluate K924A and K924D DM15 mutants. The melting temperatures of both were reduced by 3.2°C, showing that K924 contributes to *apo* protein stability as expected (Figure 8F). Interestingly, mutation of K924 did not change the affinity of the DM15 region for capped TOP sequence (data not shown); this result, however, was not unexpected for an assay conducted at equilibrium. We anticipate that the mutation affected the kinetics of binding, which were not explicitly tested.

2.4.2 The Ligand-Ready State

Some of our simulation frames captured the m^7G pocket in a conformation more amenable to TOP mRNA binding. In this state, the α7–α8 bridge is in a helical conformation similar to that of 4ZC4:B. The more rigid helical α7–α8 bridge rotates K924 away from E886, such that E886 more readily adopts the orientation required for mRNA-cap binding. Finally, the helical structure also pulls Y922 away from the m^7G pocket, reducing the steric hindrance seen in the collapsed state (Figure 11B, in green). The simulations suggest that interactions with K921—possibly mediated by hydrogen bonding with a common water molecule (Figure 13)—maintain Y922 in this open-pocket position (Figure 13).

We did not expect K921 to play so prominent a role in preventing Y922-mediated pocket occlusion. Several crystallographic conformations show that the side chains of these two amino acids are physically adjacent, but visual inspection of the simulations suggests a fairly robust interaction between the two. To provide experimental evidence in favor of this interaction, we used

a protein melting temperature assay to evaluate the K921A mutant. This mutant had a melting temperature 4.0°C below that of the wild type (Figure 8F), showing that K921 does in fact contribute substantially to *apo* protein stability.

2.4.3 The Collapsed and Ligand-Ready States May Interconvert in Vivo

Given that our 4ZC4:B simulation started from the 4ZC4:B conformation, it is not surprising that it sampled the α7–α8 bridge in a helical conformation. What is remarkable is that over the course of our simulation, the 3_{10} helix unfolded and then refolded (Video S1). This unfolding-and-refolding event caused DM15 to transition between the collapsed and ligand-ready states, suggesting that the energetic difference between the two is small. LARP1 DM15 may interconvert rapidly between both states *in vivo*.

Since these two states appear to be nearly isoenergetic, one would expect that even small perturbations to the protein in this region could bias the conformational ensemble in favor of one or the other. Indeed, mutations near or in the α7–α8 bridge have been observed in several cancers. I930V and I930T mutations have been implicated in lung adenocarcinoma and breast cancer, respectively (COSMIC (Forbes et al., 2017) Study IDs COSU417 and COSU652)[113]. Y922D has also been implicated in colon adenocarcinoma (COSMIC (Forbes et al., 2017) Study ID COSU376)[113].

These amino-acid changes were identified in tumors, but post-translational modifications of bridge and bridge-adjacent residues may also serve to regulate LARP1/5'TOP motif binding in a non-pathogenic context. Though LARP1 associates with mTORC1, there are no known phosphorylation sites within the simulated DM15 region (Hong et al., 2017; Philippe et al., 2017)[81,84]. However, a recent proteome-wide study identified the DM15 bridge residue K926 as a

ubiquitination site (Hornbeck et al., 2015; Udeshi et al., 2013)[114, 115]. The role of ubiquitination is not limited to targeting proteins for proteasome-mediated proteolysis (Mukhopadhyay and Riezman, 2007; Schnell and Hicke, 2003)[116, 117], so it is reasonable, albeit speculative, to hypothesize that ubiquitination here serves to alter TOP mRNA binding via changes in m^{7}G-pocket dynamics. While mTORC1 plays a role in overall ubiquitin-mediated degradation (Zhao and Goldberg, 2016; Zhao et al., 2015)[118, 119], ubiquitination signaling (unrelated to degradation) is also known to modulate mTORC1-pathway activity (Deng et al., 2015)[120].

2.4.4 Future Avenues for Drug Discovery: Targeting the m^{7}G Pocket
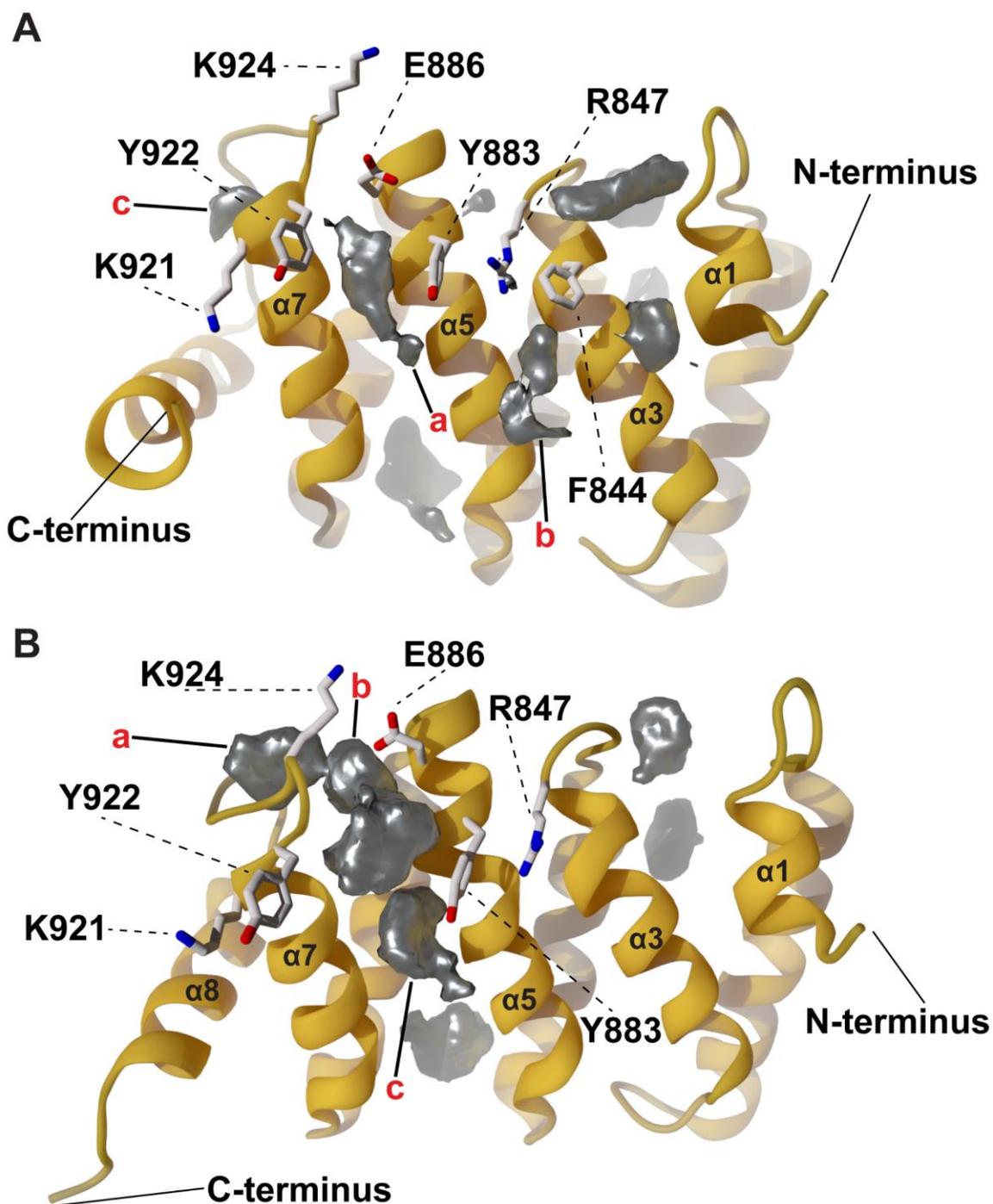
Figure 15. LARP1 DM15 druggable hotspots. **(A) The locations of persistent druggable hotspots across all twelve 4ZC4:B ensemble conformations are shown as silver surfaces. These surfaces encompass regions**

The proposed role that m^7G-pocket flexibility plays in regulating TOP mRNA binding provides novel opportunities for drug discovery. The DM15 conformations captured crystallographically show a cap-binding pocket that resembles those of other cap-binding proteins (Lahr et al., 2017)[79]. These similarities suggest that LARP1-DM15 ligands targeting the crystallographic m^7G-pocket conformation may be promiscuous, leading to unacceptable side effects.

To assess the druggability of alternate, non-crystallographic m^7G-pocket conformations, we applied FTMap (Kozakov et al., 2015; Ngan et al., 2012)[121, 122] to the 4ZC4:B-simulation conformation with the largest m^7G pocket (per POVME 2.0 (Durrant et al., 2014)[110], Figure 15B). In this conformation, a cryptic druggable pocket opens near the key α7–α8 bridge (Figure 15B, hotspot "a" in red), connected to the m^7G pocket via a narrow channel. The bridge pocket is not directly involved in m^7G binding and so is less likely to be conserved among cap-binding proteins (Lahr et al., 2017; Lahr et al., 2015)[79, 80].

The optimal pharmacological strategy for targeting the m^7G pocket will depend on the prevailing cellular role of LARP1 (e.g., sustaining vs. repressing TOP mRNA translation). LARP1 helps regulate both TOP mRNA stability and TOP mRNA translation. We and others have observed that LARP1 represses the translation of TOP mRNAs (Fonseca et al., 2015; Philippe et al., 2017)[81, 83] while simultaneously protecting these (and other) mRNAs from degradation (Aoki et al., 2013; Fonseca et al., 2015; Gentilella et al., 2017)[83, 105, 108]. If in protecting TOP mRNAs from degradation LARP1 predominantly promotes the production of ribosomes, then compounds that disrupt TOP mRNA binding could have anti-cancer properties. DM15 ligands that bind the

m[7]G pocket and extend chemical moieties into the bridge pocket could be both potent and specific direct competitors. On the other hand, some evidence suggests that LARP1 predominantly represses TOP mRNA translation (Fonseca et al., 2015; Philippe et al., 2017)[81, 83], likely by preventing the assembly of the eIF4F complex on the 5'UTR of TOP transcripts (Lahr et al., 2017; Philippe et al., 2017)[79, 81] and sequestering those transcripts in SGs (Wilbertz et al., 2019)[106]. If repression of TOP mRNA translation dominates, compounds that encourage LARP1 to adopt the ligand-ready conformation could have anti-cancer properties. Such compounds could block TOP mRNA translation, promoting transcript shuttling to SGs where they are kept in a translation-repressed state. Allosteric DM15 ligands that bind the bridge pocket alone could encourage the ligand-ready conformation, perhaps by stabilizing the α7–α8 bridge in the mRNA-amenable helical conformation. Aside from serving as anti-cancer drug leads, future ligands will also be useful chemical probes for better understanding these complex translational regulatory mechanisms.

2.4.5 The LARP1-DM15 TSS Pocket

The TSS Pocket Governs TOP mRNA Specificity

The TSS pocket, which binds the cap-adjacent cytosine that is characteristic of TOP mRNAs (Lahr et al., 2017; Lahr et al., 2015)[79, 80], was far more stable in our simulations. The amino acids comprising this pocket moved little over the course of the 4ZC4:B simulation, with RMSF values of 0.85 (R847), 0.95 (F844), 1.12 (Y883), and 1.15 (R879). R847, the amino acid that recognizes the +1C Watson-Crick face of TOP RNAs, was particularly stable.

Despite this stability, the TSS pocket presents an important opportunity for LARP1-specific drug discovery. LARP1 is unique among cap-binding proteins in its ability to recognize

the m^7GpppC sequence at the 5' ends of TOP transcripts (Lahr et al., 2017; Philippe et al., 2017)[79, 81]. It is the +1 nucleotide, or the first nucleotide after the cap (the cytosine) that apparently determines specificity.

To test this hypothesis, we used mutagenesis to change RNA specificity from a capped-TOP transcript to a capped-non-TOP transcript (m^7GpppC to m^7GpppG). We specifically selected mutations that would make the TSS pocket resemble the m^7G pocket, changing F840 to tyrosine and R847 to glutamic acid. Comparison of this so-called FYRE mutant with mutants of each individual residue shows that the conserved arginine does, indeed, determine DM15 specificity for the +1 nucleotide.

2.4.6 Future Avenues for Drug Discovery: Targeting the TSS Pocket

Compounds that exploit the unique, LARP1-specific TSS pocket may be less likely to bind promiscuously to other cap-binding proteins. The TSS pocket is itself fairly shallow; ligands that bind this pocket alone are unlikely to have high affinities. But m^7G-pocket ligands that extend moieties into the TSS pocket, as do endogenous TOP mRNAs, could be developed into therapeutics with fewer toxic side effects in the clinical setting. Our ensemble-based druggability assessment (Figure 15A), which suggests that both the m^7G and TSS pockets are druggable, supports this strategy.

2.5 STAR★Methods

2.5.1 Key Resources Table

Table 2. Key resources table.

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bacterial and Virus Strains		
NEB DH5 α competent E.coli	NEB	cat#C2987
BL21(DE3) competent E.coli	NEB	cat#C2527
Chemicals, Peptides, and Recombinant Proteins		
m7GpppG	NEB	cat#S1405S
Sypro Orange	Thermo Fischer Scientific	cat#S6650
T4 PNK	NEB	cat#M0201
vaccinia capping enzyme	NEB	cat#M2080
α -32P-GTP	Perkin Elmer	cat#BLU506H250UC
His-Pur NiNTA	Thermo Fischer Scientific	cat#88222
HiTrap Q	GE	cat#17115401
HiTrap SP	GE	cat#17115201
HiTrap Butyl	GE	cat#28411005
10K MWCO concentrator	Millipore	cat#C7715
calf intestinal phosphatase	Roche	cat#11097075001
Deposited Data		
Apo X-ray crystal structure of DM15	(Lahr et al., 2015)[80]	PDB: 4ZC4
F844YR847E DM15 X-ray crystal structure	This study	PDB: 6PW3
Oligonucleotides		
Forward and reverse primers for K924D Mutation of LARP1 DM15 5'CTGGGCCTTCTTGAAATATTCCGACGCCAAAAATTTGGAC3' 5'GTCCAAATTTTTGGCGTCGGAATATTTCAAGAAGGCCCAG3'	Sigma	N/A
Forward and reverse primers for K924A Mutation of LARP1 DM15 5'CTGGGCCTTCTTGAAATATTCCGCAGCCAAAAATTTGGAC3' 5'GTCCAAATTTTTGGCTGCGGAATATTTCAAGAAGGCCCAG3'	Sigma	N/A
Forward and reverse primers for K921A Mutation of LARP1 DM15 5'caaatttttggctttggaatatgccaagaaggcccagaacttctcc 3' 5'ggagaagttctgggccttcttggcatattccaaagccaaaaatttg3'	Sigma	N/A
Forward and reverse primers for K815A Mutation of LARP1 DM15 5'cacaacacgtctaccatgcgtatcgtaggcgctgcc3' 5'ggcagcgcctacgatacgcatggtagacgtgttgtg3'	Sigma	N/A
Forward and reverse primers for F844YR847E Mutation of LARP1 DM15 R847E 5'catctttttgttgaagtgatcctcgaggaagaaggaccagaagcg3' 5'cgcttctggtccttcttcctcgaggatcacttcaacaaaaagatg3' F844YR847E 5'catacatctttttgttgaagtgatcctcgaggaaataggaccagaagcggaagagtgtgt3' 5'acacactcttccgcttctggtcctatttcctcgaggatcacttcaacaaaaagatgtatg3'	Sigma	N/A

Splint adapter 5' CTTGAAGCAGCTGAACGCCTCCGAGGCGCCACGGAAAAGAGG 3'	Sigma	N/A
Recombinant DNA		
human LARP1 isoform 2	ThermoFisher	BC033856

Lead Contact and Materials Availability

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Jacob D. Durrant.

Experimental Models and Subject Details

NEB DH5α competent E. coli (cat# C2987) cells were used for mutagenesis and plasmid preparation of DM15. Recombinant protein expression was performed with BL21(DE3) competent E. coli cells (cat# C2527). Cells were cultured according to the manufacturer's protocols, and protein expression was performed as detailed below.

Method Details

2.5.2 Model Building and Parameterization

We considered chains B and C of the 4ZC4 LARP1 DM15 structure (Lahr et al., 2015)[80] separately. We selected these chains because they capture S923-L928, a region near the 7-methylguanosine (m7G) pocket, in two distinct conformations. To prepare each chain for simulation, we added hydrogen atoms using the PDB2PQR 2.1.1 (Dolinsky et al., 2007; Dolinsky et al., 2004)[123, 124] implementation of the PROPKA algorithm (Olsson et al., 2011)[125], with the pH set to 7.0. PDB2PQR also optimized the hydrogen-bond network. We then used the Ambertools18 tleap program (Case et al., 2017)[126] to add a water box extending 10 Å beyond the protein in all three dimensions; Cl− counterions as required to bring the system to electrical neutrality; and Na+ and Cl− counterions as required to achieve a 150 mM concentration.

We used Ambertools18 to parameterize each system. The protein and counterions were parameterized according to the Amber ff14SB force field (Maier et al., 2015)[127]. The water molecules were parameterized according to the TIP3P forcefield (Jorgensen et al., 1983)[128].

2.5.3 Minimization, Equilibration, and Production Simulation

We minimized and equilibrated both systems (4ZC4:B and 4ZC4:C) using the NAMD 2.9 molecular dynamics simulation package (Kale et al., 1999; Phillips et al., 2005)[129, 130]. The minimization protocol included four phases of 5,000 minimization steps each. We first relaxed hydrogen atoms; then hydrogen atoms and water molecules; then hydrogen atoms, water molecules, and protein side chains; and finally all atoms. Following minimization, four 0.25-ns isothermal–isobaric (NPT) simulations were used to equilibrate each system (310 K). We applied harmonic constraints to the protein backbone atoms. The associated restraining forces were gradually weakened: 1.0, 0.75, 0.5, and 0.25 kcal/mol/Å^2 for each phase, respectively. Following these equilibration steps, the systems were next subjected to extended NPT production runs. We simulated 4ZC4:B and 4ZC4:C for 500 ns each. All figures showing simulation-derived DM15 structures (Figures 6C, 8A, 8B, 11, and 15) were generated using BlendMol (Durrant, 2018)[131].

2.5.4 Root Mean Square Deviations and Fluctuations

We used MDAnalysis 0.16.2 (Michaud-Agrawal et al., 2011)[132] to align each trajectory by its C$_\alpha$. We extracted frames spaced 100 ps apart and calculated the RMS deviations (RMSDs) between each and the corresponding first frame. Early frames were discarded because their

RMSDs had not yet sufficiently equilibrated. Subsequent analysis focused on the last 88% and 90% of the 4ZC4:B and 4ZC4:C simulations, respectively.

We used frames spaced 10 ps apart to calculate RMSF values for each residue center of geometry. The RMSF values for the 4ZC4:B and 4ZC4:C simulations were calculated separately using custom Python scripts built on MDAnalysis.

2.5.5 Hotspot Analysis

We used the affinity propagation algorithm implemented in MDAnalysis to cluster each trajectory (AffinityPropagationNative, preference = −35). A representative conformation (centroid) was selected from each cluster. The set of centroids associated with each simulation constitutes a conformational ensemble. We analyzed each 4ZC4:B and 4ZC4:C cluster centroid with FTMap (Kozakov et al., 2015; Ngan et al., 2012)[121, 122] to identify druggable hotspots.

For a given protein conformation, FTMap outputs a PDB file that includes both the protein structure and multiple small organic probes docked into the identified hotspots. To more easily identify persistently druggable regions, we superimposed the FTMap outputs associated with each ensemble conformation. We wrote a Python script that converted all docked probes across all ensemble members into a single density map. For each probe atom, we generated a set of 1,000 points distributed according to a 3D Gaussian function centered on the corresponding atom, with a standard deviation of 0.5 Å. We then used MDAnalysis to bin those points into cubic voxels (0.5 Å × 0.5 Å × 0.5 Å) and to output the corresponding density map in DX format. Programs such as VMD (Humphrey et al., 1996)[133] can visualize this density as a surface at a given isovalue. Selecting different isovalues allows one to easily identify regions with varying degrees of persistent (ensemble-wide) druggability.

2.5.6 POVME Analysis

We used the POVME 2.0 algorithm (Durrant et al., 2014)[110] to track the volume of both the m^7G and TSS (transcription start site) pockets over the course of the 4ZC4:B and 4ZC4:C simulations. One thousand aligned, regularly spaced frames from each simulation were considered.

To measure pocket shapes and volumes, POVME requires users to first identify an inclusion region that encompasses the many sampled pocket conformations. We used carefully chosen spheres to define these regions. For the m^7G pocket, we found that a single inclusion sphere with radius 9 Å encompassed all conformations. The shape of the TSS pocket was more complex; the associated inclusion region was taken to be the union of four carefully chosen spheres of radii 5, 5, 4, and 3 Å, respectively. Each inclusion region was filled with probe points spaced 0.5 Å apart in the X, Y, and Z directions.

For each pocket conformation (simulation frame), probe points outside the pocket itself were removed. Specifically, POVME first removed all points within 1.09 Å of any protein atom. Second, POVME removed points outside the convex hull defined by the receptor atoms. POVME also removed any points that were not contiguous with a user-defined "seed region." This region identifies portions of the pocket that are persistently open (i.e., open in all simulation frames). For the m^7G pocket, we defined this region using a single sphere of radius 3 Å. For the TSS pocket, the region was taken to be the union of three spheres of radii 3, 3, and 2 Å, respectively.

POVME calculates the volume of each pocket conformation by counting pocket-occupying points. Each point corresponds to a cubic volume of 0.125 Å^3 (0.5 Å × 0.5 Å × 0.5 Å). POVME

also tracks how often each point is pocket occupying over the course of the simulation. It outputs a density map that can be used to determine which pocket regions are most persistently open.

2.5.7 Protein Purification

We cloned amino acids 796–946 of LARP1 isoform 2 from cDNA (OpenBioSystems, now ThermoFisher [BC033856]) into a modified pET28a+ vector as described previously (Lahr et al., 2015)[80]. Point mutants were generated using this vector as template for site-directed mutagenesis, using the mutagenic primers listed in the key resource table. The resulting constructs produced wild-type or mutant DM15 with an N-terminal His$_6$-MBP tag, followed by a tobacco etch virus (TEV) protease cleavage site and a glycine$_6$ linker.

Expression plasmids were transformed into BL21(DE3) E. coli and grown overnight on LB agar plates supplemented with 30 µg/ml kanamycin. The His$_6$-MBP-DM15 fusion protein was expressed by autoinduction (Studier, 2005)[134] for 3 hours at 37°C, and then for 18 hours at 18°C. Cells were collected by centrifugation, flash frozen in liquid nitrogen, and stored at −80°C until used.

Cells (~2g) were resuspended at 4°C by gentle stirring in 50mL NiNTA lysis buffer (50 mM Tris-HCl, pH 7.5 or 8, 400 mM NaCl, 10 mM imidazole, 10% glycerol). Protease inhibitors PMSF (1 µM final concentration), leupeptin (0.1 µM final), and aprotinin (0.1 mM final) were added. Cells were lysed by homogenization, followed by clarification via centrifugation (12,000 RPM) at 4°C for 30 minutes. The soluble fraction was nutated with 4 mL equilibrated HisPur Ni-NTA Resin (ThermoFisher) for 2 hours at 4°C. The beads were washed two times in 50 mL lysis buffer and three times with wash buffer (50 mM Tris-HCl, pH 7.5, 400 mM NaCl, 35 mM imidazole, 10% glycerol). His$_6$-MBP-DM15 fusion protein was eluted from beads in 30 mL elution

buffer (50 mM Tris-HCl, pH 7.5, 400 mM NaCl, 250 mM imidazole, 10% glycerol). The N-terminal His_6-MBP tag was removed by the addition of 2 mg TEV protease for cleavage overnight in 10K MWCO SnakeSkin dialysis tubing (ThermoFisher) in 2 L of dialysis buffer (50 mM Tris-HCl, pH 8, 150 mM NaCl, 1 mM EDTA, 10% glycerol, 1 mM DTT) at 4°C.

Cleaved DM15 protein was further purified by tandem HiTrap Q and HiTrap SP columns (GE Lifesciences). DM15 protein free of nucleic acid contaminants was eluted off the HiTrap SP column with gradient from 150 mM NaCl to 1 M NaCl over 50 mL. MBP flowed through both columns, untagged DM15 eluted at 33 mS/cm, and uncleaved fusion protein eluted at 20 mS/cm, allowing for efficient separation of target DM15 construct. Fractions containing DM15 were pooled and brought to 1 M ammonium sulfate by the dropwise addition of 3 M ammonium sulfate with gentle swirling. The protein was diluted to 40 mL in 50 mM Tris-HCl, pH 7, 1 M ammonium sulfate and loaded onto a 5 mL Butyl HP column (GE Lifesciences) at 0.5 mL/min. The Butyl HP column was eluted over a 10 CV gradient to 50 mM Tris-HCl, pH 7, 2 mM DTT. Fractions containing DM15 were collected and concentrated to 20 mg/mL using a 10K MWCO spin concentrator (Millipore) in 50 mM Hepes, pH 7, 50 mM NaCl, 2 mM DTT for crystallographic experiments, or to ~2.0 mg/mL in 50 mM Tris-HCl, pH 7.5, 250 mM NaCl, 25% glycerol, 2 mM DTT for biochemical experiments. Protein was flash frozen in 10 μL aliquots and stored at −80°C for further use.

2.5.8 Crystallization and Structure Solution

The F884Y/R847E (FYRE) mutant was concentrated to ~10 mg/mL in 25 mM Hepes, pH 7.0, and 75 mM NaCl. FYRE and m^7GpppG (NEB cat#S1405S) were incubated at a 1:1.2 ratio at room temperature (final concentrations of 480 μM (9 mg/mL) DM15 and 576 μM m^7GpppG).

Crystals were set up using hanging drop vapor diffusion with a 1:1 ratio of complex to mother liquor in a final drop volume of 3 μL. The initial crystals grown in 100 mM Hepes, pH 7.5, 0.15 M NaCl, and 15% PEG 3350 diffracted poorly. To optimize crystals for publication-quality diffraction, crystals were slowly equilibrated into 100 mM Hepes 7.5, 0.08 M NaCl, and 36% PEG 3350 using the following scheme: crystals grew overnight and were allowed to equilibrate for 24 hours in the original hit condition. Every 24 hours, 0.5 μL of a new mother liquor solution was added to the drop, and the coverslip was moved to a well containing the new mother liquor. The salt was decreased by 5 mM and the PEG 3350 was increased by 1.5% per transfer, for a final well solution of 100 mM Hepes 7.5, 0.08 M NaCl, and 36% PEG 3350.

X-ray diffraction data were collected at NLS-II 17-ID-1 (AMX) at wavelength 0.92 Å using an Eigen 9M detector. An initial map was generated by molecular replacement using a 3 Å trimmed model of chain A of 4ZC4 (Lahr et al., 2015)[80] as a search model. To remove bias from the initial search model, simulated annealing composite omit maps were used to confirm amino acid register [Phenix (Adams et al., 2010), CCP4 (Winn et al., 2011)][135, 136]. Iterative building and refinement using xyz coordinates, real-space, occupancies, and individual B-factor parameters were performed in Coot (Emsley et al., 2010)[137] and Phenix (Adams et al., 2010)[135], respectively. Figure 14C was generated with the PyMOL Molecular Graphics System (Schrödinger, LLC).

2.5.9 Protein Melting Temperature Assays

Three technical triplicates were performed using an 80 μL reaction mixture composed of 20 μL (4x) 40% glycerol/4 mM β-mercaptoethanol mix; 4 μL 1M Tris, pH 8; 1.33 μL 3 M NaCl; and 8 μL 100x SYPRO orange (Thermo Fisher Scientific, cat#S6650). Protein was added to a final concentration of 5 μM. Reactions were incubated at room temperature for 10 min. The three 20

μL technical replicates were performed using the respective reaction mixes in separate wells of a 96-well real-time PCR plate. The final concentration of the buffer components varied slightly depending on the volume of protein required to reach 5 μM final concentration. All reactions contained a final concentration of 10x SYPRO orange and 1 mM β-mercaptoethanol, but the final concentrations of the other buffer components were 50 mM-53 mM Tris pH 8; 55 mM-67mM NaCl; 37 μM-130 μM DTT; and 11%–12% glycerol.

Protein unfolding was measured by monitoring the fluorescence of SYPRO orange using the x3-m3 peak channel of the QuantStudio™ 3 System (Thermo Fisher Scientific) during a temperature ramp from 30–90°C. To calculate melting temperatures, fluorescence data were analyzed using the melt temp feature of the QuantStudio™ Design & Analysis Software, per the QuantStudio™ 3 and 5 Real-Time PCR Systems Installation, Use, and Maintenance Guide (Pub. No. MAN0010407, Rev. C.0). Three independent replicates were performed, each consisting of three averaged technical replicates.

2.5.10 RNA Preparation

The 5'-triphosphorylated 20-mer RPS6 TOP-RNA sequence (CCUCUUUUCCGUGGCGCCUC) was synthesized on a 1 μmol scale using a MerMade4 DNA synthesizer (Bioautomation, Irving, TX) with ultramild RNA and 2′-OMe RNA phosphoramidites with phenoxyacetyl (PAC), acetyl (Ac), or isopropyl-phenoxyacetyl (iPrPAC) protecting groups; CPG supports; and standard solid-phase synthesis reagents (Glen Research, Sterling, VA and Chemgenes, Wilmington, MA). The 5' triphosphate was installed on the RNA on solid support using freshly prepared tributylammonium pyrophosphate according to the procedures and protocol for 5'-phosphitylation, hydrolysis, oxidation, and substitution (with the tributylammonium

pyrophosphate) in ref. (Zlatev et al., 2012)[138]. Following synthesis, cleavage from the CPG beads and deprotection was performed using standard protocols. After deprotection, RNA was analyzed for purity using reverse phase HPLC. The HPLC system consisted of a Waters 1525 pump system, a Waters 2998 photodiode array detector, a Waters XBridge OST C18 Column 2.5 um (4.6 × 50 mm) in 0.1 M triethylamine acetate, and 80:20 acetonitrile:water in 0.1 M triethylamine acetate at 25°C. Mass spectrometry of the RNA sequence was performed on an Applied Biosystems Voyager DE-STR MALDI-TOF instrument in positive mode using a 3-hydroxypicolinic acid matrix. Mass calculated: 6445; mass found: 6443 [M-2H]. Both the TOP and non-TOP (+1G) 5'- triphosphate RNAs were capped using NEB vaccinia capping enzyme (M2080) and $\alpha-^{32}$P-GTP, and then gel extracted.

2.5.11 Electrophoretic Mobility Shift Assays

Each binding reaction contained the indicated final concentration of recombinant human LARP1 DM15, 500 counts of radiolabeled RNA (<2 nM final concentration), binding buffer (20 mM Tris–HCl, pH 8, 150 mM NaCl, 10% glycerol, 1 mM DTT), 1 µg bovine serum albumin (BSA), and 0.5 µg tRNA. Reactions were incubated on ice for 30 minutes and analyzed on an 7% polyacrylamide (29:1) native 0.5× TBE native gel. Gels were run at 4°C before being dried and exposed overnight to phosphor screens and scanned with a Fuji plate reader. Pixel density of shifted (complexed) RNA over total counts was quantified using Image Quant and graphed using GraphPad Prism.

2.5.12 Quantification and Statistical Analysis

We used IBM® SPSS® Statistics 25 for statistical analyses. To compare the means of two populations, we first used Levene's Test of Equal Variances to determine whether equal population variances could be assumed. We then used an independent-samples two-tailed t-test to assess the difference in populations means, selecting the appropriate p-value given the conclusion of Levene's test.

For protein melting temperature assays, three independent replicates were performed (n = 3).

For the $\alpha7$-$\alpha8$-bridge probability distribution analyses, we compared helical and loop datasets (helical bridge, n = 117; disordered bridge, n = 883).

EMSAs (n = 4) where analyzed with ImageQuantTL by quantifying the free RNA (unbound) pixel density and bound RNA (complex) pixel density. The ratio of bound over total pixel density was plotted as a function of protein quantity using GraphPad Prism 7 (Prism version 7.00 for Windows, GraphPad Software, La Jolla California USA, www.graphpad.com). Curve fitting and affinity calculations were performed using the model representing one site-specific binding to saturation.

2.5.13 Data and Code Availability

We will share data, materials, and computer code with the community for academic and non-commercial use upon reasonable request, within a reasonable amount of time. We will provide relevant reagent samples, as requested.

Coordinates of the structure described in this article have been deposited in the PDB with accession number PDB: 6PW3.

2.6 Acknowledgements

2.7 Author Contributions

A.J.B. and J.D.D. designed the research. K.C.C., R.M.L., A.J.B., and J.D.D. performed the research. S.M. and S.R.D. provided reagents. K.C.C., R.M.L., J.C.K., A.J.B., and J.D.D. analyzed the data. K.C.C., R.M.L., J.C.K., B.D.F., A.J.B., and J.D.D. wrote the paper.

2.8 Declaration of Interests

The authors declare no competing interests.

3.0 Identification of Auxiliary LARP1 DM15 Binding Pockets

3.1 Introduction

A cell's need for protein production varies, and so it is crucial that this production is tunable. Environmental conditions, such as availability of nutrients or the presence of stressors, provide signals for the cell, and in turn the cell reacts by altering protein production. A key step in this process is the production of the translation machinery, such as the ribosome[3, 11]. The mRNAs that encode ribosomal proteins contain characteristic untranslated nucleotides that serve regulatory purposes, including responding to signals for accelerating or halting mRNA translation. These transcripts are referred to as terminal oligopyrimidine (TOP) mRNAs because they have an invariant +1C followed by a stretch of pyrimidines at their 5' ends[2, 28, 29]. The mammalian target of rapamycin complex 1 (mTORC1) regulates the translation of these TOP mRNAs[30]. Unsurprisingly, mTORC1 dysregulation has been connected to multiple cancers[10, 33]. Due to the diverse roles that this kinase complex has in cells, many off-target effects are induced when it is targeted pharmacologically[32, 64, 139, 140]. Targeting nodes downstream of mTORC1 could reduce these off-target effects, while still being an effective therapeutic strategy.
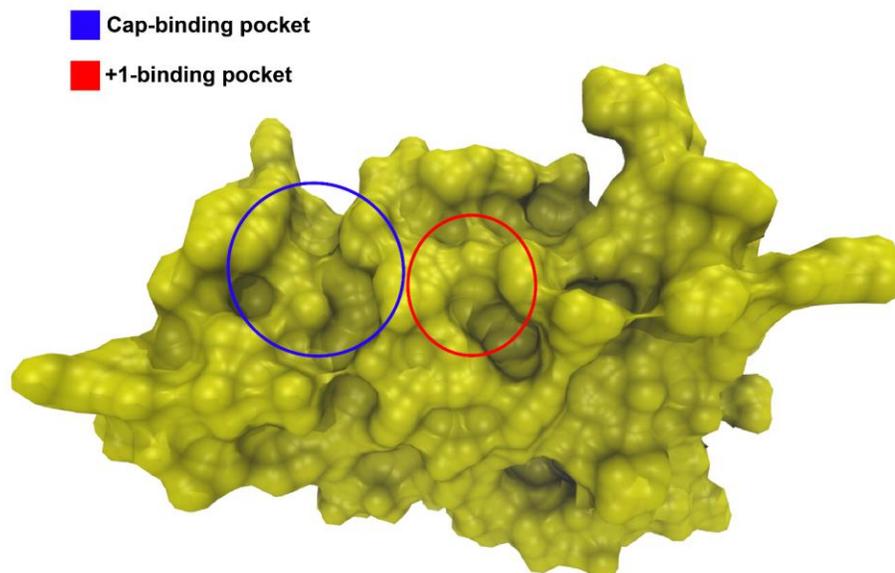
Figure 16. Cap- and +1-binding pockets of LARP1 DM15. A Surface representation of the TOP mRNA binding pockets of LARP1 DM15 (PDB ID: 4ZC4:B)[80]. Blue circle indicates the cap-binding pocket, and the red circle indicates the +1-binding pocket. All subsequently indicated orientations are relative to this pose.

La related protein 1 (LARP1) has recently been identified as a factor that regulates the translation of TOP mRNAs downstream of mTORC1[2, 81-83]. LARP1 has also been identified as a substrate of mTORC1, and phosphorylation by mTORC1 alters its association with TOP mRNAs[35, 82]. The DM15 region, located at the LARP1 C-terminus, recognizes the 7-methylguanosine cap and invariant +1C of TOP mRNAs (Figure 16)[79, 80]. All subsequently indicated orientations are relative to Figure 16. Our previous work has shown that the +1-pocket of the DM15 region drives specificity for TOP mRNAs[98]. We also showed that the cap-binding pocket transitions between an open "ligand ready" and a closed collapsed state. This transition was correlated with the secondary structure of the adjacent α7- α8 bridge, indicating that sites beyond where the RNA binds can play a role in modulating TOP mRNA translational regulation.

Currently the literature supports seemingly contradictory repressive and stimulatory roles of LARP1 DM15 in the translation of TOP transcripts[2, 81-83, 105]. When LARP1 DM15 binds the cap and +1 nucleotide of TOP mRNAs, it occludes the translation initiation factor eIF4E from binding to the cap and therefore represses translation initiation of these transcripts[79, 80]. LARP1 knockdown results in a greater abundance of TOP transcripts associated with heavy polysome fractions, providing evidence for a repressive role[83]. However, Fonseca et al. also determined that overexpression of LARP1 resulted in greater TOP mRNA abundance, indicating a role in stabilizing these transcripts. In contrast, a separate study showed that LARP1 knockdown results in a greater abundance of TOP transcripts associated with lighter polysome fractions, indicating a role of stimulating translation[82]. Long-term LARP1 knockdown has also been shown to result in a greater abundance of proteins that TOP mRNAs encode[82]. Allosteric regulation of the LARP1 DM15 region, via mTORC1 phosphorylation, has been shown to alter the affinity for TOP mRNA and could be the key to reconciling these seemingly contradictory data[35]. As of now, none of the LARP1 DM15 structures contain the residues of the characterized mTORC1 phosphorylation sites. However, investigating the effects of potentially allosteric pockets within the DM15 region could also provide insight into the allosteric mechanisms tied to the TOP mRNA binding pockets.

The connection between LARP1 and cancer also supports an overall enhancing role in TOP mRNA translation, because the cancer associated cell growth and proliferation requires an increased rate of protein synthesis. High LARP1 transcript and protein levels have been associated with several carcinomas[71, 85, 86, 88, 94]. LARP1 promotes cell growth and *in vivo* tumorigenesis [71, 85, 86]. LARP1 enhancement of TOP mRNA translation or stabilization would be consistent with these data and lead to an increase in the abundance of the machinery (or messages encoding it) needed for ribosome biogenesis that is necessary for cell growth and proliferation. LARP1 is an attractive

pharmacological target because it is downstream of mTORC1, and therefore targeting this protein may decrease off-target effects.

Although the cap- and +1-binding pockets have been well characterized, other pockets in the DM15 region have not, and may be pharmacologically exploitable. Targeting the cap-binding pocket alone may lead to low specificity ligands, due to the abundance of other cap-binding proteins with similar binding pockets. The +1 pocket is likely more difficult to target because it is shallow[79, 80, 98]. Alternative pockets that are allosterically linked to the cap- and/or +1-binding pocket may allow for greater compound specificity because they likely will be more distinct, while still modulating DM15/TOP mRNA binding. Investigating alternative pockets also provides greater insight into the biological mechanism by which LARP1 regulates TOP mRNA translation and could shed light on the seemingly contradictory evidence.

Here we characterize three putative binding pockets in the LARP1 DM15 region, two of which are predicted to be allosterically linked to the cap- or +1-binding pocket. We detail the conservation and electrostatics of these pockets, and investigate their dynamics. Utilizing POVME 2.0 to measure pocket volume, we identified the most open and closed states, gaining insight into gating mechanisms[110]. Subsequently, we utilized dynamic cross correlation (dCC) and the weighted implementation of sub-optimal pathways (WISP) computer program, to identify putative allosteric pathways linking both allosteric pockets to either the cap- or +1-binding pockets[141, 142]. Finally, co-evolution analysis provides evidence leading to a putative novel model of autoregulation, whereby we hypothesize that another domain in LARP1 binds to the DM15 region, acting as an allosteric signal to modulate TOP mRNA binding.

3.2 Materials and Methods

3.2.1 Molecular Dynamics Simulation, Electrostatics, Conservation, and Cancer Mutations

We used the 500ns simulation of the LARP1 4ZC4 B chain from our previous study for the subsequent analyses[98]. We characterized the newly identified pockets via electrostatics, conservation, and prevalence of cancer-associated mutations. PDB2PQR was used to generate the PQR files, and the APBS VMD plugin (with APBS 3.0) was subsequently used to calculate electrostatics and the electrostatics were visualized with VMD[123, 124, 133, 143]. Conservation was projected onto the surface of the 4ZC4:B chain structure, or a conformation from the simulation with the ConSurf server[144, 145]. The COSMIC database was used to check for cancer-associated mutations among the residues within the newly identified pockets[113]. The mutagenesis plugin of the PyMOL Molecular Graphics System (Schrödinger, LLC) was used in order to create a structure with the R824W cancer mutation.

3.2.2 POVME Analysis

We used the POVME 2.0 algorithm as previously described in Cassidy, Lahr et al. 2019 to track the volume of the newly identified pockets[110]. A union of four inclusion spheres of radii 5, 4.5, 3, and 3 Å were used for PQ810. For PR824, a union of three inclusion spheres of radii 5, 3, and 3 Å were used and a contiguous point sphere of radius 2 Å was also used to further refine selection. Finally, for PY880, we chose a union of two inclusion spheres of radii 6 Å and 7 Å; a contiguous point sphere of radius 4 Å was also used. For all pockets, a pocket was considered open if in that frame the volume was at least 30 Å^3. The results were visualized with VMD[133].

3.2.3 Dynamic Cross-Correlation

We utilized dynamic cross-correlation (dCC) to identify putative source residues in the newly identified pockets that may have a putative allosteric relationship with residues in the respective cap- or +1-binding pocket. The software suite MD-TASK was used to calculate a residue NxN dCC matrix based on the α carbons of the protein[141]. No stride was utilized, and default parameters were used. The average magnitude of dCC (for the respective cap- or +1-binding pocket residue) was used as a cutoff for considering residues in the newly identified pockets as potential allosteric sources.

3.2.4 Putative Allosteric Pathway Analysis

We utilized weighted implementation of suboptimal paths (WISP) to identify putative allosteric pathways connecting the previously identified sources with sinks either in the cap- or +1-binding pockets[142]. One thousand paths were identified for each source/sink pair, and a contact distance limit of 4.5 Å was used. The following source/sink pairs were used: R824/R847, Y880/E886 based on the dCC data. Custom python scripts were used to determine the most prevalent residues in the pathways, as well as the prevalence of cancer-associated mutations in the pathways. LARP1 cancer mutations were taken from the COSMIC database[113]. The results were then visualized with VMD[133].

3.2.5 Coevolution Analysis and Shannon Entropy

To identify putative interdomain interactions involving these newly identified pockets, we used coevolution analysis. Shannon Entropy was calculated to gain information on how to design protein constructs of the other LARP1 RNA binding domains, for future in vitro tests of these interdomain interactions. We used the Evol application in the python package ProDy to refine the multiple sequence alignment (MSA) and perform the analyses[146]. The PTHR22792_SF51 Ensembl (Release 96) Family alignment was used for both analyses, and the MSA was refined using human LARP1 "Q6PKG0" and a row occupancy parameter of 0.8[147]. Shannon Entropy was calculated for this refined MSA[146]. A direction information matrix was created using the MSA and rank ordered by z-score. Custom R scripts were used to sort the data by residue, including those lining the newly identified pockets.

3.3 Results

3.3.1 Identification of Three Novel Pockets in the LARP1 DM15 Region

While our previous analysis of the LARP1 simulation focused primarily on the cap- and TSS (+1)-binding pockets, the FTmap results suggest that other druggable pockets exist in the DM15 region. Upon visual reinspection the LARP1 DM15 4ZC4:B simulation and the 12 previously identified cluster centroids for other pockets, we identified a druggable pocket near the N-terminus of the construct on the convex face of LARP1 DM15. It is lined by residues on the α1-α2 bridge (L803, K804, N806, and F808), α2 (Q810, Y813, and R817), α3 (F845, and H849), α3-

α4 bridge (F850, and N851), and α4 (K853, M854, and E857) (Figure 17A, 17D). This shallow

pocket was statically open for most of the trajectory, via visual inspection. To gain insight into the

importance of this pocket we investigated the level of conservation of the lining residues.
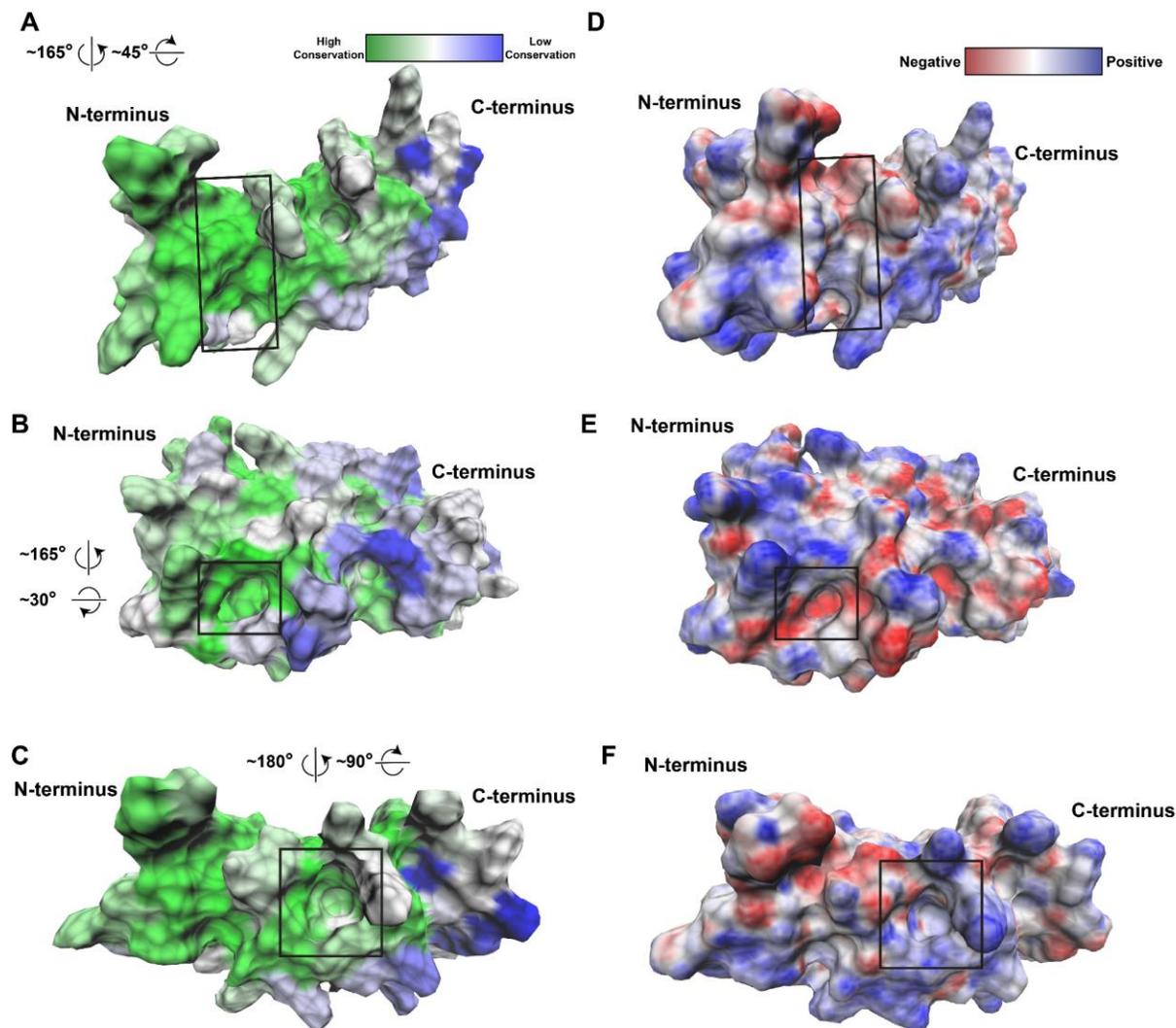
Figure 17. Identification of novel LARP1 DM15 pockets. (A,B,C) Residue conservation projected onto a frame from the LARP1 DM15 simulation for PQ810 (A), PR824 (B), and PY880 (C). Electrostatics projected onto a frame of the LARP1 DM15 simulation for PQ810 (D), PR824 (E), and PY880 (F).

Projection of the consurf calculated conservation onto the LARP1 DM15 structure

indicates that the residues comprising this pocket are mostly well conserved (Figure 17A)[144, 145].

Notably, this conservation does not extend very far into adjacent regions. Electrostatics were also

calculated and indicate a neutral groove throughout the shallow pocket, that is lined by a few charged residues (Figure 17D). From this point forward this pocket will be referred to as PQ810 (Pocket, Q810).

Further inspection revealed two dynamic pockets, and the same conservation data was projected onto frame in which these pockets are open. The first, which has not been identified in any crystal structure, is positioned near the N-terminus of the construct on the convex side of LARP1 DM15 between α2 and α4 and is lined by residues R824, L827, I829, G830, N831, S832, N836, E864, D865, Y870, and Y872 (Figure 17B, 17E). Strikingly, when conservation is mapped to a frame in which the pocket is open, a very well conserved patch is identified (Figure 17B). Again, this high conservation does not extend far into adjacent regions, indicating this is a local structural hotspot of high conservation. Electrostatics calculations indicate that this pocket is negatively charged (Figure 17E). From this point forward this pocket will be referred to as PR824 (Pocket, R824).

The second cryptic pocket is positioned between the α3-α4 and α5-α6 bridges, and lined by residues L846, R847, D848, H849, F850, Y880, G883, L885, K887, K888, Y889, R890, and I893 (Figure 17C, 17F). Conservation was mapped to the structure for a frame in which this pocket is open and indicates that this pocket is conserved, but not as well as PQ810 or PR824 (Figure 17C). Electrostatics mapped to the structure indicate a slight positive charge for this pocket (Figure 17F). From this point forward this pocket will be referred to as PY880.

3.3.2 Pocket Dynamics
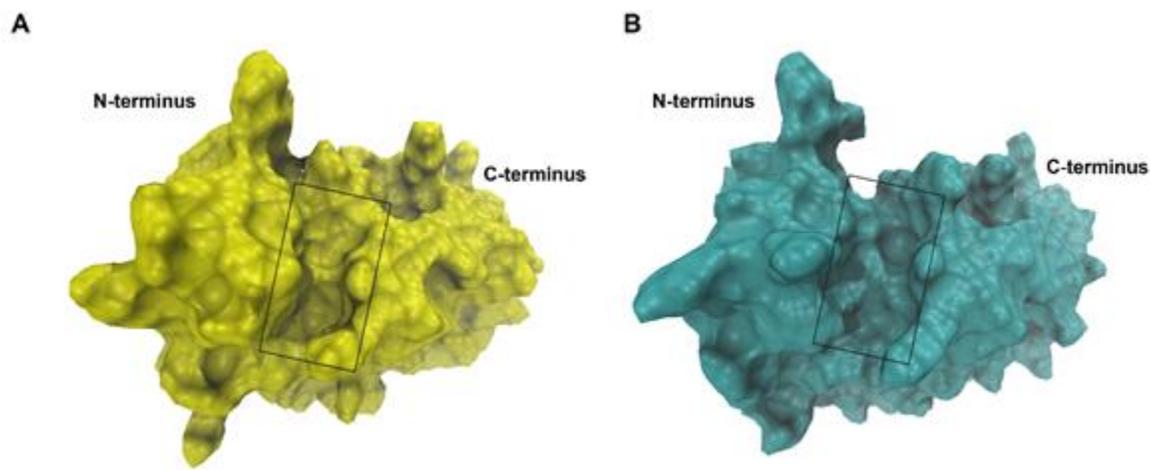
Figure 18. The dynamics of PQ810. (A) A representative frame of the smallest state of PQ810 per POVME 2.0[110]. (B) A representative frame of the largest state of PQ810 per POVME 2.0.

We next investigated the dynamics of these newly identified pockets. We used POVME 2.0 to calculate the pocket volumes of 1,000 equally spaced frames taken from the 4ZC4:B simulation as done previously in Cassidy, Lahr et al. 2019[98, 110]. PQ810 was notably very stable, and even the smallest state was still open (Figure 18).
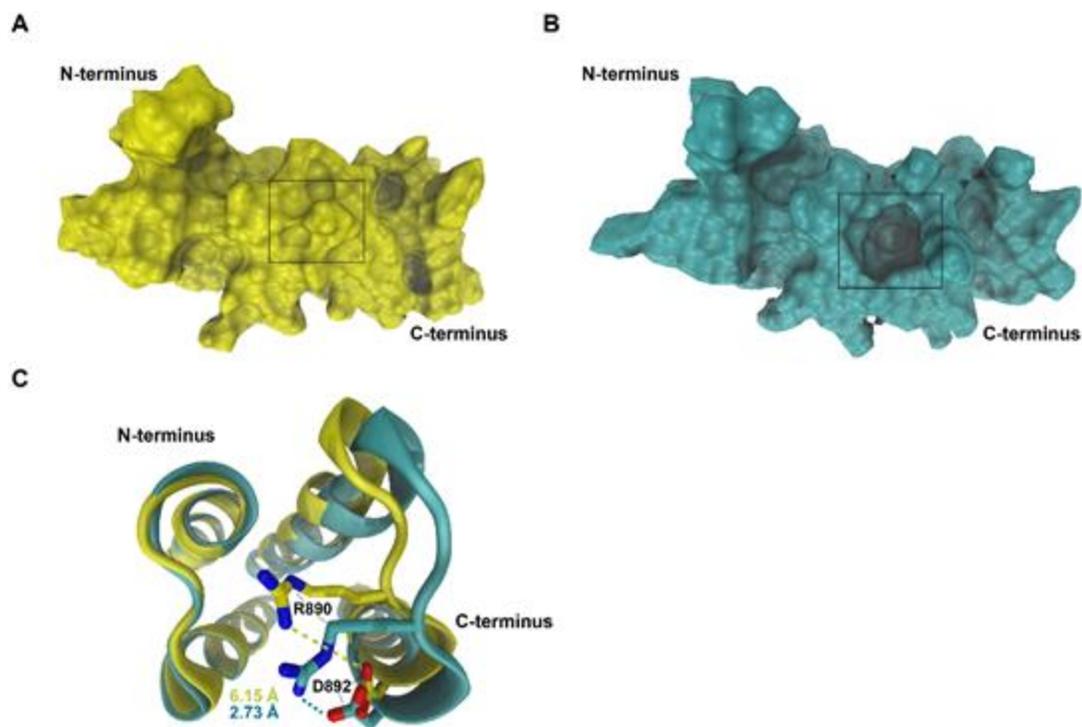
Figure 19. Dynamics of PY880. (A) A representative frame of the smallest state of PY880 per POVME 2.0[110]. (B) A representative frame of the largest state of PY880 per POVME 2.0. (C) A salt bridge between R890 and D892 seemingly gates access to PY880. This interaction is robust in the largest state of PY880, and not present in the smallest state of PY880.

In contrast, PY880 is a dynamic pocket (Figure 19). This pocket was open 81.6% of the analyzed trajectory, with volumes that ranged from 0 to 207.125 Å^3. This analysis revealed that PY880 opened and closed throughout the simulation (Figure 19A, 19B). Key residues and transient interactions seem to be gating the pocket. A salt bridge between R890 and D892 appears to sequester R890 away by "sliding the top off the jar", allowing for opening of the pocket (Figure 19C). The R890 amino and D892 carboxyl groups are within close proximity in the largest state of PY880 (2.73 Å), and distant in the smallest state (6.15 Å) (Figure 19C).
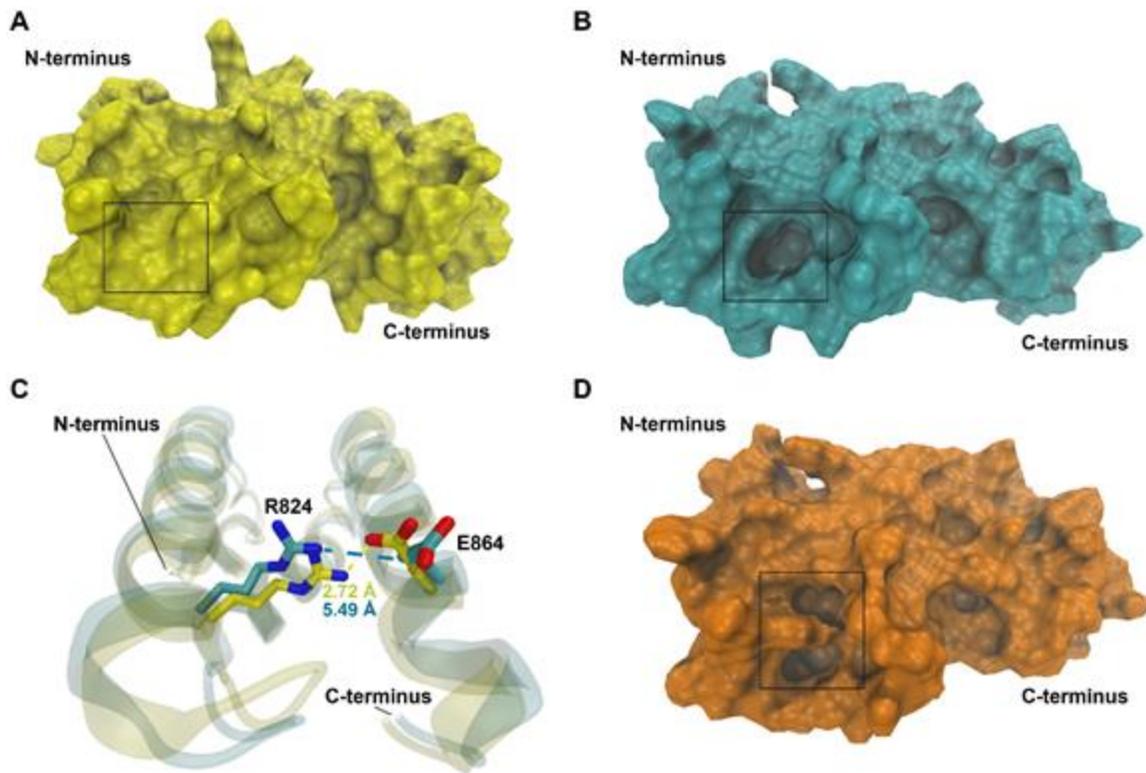
Figure 20. Dynamics of PR824. (A) A representative frame of the smallest state of PR824 per POVME 2.0[110].
(B) A representative frame of the largest state of PR824 per POVME 2.0. (C) A salt bridge between R824 and
E864 seemingly gates access to PR824, in conjunction with the clamshell like closing of the α2- α3 bridge. This
interaction is robust in the largest state of PR824, and not present in the smallest state. (D) Cancer-associated
mutation R824W was created in silico, to see its effects on PR824, and this mutation occludes access to the
pocket.

PR824 was even more rarely open, and is only open 1.6% of the analyzed trajectory. The

pocket volumes ranged from 0 to 66.25 Å^3. Upon inspection of the most open and closed states of

the pocket (Figure 20A, 20B), key residues and transient interactions were found to gate the pocket.

Notably, the position of R824 as well as the α2-α3 bridge seemingly determine the openness of the

pocket. R824 and the α2-α3 bridge converge in a clam-shell-like fashion in the closed state,

occluding the pocket. This closed position of R824 is stabilized by the formation of a salt-bridge

with E864 on α4. In the smallest closed state of PR824, this interaction is robust and the distance

between the R824 amino and E886 carboxyl groups is 2.72 Å. However, in the largest open state, this interaction is not present and the respective distance is 5.49 Å (Figure 20C).

Due to the importance of R824 in this gating mechanism, we searched for any reported cancer-associated mutations in this region. Strikingly, R824 is not only tied for the residue in DM15 with the most missense cancer-associated mutations, but in all of LARP1 at the time of analysis. This mutation was found in breast cancer and is also predicted to be pathogenic, per the COSMIC-provided FATHMM score of 0.98[113]. We created the most common missense mutation of R824 *in silico*, R824W, in order to see its effect on this newly discovered cryptic pocket. Even when this mutation was made within the most open state of the PR824 pocket, the increase in surface area of the tryptophan occludes the opening of the pocket (Figure 20D).

3.3.3 Putative Allosteric Pathways
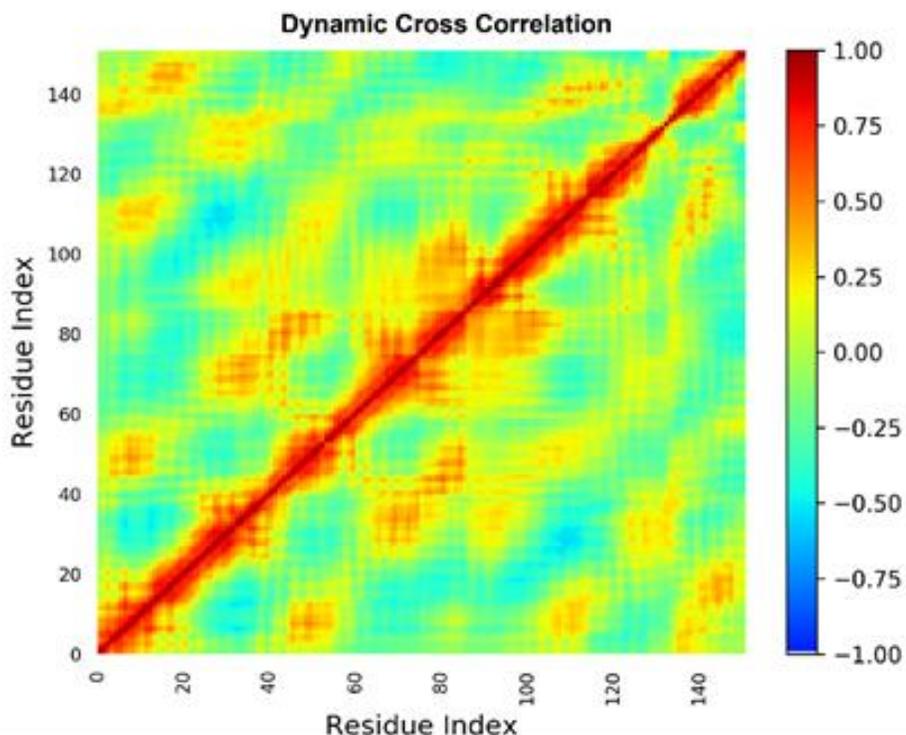
Figure 21. Dynamic Cross correlation for LARP1 DM15 4ZC4:B simulation. The dynamic cross correlation heatmap for LARP1 DM15 represents the correlation of motion between two residues (-1 to 1). Numbering from the simulation is used for residue index (residue index + 795 = isoform 2 numbering).

Although it is possible that these pockets have novel biological functions, we next investigated whether the PY880 and PR824 pockets are allosterically coupled to either the cap- or +1-pocket. We first used MDtask to calculate dynamic cross correlation, which scores the correlated motions of residues, to find putative source residues in the PY880 or PR824 pockets for sink residues in the cap- or +1-pocket (Figure 21)[141]. PY880 is in close proximity to the cap-binding pocket and several residues lining PY880 had high levels of correlation with E886, the cap-binding pocket residue that hydrogen bonds with the Watson-Crick face of the 7-methylguanosine cap of TOP mRNAs. Y880 was chosen as the source residue for PY880 because

76

the side chain is directly within the pocket, and positioned such that the Oη could bind with a substrate. We then utilized weighted implementation of suboptimal paths (WISP) to determine the top 1,000 putative allosteric pathways[142].The top 10 pathways included several short paths of similar length and residue composition. There was a total of 5 unique residues in the top 10 paths, outside of the source and sink. The most optimal path was Y880:G884:E886 (Figure 22A). Custom python scripts were used to identify the most common residues in the top 1000 putative allosteric pathways, as well as to identify cancer-associated mutations present in these pathways (Figure 22B). Notably, Y881, G884 and L885 were in more than 50% of the paths. Although residues with known cancer-associated mutations, which could disrupt allosteric communication, were present in a portion of the top 1000 paths, none were in more than 30% of them.
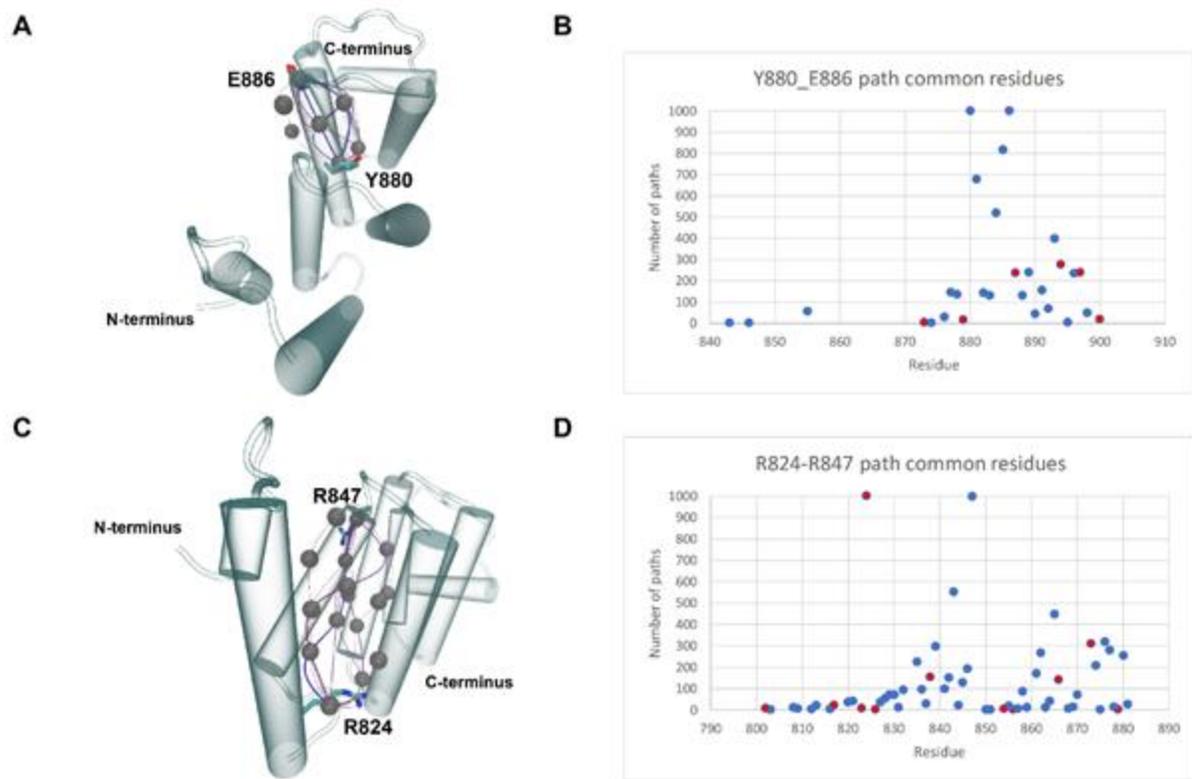
Figure 22. Putative allosteric pathways couple PY880 and PR824 with the TOP mRNA binding pockets. (A) The top 10 allosteric pathways between Y880 (PY880) and E886, as determined by WISP[142]. (B) The most prevalent residues in the top 1000 allosteric pathways between Y880 and E886 as determined by WISP. Residues with known cancer mutations are in red. (C) The top 10 allosteric pathways between R824 (PR824) and R847, as determined by WISP. (D) The most prevalent residues in the top 1000 allosteric pathways between R824 and R847 as determined by WISP. Residues with known cancer mutations are in red.

We also used dynamic cross correlation to identify a putative source residue in PR824. There were many residues with a similar magnitude of cross correlation with R847 in the TSS-binding pocket, which hydrogen bonds with the Watson-Crick face of the invariant +1C of TOP mRNA. R824 was chosen as the putative source due to its biological relevance as the residue with the most prevalent missense cancer mutations in LARP1. WISP was again used to calculate the top 1,000 putative allosteric pathways, and the top 10 paths included several short paths of similar length but that differed more in residue composition than PY880. (Figure 22C). There was a total

of 14 unique residues in the top 10 paths, outside of the source and sink. The most optimal path was R824:M835:F839:S843:R847. We again used custom python scripts to identify the most common residues and identify cancer mutants (Figure 22D). There was a much larger spread of residue prevalence in the top 1000 paths compared to PY880. Only one residue, S843, was present in more than 50% of the paths. One residue, G873, with a known cancer-associated mutation, was in greater than 30% of the paths.

3.3.4 Interdomain Interactions May Allosterically Regulate LARP1 DM15 TOP mRNA Binding

While the identified putative allosteric pathways could transduce a signal from these pockets to the TOP mRNA binding pockets, substrate binding event that initiates this signal is unknown. We hypothesized that interdomain interactions may play this role. To test this hypothesis, we utilized a coevolution analysis.

Table 3. Z-score normalized coevolution data highlighting the putative interdomain interacting residues between the LARP1 La-Module and DM15 region.

La-Module residue	DM15 residue	Coevolution Z-score
636	852	10.16
637	852	10.14
626	809	6.19
626	810	6.19
622	809	5.85
622	810	5.85

616	862	5.63
625	817	5.04
625	822	5.04
623	859	4.73
619	829	4.59
625	818	3.76
637	859	3.67
626	852	3.57
611	870	3.46
626	817	3.4
626	822	3.4
616	852	3.37
622	852	3.34
621	817	3.33
621	822	3.33
619	859	3.31
625	809	3.29
625	810	3.29
636	859	3.19
616	863	3.14
626	817	3.09
626	822	3.09
627	817	3.09

| 627 | 822 | 3.09 |
| 608 | 807 | 3.08 |

We created a direct information matrix using the coevol package of prody to evaluate direct couplings of residues throughout LARP1[146]. This indicated a stretch of residues in the La-Module of LARP1 (608-637) that repeatedly coupled with residues in and around PQ810 and PR824 (Table 2) (Figure 23). PQ810 could be allosterically tied to the cap- or +1-binding pocket on its own, but it is also very close in proximity to PR824.
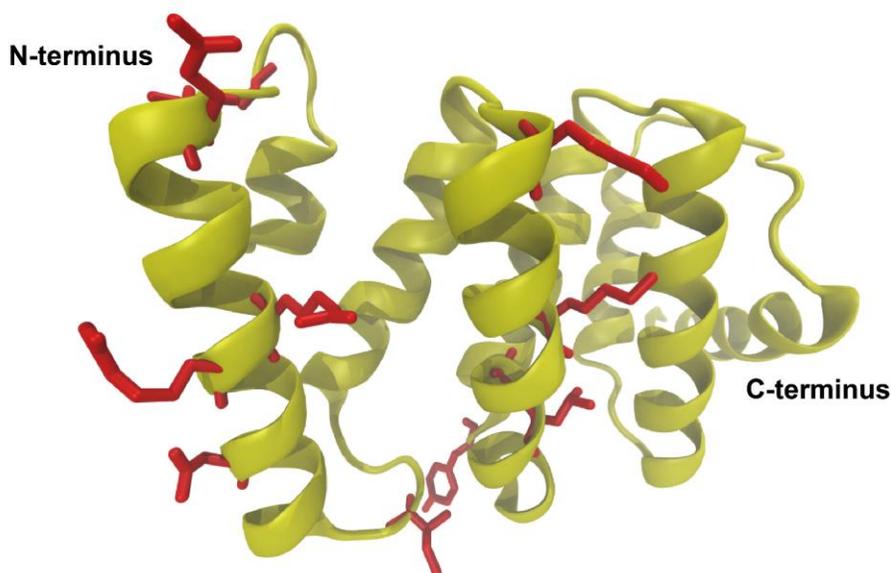
Figure 23. The putative LARP1 DM15 interdomain binding interface. LARP1 DM15 with residues that putatively coevoled with 608-637 in the La-Module highlighted in red. See also Table 2.

Although highly speculative, it is possible that the La-module binding to PQ810 anchors the interaction, and causes PR824 to be more likely in an open state. The coevolution analysis also revealed a residue within the La-module K495, that is directly coupled with D865 in PR824. K495 could then thread into PR824 after the 608-637 region of the La-module anchors the La-module/DM15 interaction, allosterically regulating the TSS binding pocket.
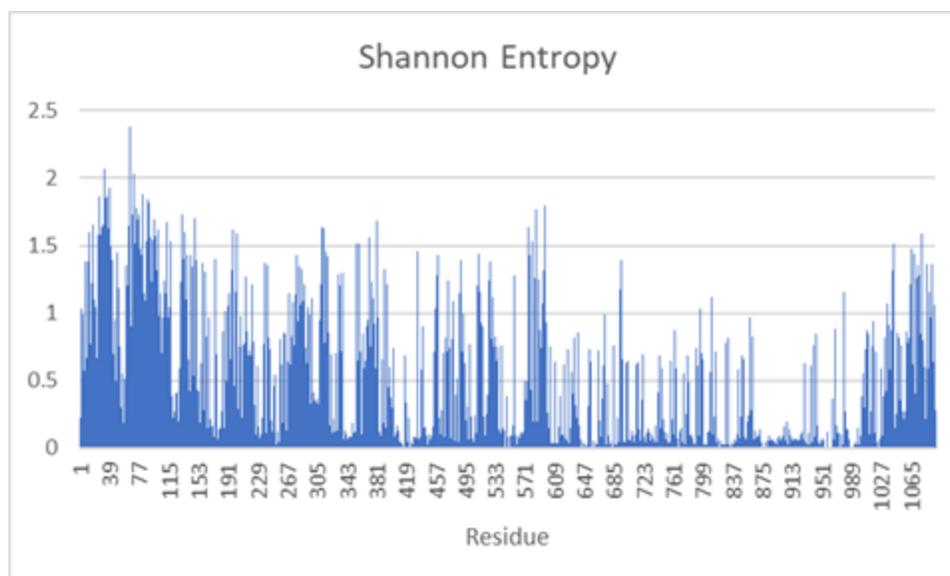
Figure 24. Shannon Entropy analysis of LARP1. Shannon (Sequence) entropy has been shown to correlate positively with structural flexibility, and so was calculated to help guide future LARP1 construct design (Isoform 1 numbering)[146]. Constructs of the La-module can be used in order to test if the La-module interacts with the DM15 region as hypothesized.

Currently there are no structures of the La-module of LARP1, and so we calculated Shannon Entropy using the coevol package of prody to help guide future construct design (Figure 24)[146]. Shannon Entropy can be used to predict structure flexibility, and could be used to determine construct boundaries by identifying hinge regions.

3.4 Discussion

LARP1 plays a central role in regulating ribosome biogenesis[2]. Unsurprisingly, high levels of LARP1 are correlated with multiple cancers, however so far LARP1 has not been targeted therapeutically[85-87]. Two main problems arise when considering LARP1 as a drug target. First, targeting the more voluminous cap-binding pocket of LARP1 DM15 may lead to compounds with

low specificity due to the existence of other crucial cap-binding proteins (e.g. eIF4E). Second, the dominant role of LARP1 in regulating TOP mRNA translation (repression or stimulation) is not known.

Identification of allosteric pockets coupled to the cap- or +1-binding pockets would aid in solving both these problems. Targeting allosteric pockets could likely increase compound specificity by targeting more unique chemical environments than the cap-binding pocket. Allosteric pockets that could be exploited to enhance TOP mRNA binding would offer an alternative strategy to occluding RNA binding, and could be necessary to properly therapeutically target LARP1 in cancer. Characterization of allosteric pockets would also shed light on how the binding of the LARP1 DM15 region to TOP mRNA \is regulated, further explaining its role in translational regulation.

3.4.1 PY880: A Transient Pocket Gated by a "Lid" Dynamic

PY880 is present in previous crystal structures, but our simulations show how transient and potentially voluminous it is. Specifically, we show that a key interaction between R890 and D892 gates access to Y880. R890 acts as a "lid" occluding access, and, in the conformation with the largest pocket volume, this "lid" slides off via a robust interaction with D892. The largest pocket conformation also provides more insight into the types of endogenous ligands that could potentially bind, however unlike for PQ810 and PR824, we do not yet have evidence for a specific ligand. The hydroxyl group of Y880 is accessible at the base of the pocket and is a good hydrogen bonding partner for potential ligands. We provide evidence suggesting PY880 is allosterically coupled to E886 in the cap-binding pocket, and perturbation at this site could alter the ability of LARP1 DM15 to bind TOP mRNA.

3.4.2 PR824: A Cryptic Pocket Gated by a "Clam-Shell" Dynamic

There are currently no crystal structures in which PR824 is present, and it is rarely open throughout our simulation. We show how a key interaction (R824-E864) along with the movement of the α2-α3 loop gates this pocket. These dynamics are similar to a "clam-shell" opening and closing, with R824 and the α2-α3 loop as the two halves of the shell. In the orientation depicted, R824 acts as the top half closing down, and the α2-α3 loop acts as the bottom half coming up to ultimately occlude the pocket. R824 was also found to be tied for the residue in all of LARP1 with the greatest number of known cancer missense mutations[113]. We show evidence suggesting PR824 is allosterically coupled with R847 in the +1-binding pocket, and perturbation at this site could alter the ability for LARP1 DM15 to bind TOP mRNA.

3.4.3 PQ810 and PR824: Novel Pockets Potentially Involved in LARP1 Autoregulation

The two pockets identified near the N-terminus of LARP1 DM15, PQ810 and PR824, differ substantially in their dynamics. The static and dynamic natures of PQ810 and PR824, respectively, indicate seemingly different endogenous functions. However, our coevolution analysis indicates both regions may bind to portions of the La-module of LARP1. PQ810 and PR824 may be involved in a set of interactions contributing to a novel and previously unpredicted autoregulatory mechanism of LARP1/TOP mRNA binding, via influencing +1-binding pocket dynamics. It is important to note that the Ensembl family alignment of LARP1 sequences that was used is not fully comprehensive, and duplicates may exist. For this reason, this analysis was used as a starting point as a search for a potential binding partner. The more static PQ810 could first bind to the La-module (residues 608-632), due to being in an open state a considerably greater

amount of time. This interaction may cause an allosteric effect in LARP1 DM15, potentially altering the dynamics of the cap- or +1-binding pockets. These interactions could also influence the energy landscape of adjacent PR824, shifting it to one in which the open state of PR824 is more probable. Once this "anchoring" event takes place at PQ810, the now more probabilistically open PR824 could interact with the La-module, via K495 binding to PR824. This could in turn induce a signal from PR824 to the +1-binding pocket, which we show are potentially allosterically coupled. Notably, a mutation observed in cancer, R824W, the residue with the greatest number of missense cancer mutations in all of LARP1, occludes access to PR824. This mutant could be insensitive to the allosteric changes induced by the La-module binding to PR824. Alternatively, this mutation could induce its own allosteric effect. Targeting either of these pockets with compounds could also turn off this putative autoregulation, making LARP1 DM15 either unable to interact with or insensitive to La-module binding. If true, this approach could be used in cells to more fully characterize the translational regulatory role of LARP1.

3.4.4 Conclusions

We characterized three pockets in the LARP1 DM15 region with either indirect or direct potential allosteric couplings to the cap- or +1-binding pockets. These pockets provide new unique chemical environments to investigate both in their roles in the regulation of TOP mRNA binding and pharmacologically. Depending on the mechanistic characterization of these allosteric relationships, compounds targeting these pockets may enhance or prevent TOP mRNA binding and subsequently alter the impact of LARP1 on TOP mRNA translation. This flexibility in strategy would allow for a more promising outlook on therapeutically targeting LARP1 DM15 in cancer. These compounds could also be utilized as probes to help define the nature of LARP1 TOP mRNA

translational regulation, as well as characterizing the potential allosteric network governing this modulation.

4.0 Allosteric Cancer-Associated LARP1 Mutation R824W Alters +1-Binding Pocket Dynamics of LARP1 DM15

4.1 Introduction

Protein synthesis regulation is a key part of cell growth and proliferation. Dysregulation of this crucial process can contribute to disease by providing the resources needed for rampant cell division, a key feature of most cancers[32, 50, 53, 56, 58, 63, 148]. mTORC1, a key piece of this regulatory machinery, has been targeted therapeutically in the treatment of kidney cancer[10, 31-33, 40, 50, 56, 63, 148, 149]. However, due to the multifaceted nature of mTORC1's influence in the cell, many side effects have been recorded[64, 65]. LARP1 is another node in this regulatory process. It resides downstream of mTORC1 and is known to regulate a class of transcripts that encodes all ribosomal proteins[2]. This group of transcripts is known as TOP mRNAs because they have a terminal oligopyrimidine motif at their 5' ends[29]. Furthermore, the overexpression of LARP1 has been observed in several epithelial cancers, including lung, liver, ovarian, and cervical cancer[71, 85-88, 93, 95]. Because of these correlations and the role of LARP1 in the mTORC1 cascade, it is vital to further dissect the role of LARP1 in translation regulation.

There is seemingly contradictory evidence suggesting a stimulatory role and a repressive role of LARP1 in TOP mRNA translation[2, 81-84]. In addition, LARP1 has been proposed to stabilize TOP transcripts—a function that can be stimulatory or neutral in the context of TOP mRNAs[2, 82, 83]. Further, recent work has shown that post translational modifications cause a change in the functions of LARP1, including its ability to bind TOP mRNAs[35, 81]. mTORC1 phosphorylation of LARP1 was shown to hinder or enhance the ability of the conserved LARP1 C-terminal cap-

binding region, the DM15 region, to bind TOP mRNAs[35]. Allosteric regulation could explain these seemingly contradictory data. In Chapter 3, I detailed a newly identified pocket, PR824, that has a potential allosteric relationship with the +1-binding pocket. I also showed how a LARP1 breast cancer-associated mutation could alter this pocket.

Given an allosteric relationship between PR824 and the +1-binding pocket, we hypothesize that PR824 could be targeted pharmacologically to alter the role of the LARP1 DM15 region in TOP mRNA translation regulation. Indeed, characterizing the effects of the R824W mutation may also shed light on how this mutant could have a pathogenic effect in breast cancer.

Here I simulate LARP1 DM15 4ZC4:B R824W for 2 µs to investigate the effect of this mutation and the allosteric relationship of PR824 with the+1-binding pocket. The previously analyzed 500 ns LARP1 DM15 4ZC4:B simulation was extended to ~2µs to more fully capture the representative conformations that the molecule could have in a simulated aqueous environment[98]. This extended simulation provides new insight into the dynamics of the cap- and +1-binding pockets of LARP1 DM15, with the identification of an alternative +1-binding pocket. I detail how the +1-binding pocket dynamics change in the context of the R824W mutation, as compared with the wild-type protein, and how this mutation increases the probability of the +1-binding pocket obtaining an alternative conformation, which could alter affinity for TOP mRNAs or switch the binding partner to 3' UTRs. Lastly, I identify the residue-residue contact and community differences that lead to the altered dynamics of the cap- and +1-binding pocket by using the difference contact network analysis (dCNA) method.

4.2 Materials and Methods

4.2.1 Model Building and Parameterization

The LARP1 DM15 4ZC4:B R824W mutant from Chapter 3 was used as the starting structure for a simulation. Model Building and parameterization was carried out as in Cassidy, Lahr et al. 2019[98]. VMD was used for all structure visualizations[133].

4.2.2 Minimization, Equilibration, and Production Simulation

Minimization, equilibration and production simulations were carried out as previously described in Cassidy, Lahr et al. 2019[98]. The LARP1 DM15 4ZC4:B simulation from Cassidy, Lahr et al. 2019 was extended to ~2 µs (2033.36 ns), and the first 60 ns was again removed out of concern the simulation had not yet equilibrated[98]. All analyses except for the RMS distance (RMSD), use this truncated LARP1 DM15 4ZC4:B simulation (1973.36 ns). LARP1 DM15 4ZC4:B R824W was simulated for ~2 µs (2000.88 ns).

4.2.3 RMSD, RMSF

We extracted frames every 10 ps and aligned each trajectory to its respective first frame by Cα using MDAnalysis 0.20.1[132]. We calculated RMS distance and RMSF as previously described in Cassidy, Lahr et al. 2019.[98]

4.2.4 POVME Analysis

POVME 2.2 was used to calculate pocket volumes as described in Cassidy, Lahr et al. 2019 with POVME 2.0, but every frame was used for both the extended LARP1 DM15 4ZC4:B simulation (1973.36 ns) and the LARP1 DM15 4ZC4:B R824W simulation (2000.88 ns)[98, 110]. The two trajectories were aligned to the first frame of the truncated LARP1 DM15 4ZC4:B simulation (1973.36 ns) via the Cα of residues 796-930. This allowed for consistency between the point spheres chosen to measure pocket volume. Residues at the C-terminus were excluded from the alignment selection due to their high flexibility differences between simulations, which is likely an artifact of the truncated construct.

A single inclusion sphere with a radius of 9 Å was used to measure the volume of the cap-binding pocket. To further refine the selection, a contiguous sphere with a radius of 3 Å was used for the cap-binding pocket. Four inclusion spheres with radii of 3 Å, 5 Å, 5.5 Å, and 7 Å were used to measure the volume of the +1-binding pocket. Three inclusion spheres with radii of 3 Å, 3 Å, 5 Å were used to measure the volume of PR824. To further refine the selection, a contiguous sphere with a radius of 2 Å was used.

4.2.5 Probability Distributions

Bond lengths were calculated with VMD 1.9.3 for each frame of the respective trajectories[133]. The largest and smallest bins of the probability distribution were chosen based off the observed minimum and maximum values for the respective metric. A frame was assigned to a bin if the metric value was greater than or equal to the bin value, but less than the next higher value bin.

4.2.6 Residue-Residue Contact Analysis

The difference residue-residue contact analysis and difference community analysis were performed using the difference contact analysis (dCNA) method and tool described in Yao et al. 2018[150]. The default values for all model parameters were used.

4.2.7 Clustering Analysis

The AmberTools20 cpptraj 4.25.6 implementation of the hierarchical agglomerative algorithm was used for clustering analysis[151-153]. A sieve value of 10, an epsilon of 2.5, and average-linking were used. The residues (non hydrogen atoms) selected for the dme mask were: R840, F844, R847, R879, S882, Y883, E886, and Y922.

4.3 Results

Our previous analysis of the 500 ns simulation of LARP1 DM15 4ZC4:B shed light on the dynamics of the cap- and +1-binding pockets[98]. However, I wanted to further investigate the allosteric effects of the R824W cancer-associated mutation. This residue lines pocket PR824, identified in Chapter 3, which has a potentially allosteric link with the +1-binding pocket. I hypothesized that the R824W mutation would alter the dynamics of the +1-binding pocket. I simulated LARP1 DM15 4ZC4:B R824W (see methods in Chapter 3) for 2 µs to investigate this hypothesis. This simulation from here on is referred to as the R824W simulation. I also extended

the 500 ns LARP1 DM15 4ZC4:B simulation described in Chapter 2 to 2 µs to more fully investigate the full representative conformational ensemble[98]. This simulation is from here on referred to as the WT simulation in this chapter.

4.3.1 Simulation Equilibration and PR824 Pocket Volume
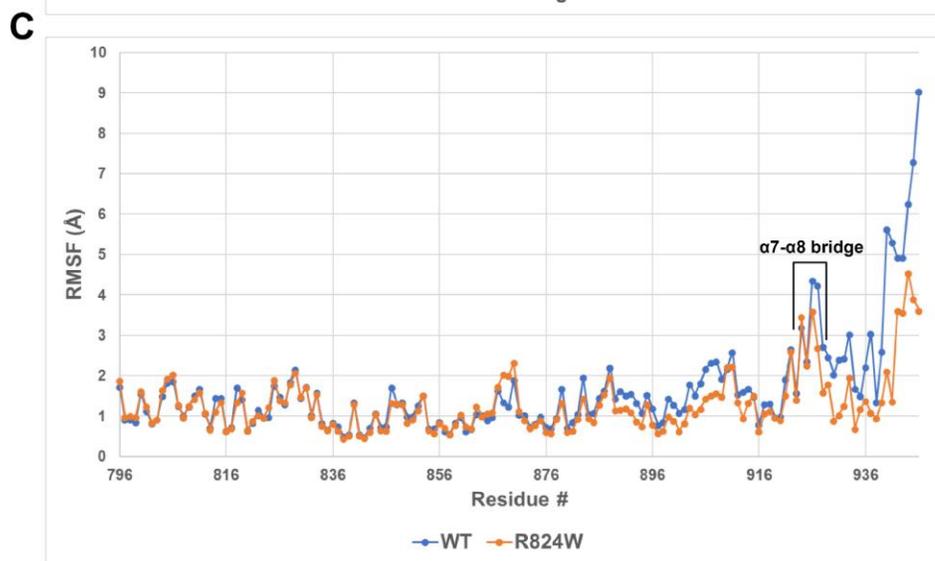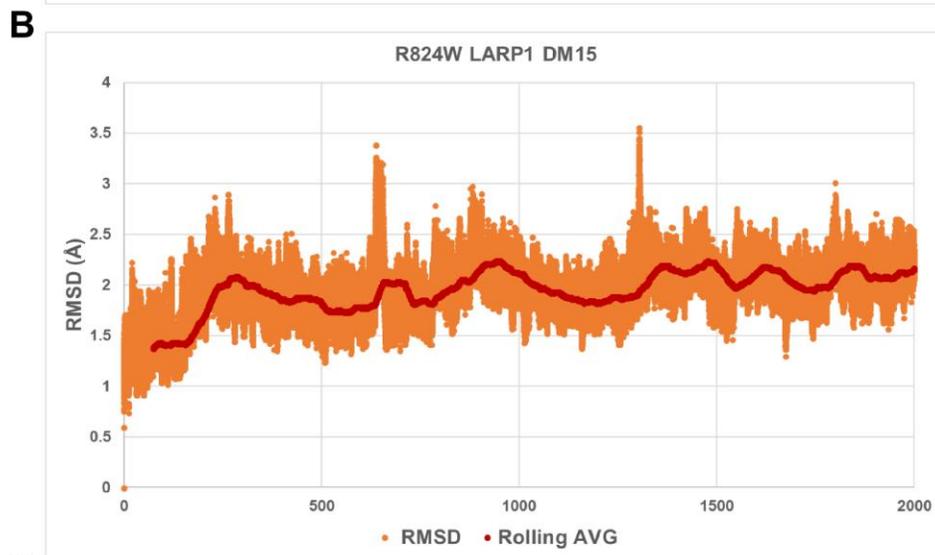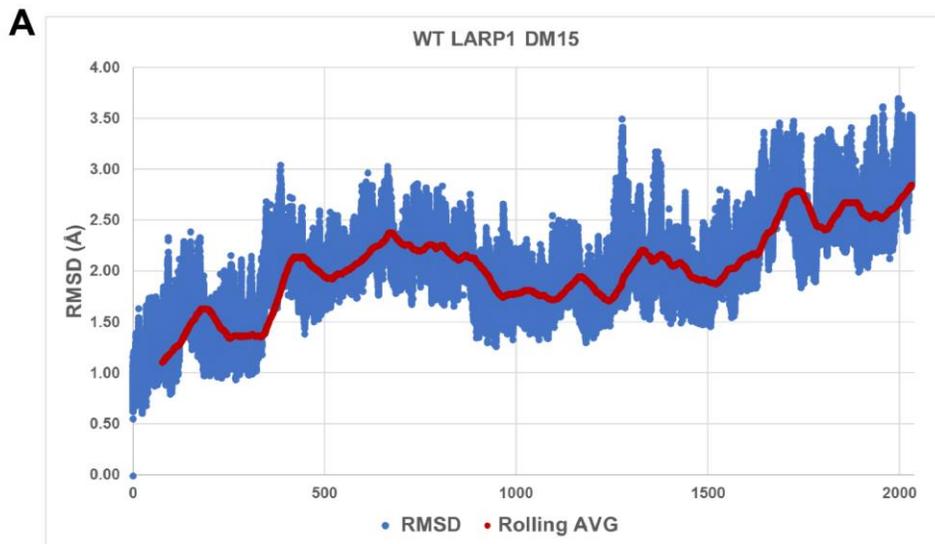
Figure 25. WT and R824W RMSD/RMSF analysis. (A) RMSD of each frame of the WT simulation in reference to the first frame. Raw data in blue, rolling average in orange. (B) RMSD of each frame of the R824W simulation in reference to the first frame. Raw data in blue, rolling average in orange. (C) The RMSF of the WT and R824W simulations, calculated via center-of-geometry per residue. WT simulation values in blue, R824W simulation values in orange.

For both simulations, I calculated the C_α root mean square distance (RMSD), for each frame in reference to the first frame of the respective simulation, to ensure both simulations had equilibrated (Figure 25). The first 60 ns of the LARP1 DM15 4ZC4:B simulation were discarded, out of caution that the simulation had not equilibrated. The R824W simulation had stabilized by the beginning of the production run, by visual inspection of the RMSD, so no frames were removed.

We then calculated the root mean square flucutation (RMSF) per residue via center-of-geometry (Figure 25). The most striking difference was in the most C-terminal residues (residues 930-946), but since this construct design was truncated artifically (based on sequence conservation) not much can be inferred from this data. Overall trends of flexibility were consistent between the WT and R824W simulation. Notably, the α7-α8 bridge, S923-L928, is a peak in flexibility for both simulations. This trend was seen previously in the 500 ns simulation detailed in Chapter 2[98]. Other slight differences in RMSF were observed, but none were larger than the differences between the previous 4ZC4:B and 4ZC4:C simulations. Since these slight differences were less than those observed between the two 500 ns WT simulations, it is difficult to discern if these differences were due to the R824W mutation and functionally meaningful or just background noise. Thus, we moved onto other analyses to investigate any differences between the WT and R824W simulations.
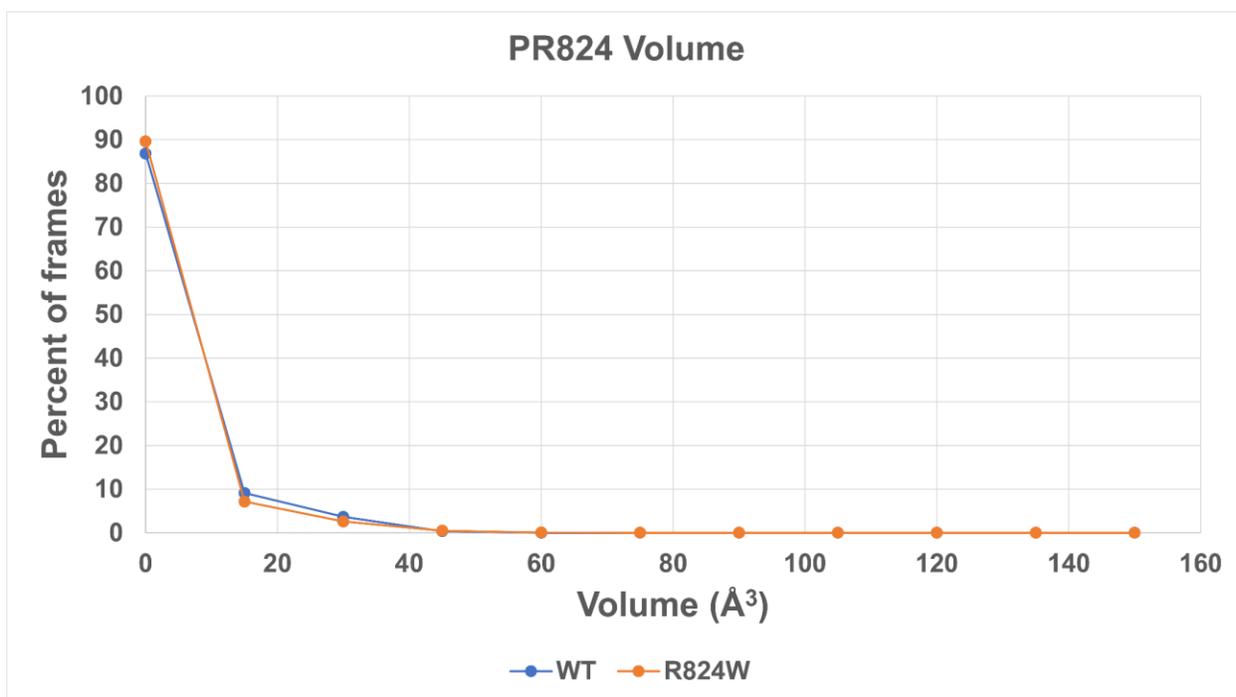
Figure 26. PR824 pocket volume. Probability distribution of the PR824 pocket volumes for both the WT and R824W simulations. POVME 2.2 was used to calculate PR824 pocket volume[110].

I hypothesized that the R824W mutation would decrease the volume of PR824 throughout the simulation, due to the larger side chain of tryptophan. To analyze this, I used POVME 2.2 to measure the PR824 pocket volume over the course of both the WT and R824W 2 µs simulations[110]. The average PR824 pocket volume was smaller in the R824W simulation than the WT simulation, 3.17 $Å^3$ and 3.84 $Å^3$ respectively. Conversely, the R824W simulation had a larger maximum PR824 volume (WT: 83.1 $Å^3$, R824W: 145 $Å^3$). In both simulations, the PR824 volume was less than 15 $Å^3$ for a large percentage of frames (WT: 86.8%, R824W: 89.6%) (Figure 26). The R824W simulation showed only a slight increase in the percentage of frames where PR824 had a volume of less than 15 $Å^3$. The decreases in the PR824 pocket volume in the R824W simulation were small, and this could be because a large percentage of frames already have small pocket volumes in the WT simulation.

4.3.2 Clustering Analysis Reveals Novel +1-Binding Pocket Conformation
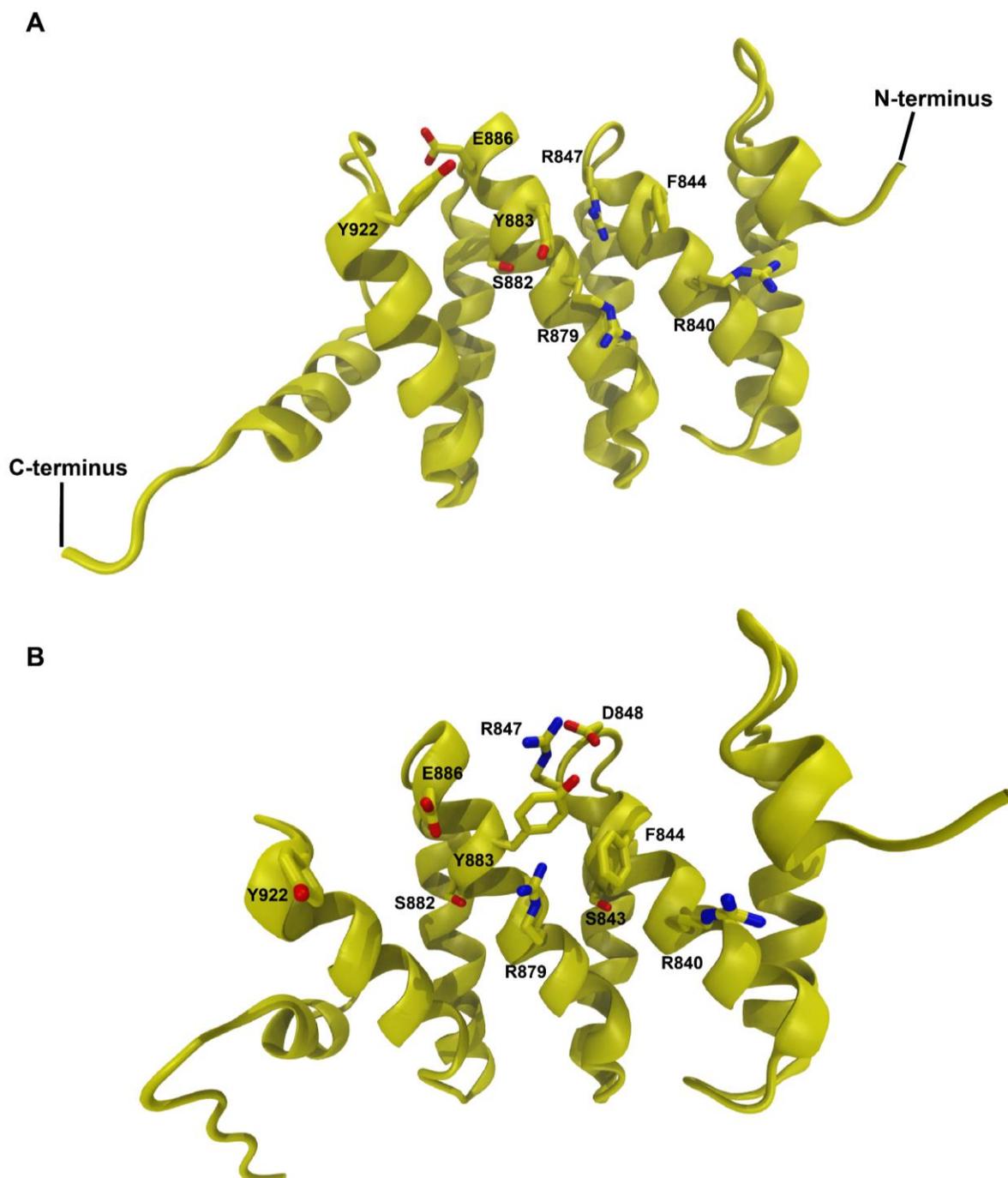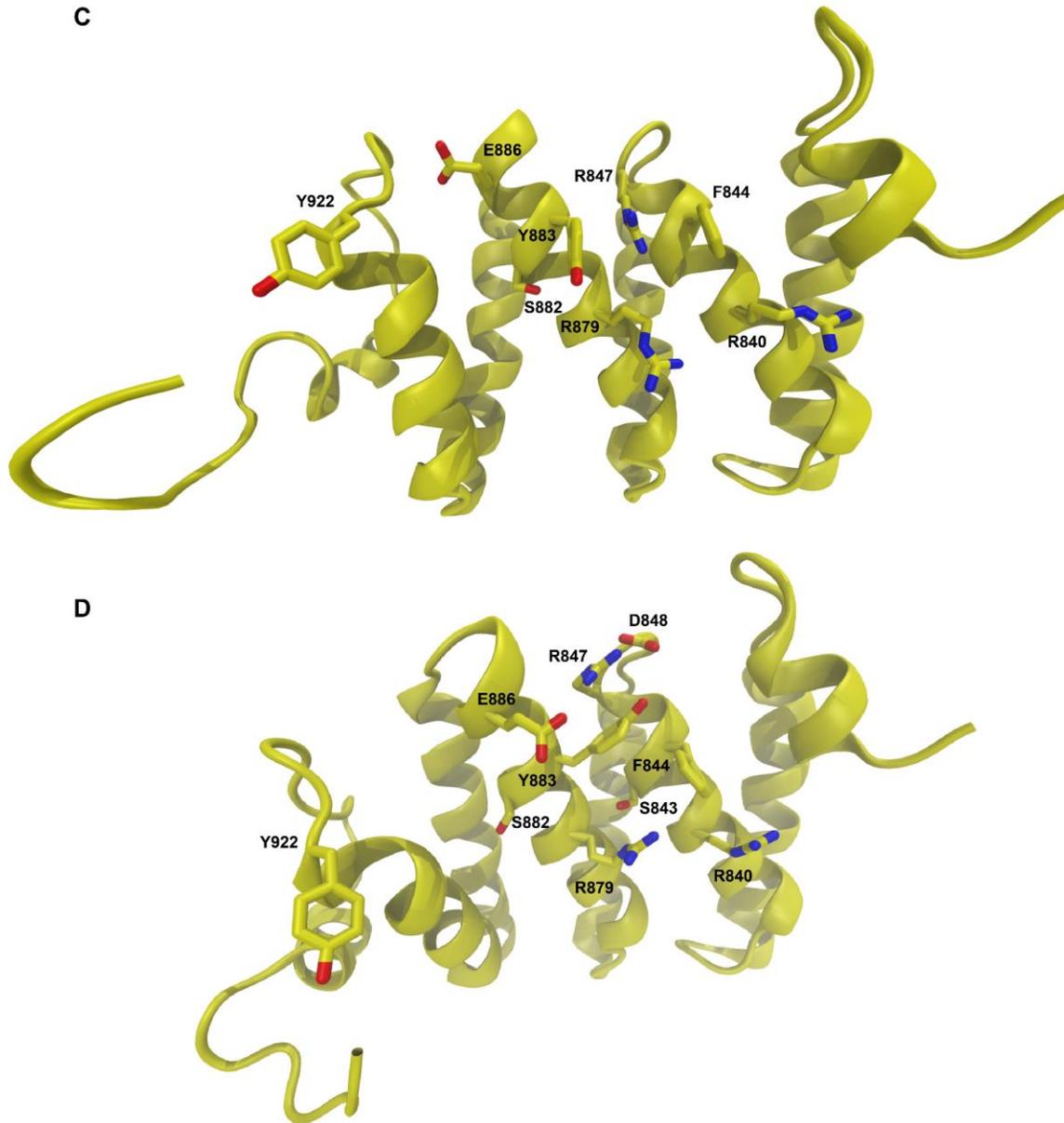
Figure 27. Clustering analysis of the WT simulation reveals a novel +1-binding pocket conformation. (A,B,C,D) Representative frames of the four clusters identified from the WT simulation with the hierarchical agglomerative clustering algorithm[151-153]. (A) Y922 occludes the cap-binding pocket. (B) R847 flips out of the +1-binding pocket and interacts with D848. Y883 flips into the +1-binding pocket. These changes in residue positions create an alternate +1-binding pocket. (C) An open cap-binding pocket, and canonical +1-binding pocket. (D) A similar alternative +1-binding pocket as in B.

The extended 2μs WT simulation offers a more complete look at the microstates accessible to WT LARP1 DM15 over the previous 500 ns simulation. To investigate all relevant microstates, I used the hierarchical agglomerative clustering algorithm to extract distinct microstates from the WT simulation; these microstates were also used for comparison with the R824W simulation to detect any changes in the cap- and +1-binding pockets[151-153]. This "bottom-up" hierarchical approach starts with each frame as its own cluster, and they are merged based on the distance-RMSD. With the epsilon set to 2.5 the minimum distance between clusters is 2.5. Four clusters were identified and representative frames (those closest to each respective cluster centroid) were extracted from the WT simulation (Figure 27). One of these clusters captured the previously identified closed state of the cap-binding pocket, wherein Y922 is flipped into the cap-binding pocket and occludes access (Figure 27A). This microstate also has key residues of the +1-binding pocket (Y883, R847) in positions similar to those that they occupy in both crystal structures and in our previous 500 ns simulations. This conformation of the +1-binding pocket will be referred to as the "canonical +1-binding pocket" for the rest of the chapter.

Shockingly, a key +1-binding pocket recognition residue, R847, was flipped out of the pocket in two of these microstates (Figure 27B, 27D). In these microstates, R847 interacts with D848 via a salt bridge, and Y883 flips into the +1-binding pocket potentially interacting with R847 via a cation-π interaction. Both of these interactions would stabilize this microstate, leading to an alternative +1-binding pocket. For the rest of the chapter, a microstate is deemed to have an "alternative +1-binding pocket" when R847 is observed to interact with D848, and/or Y883 is flipped into the pocket. Both of these two microstates also had open cap-binding pockets. The other microstate captured an open cap- and canonical +1-binding pocket (Figure 27C).
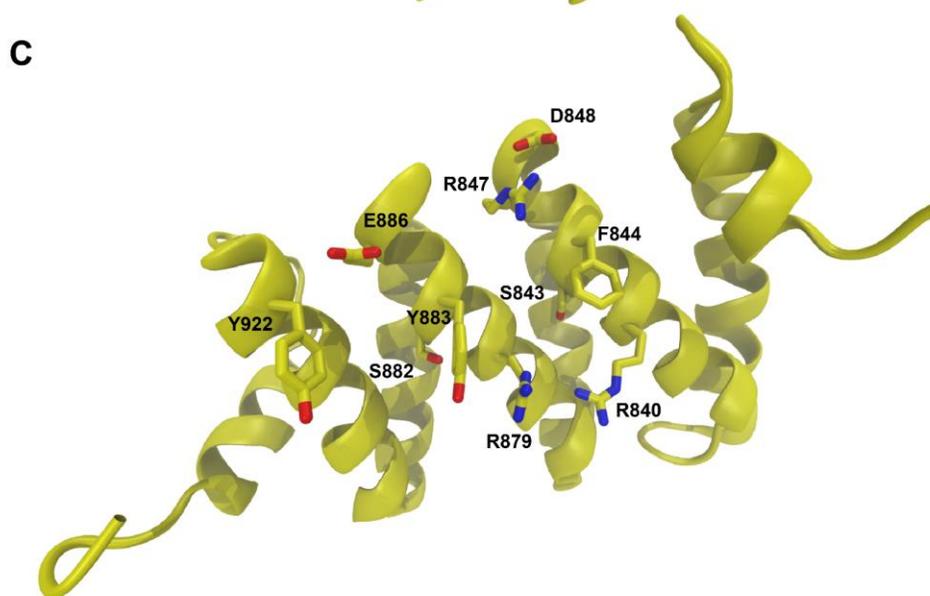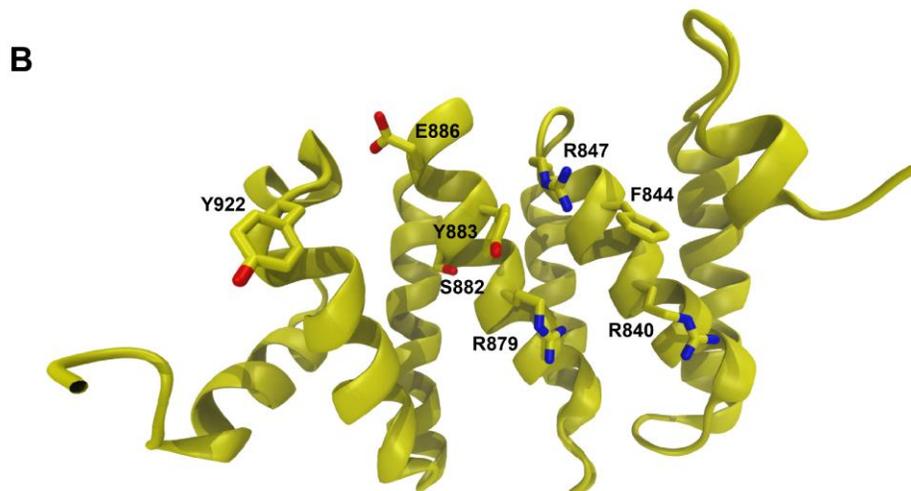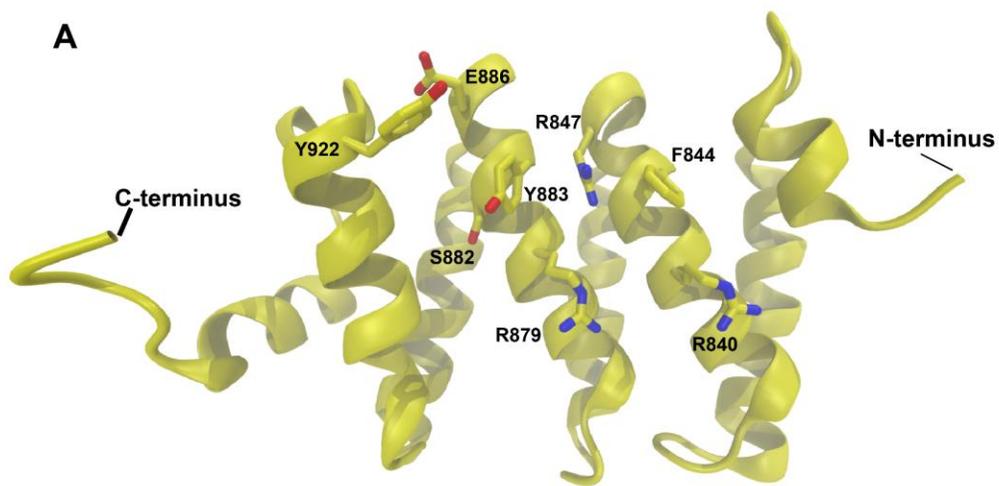
Figure 28. R824W simulation clustering analysis reveals another alternative +1-binding pocket conformation. (A,B,C) Representative frames of the three clusters identified from the R824W simulation with the hierarchical agglomerative clustering algorithm[151-153]. (A) Y922 occludes the cap-binding pocket. (B) An open cap-binding pocket, and canonical +1-binding pocket. (C) R847 flips out of the+1-binding pocket and interacts with D848. This change in residue position creates an alternate +1-binding pocket.

The same clustering analysis was done with the R824W simulation. Three clusters were identified and representative frames were extracted (Figure 28). The cap-binding pocket of the first microstate was collapsed and occluded by Y922; this microstate had a canonical +1-binding pocket (Figure 28A). The second microstate had an open cap-binding pocket, and a canonical +1-binding pocket (Figure 28B). Notably, the third microstate had an alternative +1-binding pocket, where R847 is flipped out of the pocket and interacts with D848 (Figure 28C). This alternative +1-binding pocket differed from that observed in the WT simulation; in contrast to the WT alternative +1-binding pocket, Y883 was not flipped into the +1-binding pocket.

I hypothesized the alternative +1-binding pockets offered distinct chemical environments compared to the canonical +1-binding pocket. The space R847 occupies in the canonical conformation is partially open in the R824W simulation alternative +1-binding pocket. Consequently, there is more space to accommodate a π-π stacking interaction with Y883 and F844. Although R847 no longer resides in the pocket, S843 could compensate as a hydrogen bonding partner for the pocket ligand.

4.3.3 R824W Alters the Dynamics of Residues in the +1-Binding Pocket
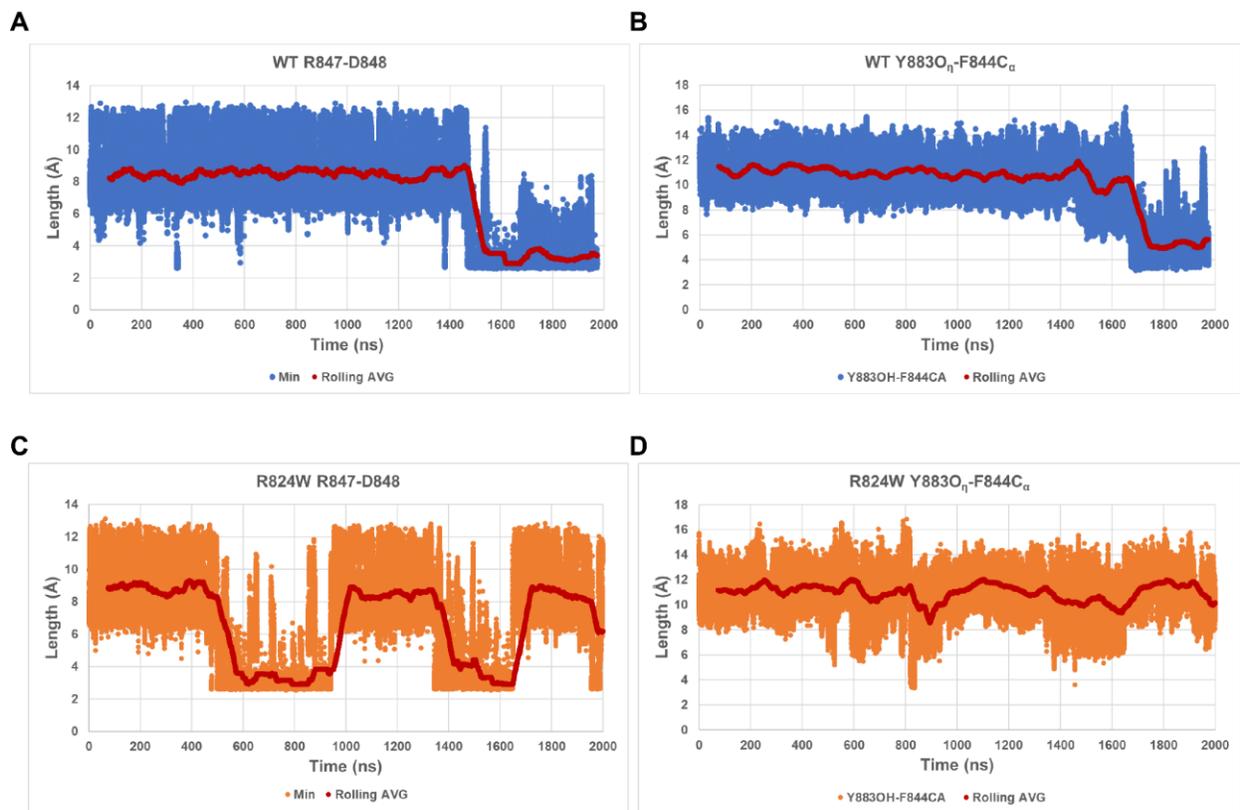
Figure 29. R847-D848 and Y883-F844 distances for WT and R824W simulations. (A) The minimum distance between the $N_{\eta 1}$ and $N_{\eta 2}$ of R847 and the $O_{\delta 1}$ and $O_{\delta 2}$ of D848 for the WT simulation (B) The distance between the O_{η} of Y883 and the C_{α} of F844 in the WT simulation. (B) The minimum distance between the $N_{\eta 1}$ and $N_{\eta 2}$ of R847 and $O_{\delta 1}$ and $O_{\delta 2}$ of D848 for the R824W simulation (B) The distance between the O_{η} of Y883 and the C_{α} of F844 in the R824W simulation. Raw data in blue, rolling average in orange.

Having identified alternative +1-binding pockets in both the WT and R824W simulations, I investigated how often these occurred. A difference in the probability for the alternate +1-binding pocket microstate, between WT and R824W LARP1 DM15, could mean a change in affinity for TOP mRNAs, a change in specificity for the +1 position, or even a shift to a different binding partner[79, 80, 84, 86, 98]. I first measured the distances between the $N_{\eta 1}$ and $N_{\eta 2}$ of R847 and the $O_{\delta 1}$ and $O_{\delta 2}$ of D848 in the WT simulation. Next, I determined the minimum distance between these

atoms for each frame and designated a cutoff of 4 Å or less for concluding the two residues were interacting (Figure 29A). I found that for the first ~1.4 μs of the simulation, the interaction between these side chains was scarce. However, after ~1.4 μs this interaction became much more common. The rolling average hovers around 8 Å until ~1.4 μs, upon which it drops to below 4 Å for the remainder of the simulation. This R847-D848 distance was less than or equal to 4 Å for 22.2% of the simulation, and the average length was 7.14 Å.

I also wanted to investigate how often Y883 flipped into the pocket, because this shift in residue position also alters the chemical environment of the +1-binding pocket and could drive a shift in affinity, specificity, or binding partner. To do this, I measured the distance between the O_η of Y883 and the C_α of F844 (Figure 29 B). I found that at ~1.7 μs there was a large decrease in the rolling average of this distance, which persisted until the end of the simulation. This coincides with R847 flipping out of the pocket to interact with D848. To see if either R847 or Y883 returned to their canonical +1-binding pocket positions, I extended the simulation another 100 ns. Neither R847 nor Y883 returned to their canonical position in this extended portion of the simulation. This suggests there is an energetic barrier between the canonical and alternate +1-binding pockets in the WT simulation, and the two microstates do not readily interconvert.

When we performed the same R847-D848 distance analysis with the R824W simulation we found four transitions between states (Figure 29C). The rolling average of the distance starts at ~8 Å and decreases to ~3 Å at ~500 ns, indicating the formation of an interaction between these side chains. Then at ~950 ns this distance increases to ~8 Å again, suggesting this interaction has broken. At ~1.35 μs there is another transition to ~4 Å, and the distance continues to decrease. Then at ~1.7 μs, there is another transition to ~8 Å. At the very end of the simulation, we see another transition starting with the rolling average decreasing similarly. However, the distance

does not reach below 4 Å. This is likely due to the simulation ending before the full transition occurs. The R847-D848 distance was less than or equal to 4 Å for 33.1% of the R824W simulation, a greater percentage of the time than the WT simulation. The average R847-D848 distance for the R824W simulation was less than that of the WT simulation, 6.59 Å. The same Y883-F844 distance analysis was performed with the R824W simulation. In contrast to the WT simulation, there was no transition to a distance less than 6 Å (Figure 29D).
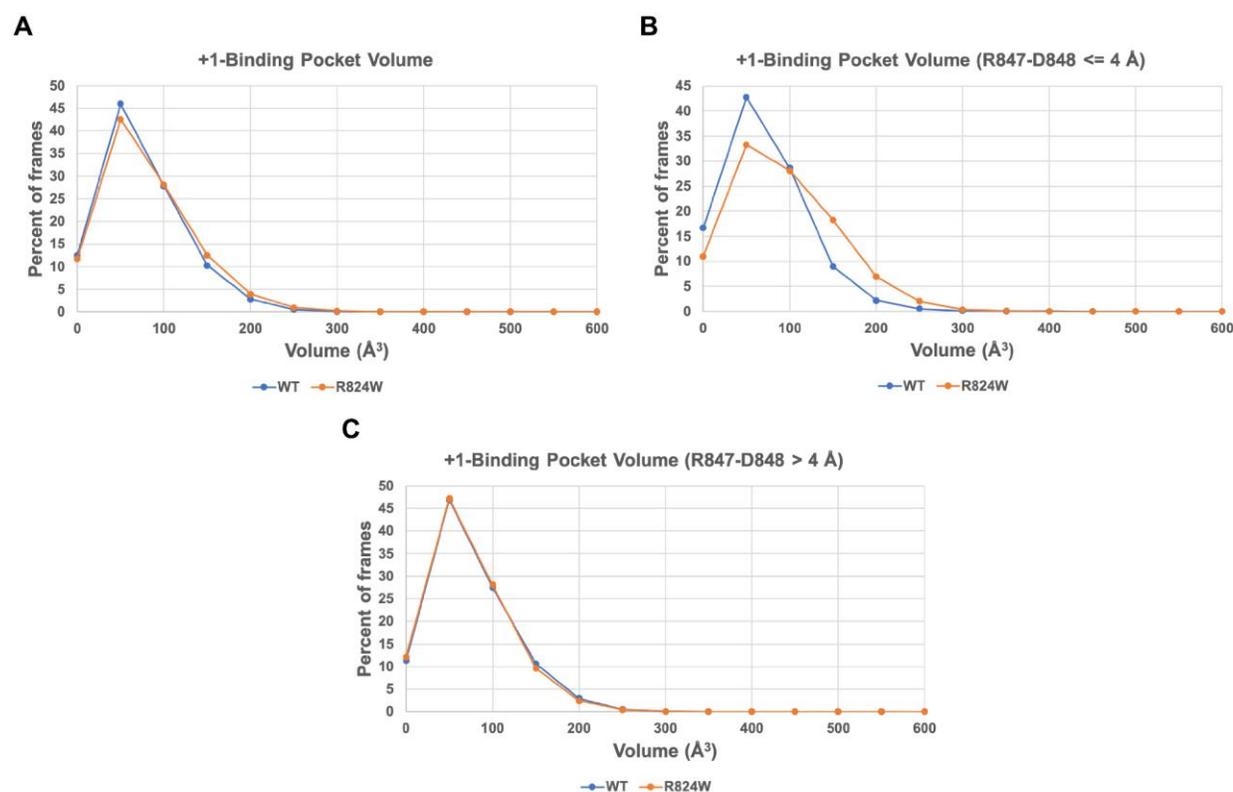
Figure 30. +1-binding pocket volume. (A,B,C) POVME 2.2 was used to calculate the +1-binding pocket volume in both the WT and R824W simulations[110]. (A) The +1-binding pocket volume for the WT and R824W simulations. (B) The +1-binding pocket volume, when the minimum distance between the Nη1 and Nη2 of R847 and the Oδ1 and Oδ2 of D848 is less than or equal to 4 Å. (C) The +1-binding pocket volume, when the minimum distance between the Nη1 and Nη2 of R847 and the Oδ1 and Oδ2 of D848 is greater than 4 Å.

These data show more transitions to an alternate +1-binding pocket in the R824W simulation compared to the WT simulation; this suggests that the R824W mutation could alter the probability of whether LARP1 DM15 is in a state capable of binding TOP mRNAs. The volume of the +1-binding pocket also contributes to determining if LARP1 DM15 is capable of binding TOP mRNA. To investigate the impact of the R824W mutation on the +1-binding pocket volume, I used POVME 2.2 to measure the +1-binding pocket volume over the course of each simulation (Figure 30)[110]. I plotted probability distributions of these volumes to investigate any differences. There did not seem to be any difference (Figure 30A); however, after isolating the frames in which the previously defined R847-D848 distance was less than or equal to 4 Å, we see a shift of the probability to the right, indicating a greater volume of this pocket is more probable in the R824W simulation (Figure 30B); when the frames with a R847-D848 distance greater than 4 Å are isolated, we observe no difference between the WT and R824W simulation. Therefore, the R824W alternative +1-binding pocket specifically is more likely to have a greater volume than that of the WT alternative +1-binding pocket. These data also suggest that the R824W mutation is causing a shift in the +1-binding pocket dynamics, and therefore could be altering the affinity for TOP mRNAs or altering the binding partner.

4.3.4 The Cap-Binding Pocket's Chemical Environment is also Affected by R824W
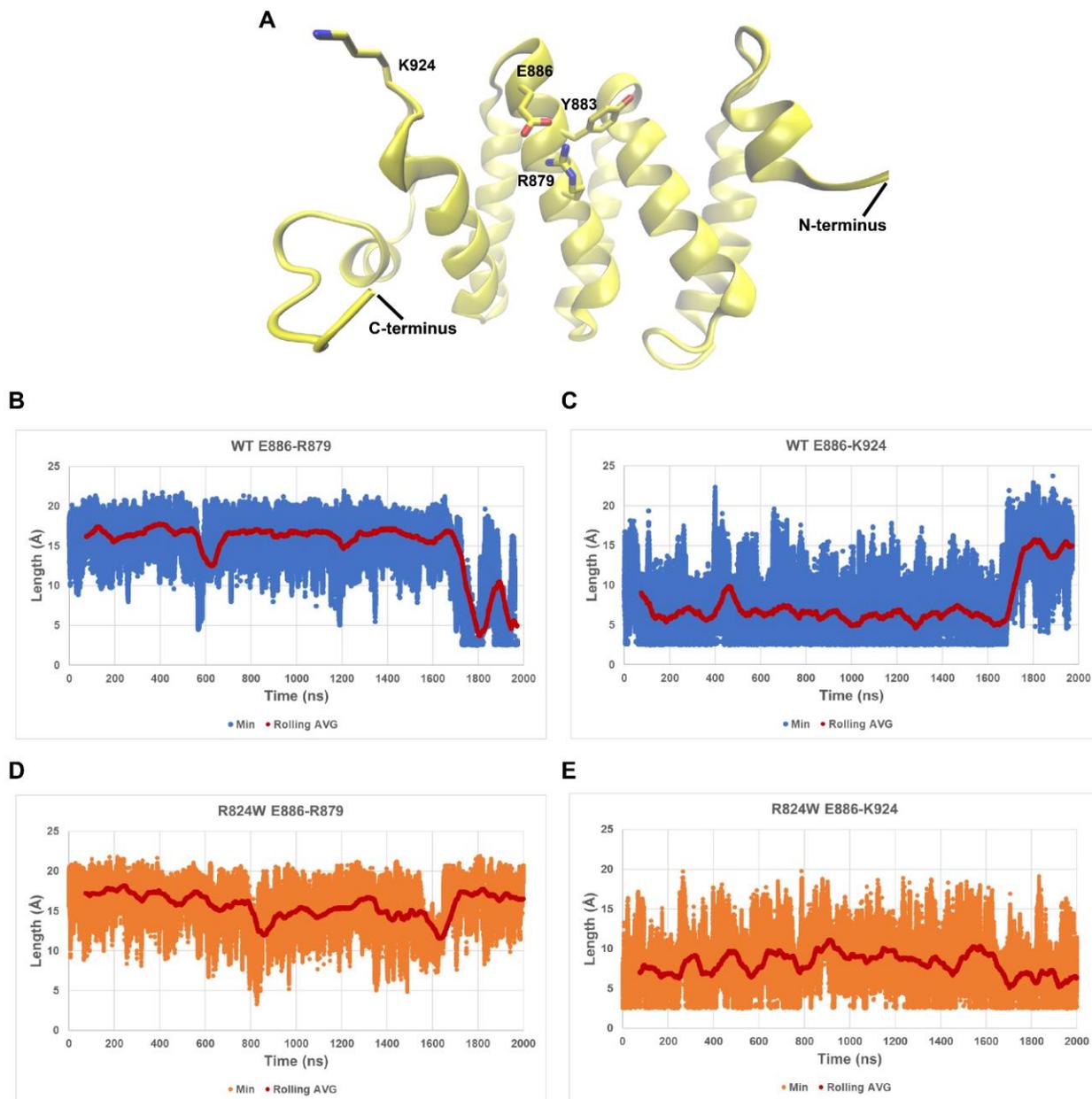
Figure 31. Cap-binding pocket dynamics are affected by the R824W mutation. (A) Representative microstate in which E886 interacts with R879 in the WT simulation. (B) The minimum distance between the Nη1 and Nη2 of R879 and the Oε1 and Oε2 of E886 for the WT simulation. (C) The minimum distance between the Nζ of K924 and the Oε1 and Oε2 of E886 for the WT simulation. (D) The minimum distance between the Nη1

The R824W mutation may also have an effect on the dynamics of the cap-binding pocket, which could affect the ability of LARP1 DM15 to bind capped mRNAs. I next investigated if the R824W mutation had any effect on the cap-binding pocket (Figure 31). Upon visual inspection of the WT trajectory, we identified a novel interaction between E886 and R879. The flipping of Y883 into the WT alternative +1-binding pocket allows for E886 and R879 to form a salt bridge (Figure 31A). This interaction has not been previously identified in LARP1 DM15 crystal structures to our knowledge.

We then measured the distances between the $N_{\eta 1}$ and $N_{\eta 2}$ of R879 and the $O_{\epsilon 1}$ and $O_{\epsilon 2}$ of E886 in the WT simulation. Next, we determined the minimum distance for each frame and classified the two residues as interacting if the distance between them was less than or equal to 4 Å (Figure 31B). The rolling average for this distance is greater than 10 Å for the first ~1.7 μs of the simulation. At this point in the simulation there was a sharp decrease to just below 5 Å. Then there was a sharp increase to ~10 Å from 1.8 μs to 1.95 μs, and then another sharp decrease to ~5Å. The previously described E886-R879 distance was less than or equal to 4 Å for 6.52% of the WT simulation. The average length was 15 Å.

Based on our initial 500 ns simulation of the wild-type DM15 region, we had proposed that K924 could sequester E886 away from the cap-binding pocket[98]. I hypothesized this interaction to be mutually exclusive of the E886-R879 interaction, because K924 would be too far away to interact with E886 if E886 is interacting with R879. I then measured the distances between the N_{ζ} of K924 and the $O_{\epsilon 1}$ and $O_{\epsilon 2}$ of E886 in the WT simulation. Next, we determined the minimum distance for each frame and deemed the two residues to be interacting if the distance was less than or equal to 4 Å Unsurprisingly, we observed a sharp increase in the E886-K924 distance and a

concomitant simultaneous decrease in the E886-R879 distance (Figure 31C). The previously defined E886-K924 distance was less than or equal to 4 Å for 22% of the WT simulation; The average length was 7.7 Å.

The R824W simulation was then investigated for a similar E886-R879 interaction, which could alter the affinity for capped mRNAs. We performed the same distance analysis with the defined E886-R879 interaction (Figure 31D). Unlike with the WT simulation, there was no sharp decrease in rolling average. In fact, the rolling average never dipped below 10 Å, indicating that this interaction is absent from the R824W simulation; the R824W mutation seems to have altered the dynamics of the cap-binding pocket in such a way that eliminates this interaction. I then investigated the previously defined E886-K924 distance in the R824W simulation. As expected, there was no sharp increase in distance for the E886-K924 interaction in the R824W simulation (Figure 31E). The previously defined E886-K924 distance was less than or equal to 4 Å for 13.8% of the simulation. The average length was 8.04 Å.
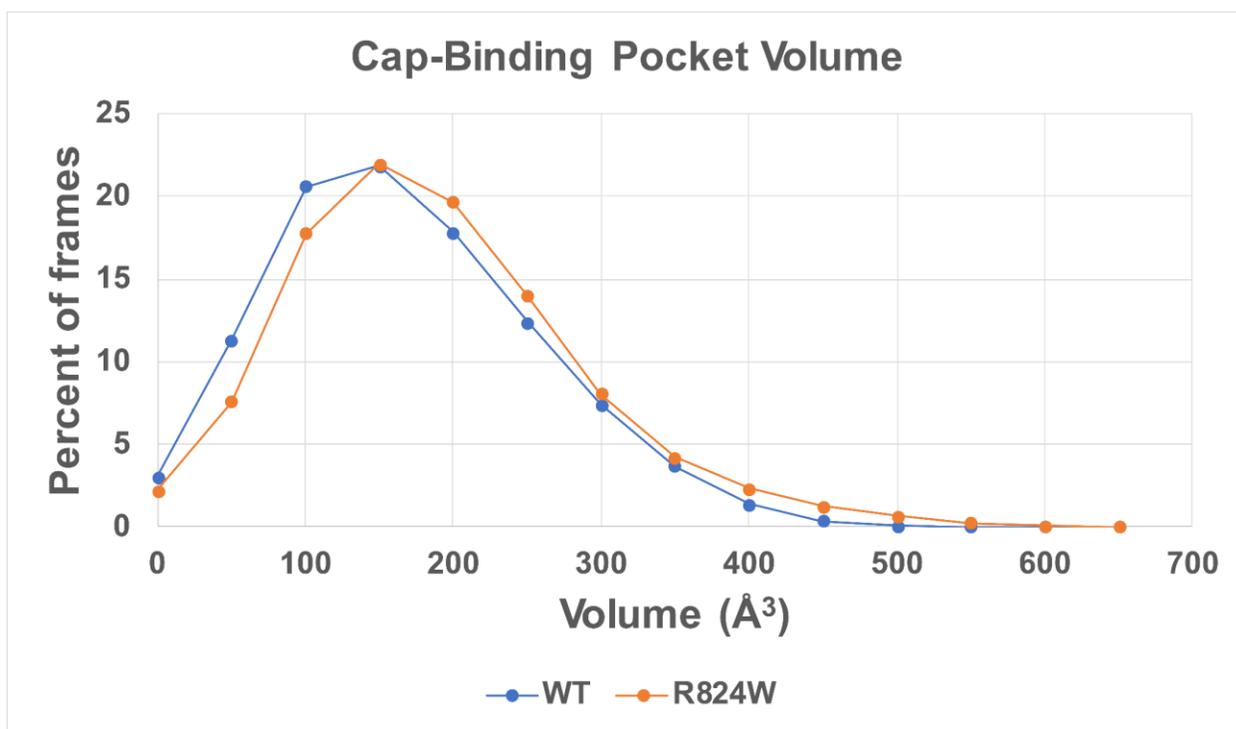
Figure 32. Cap-binding pocket volume. Probability distribution of the cap-binding pocket volumes for both the WT and R824W simulations. POVME 2.2 was used to calculate cap-binding pocket volume[110].

After having identified the R824W mutation alters the cap-binding pocket dynamics by changing the probability of inter-residue interactions, I used POVME 2.2 to measure the cap-binding pocket volume across both the WT and R824W simulations[110]. A change in the cap-binding pocket volume due to the R824W mutation could also mean a change in the ability of LARP1 DM15 to bind capped mRNAs. The probability distribution for the cap-binding pocket shows that the R824W simulation has a slight shift towards a more voluminous cap-binding pocket than the WT simulation does (Figure 32). Since the effect is rather small, it is difficult to predict if this would have a functional effect.

4.3.5 Contacts Between Residues, and Residue Communities Are Affected by the R824W Mutation Leading to a Different Alternate +1 Pocket Conformation
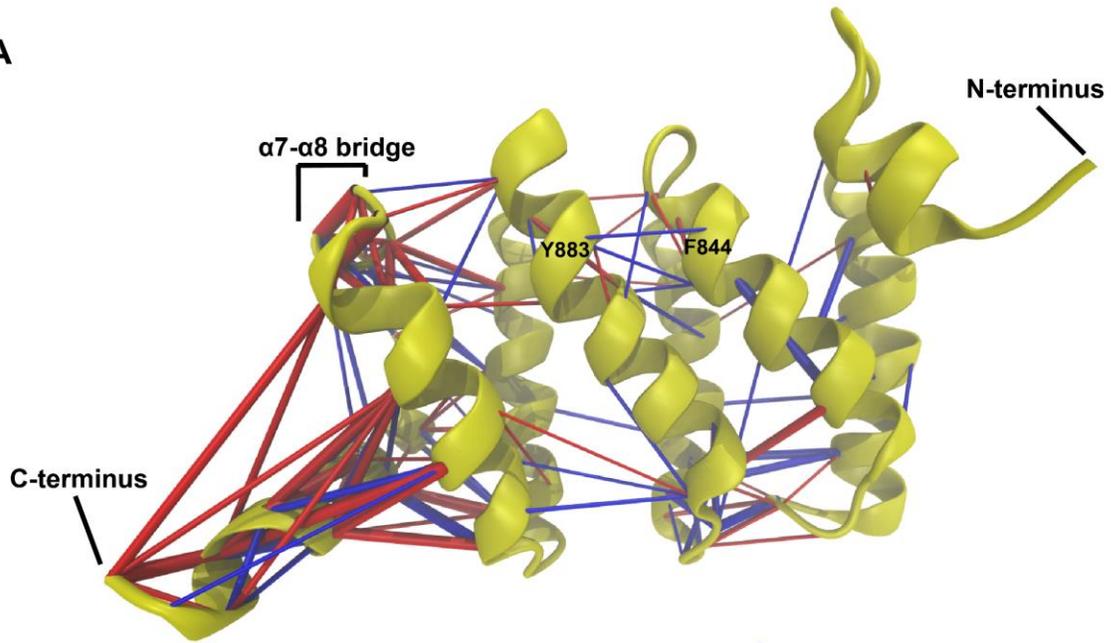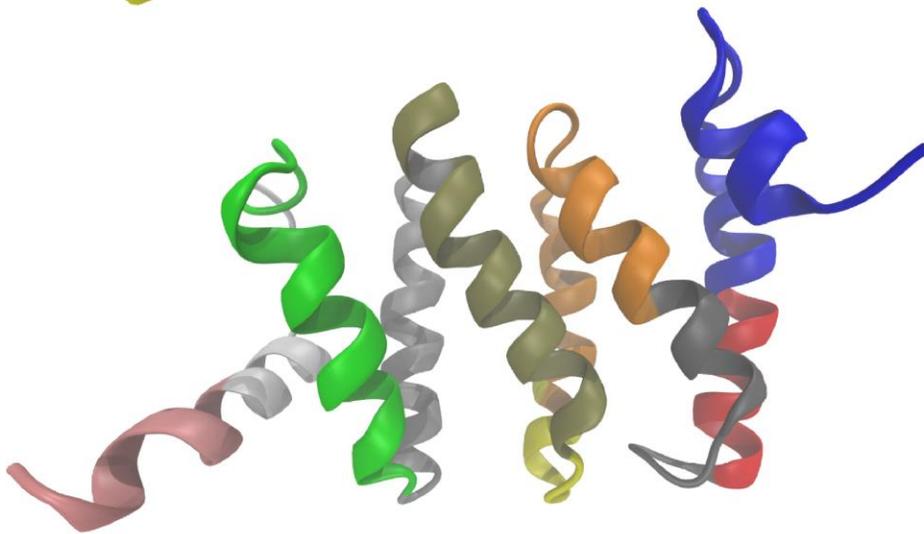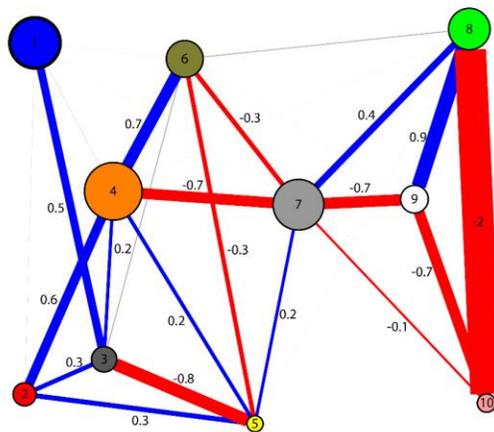
Figure 33. Difference contact network analysis reveals the effects of the R824W mutation on the α7-α8 bridge, +1-binding pocket, and PR824 dynamics[150]. (A) Residue-residue difference contact network for the WT and R824W simulations. Blue indicates higher probability of a residue-residue contact in the WT simulation (or a lower probability in the R824W simulation); red indicates lower probability of a residue-residue contact in the WT simulation (or a higher probability in the R824W simulation. Magnitude of the contact probability difference is indicated by the thickness of the cylinders. (B,C) Ten residue communities were identified from the conseus contact network. (B) The residue communities mapped onto the LARP1 DM15 structure from a frame of the WT simulation. (C) Community-community difference contact netowrk. Each node represents the corresponding colored residue community in B. The radius of each node is proportional to the number of residues in that community. Blue lines indicate a higher probability of a community-commiunity contact in the WT simulation; red lines indicate lower probability of a community-community contact in the WT simulation. Magnitude of the contact probability difference is indicated by the thickness of the lines.

I next employed the dCNA method to further investigate any allosteric effect imparted by the R824W mutation[150, 154-156]. The residue-residue difference contact network for the WT and R824W simulations revealed several clusters of changes in contact probability (Figure 33A). These residue-residue contact probability changes are mapped onto WT LARP1 with colored cylinders (cylinders connect the two residues). Blue indicates higher probability of a residue-residue contact in the WT simulation; red indicates a lower probability of a residue-residue contact in the WT simulation; the size of the cylinders indicates the magnitude of the change in contact probability. We see that there are several residue-residue contacts with a lower probability of formation in the WT simulation at the C-terminus. This involves several residues at the artificial truncation; therefore, it is hard to infer any meaningful information from these.

We also observe a higher probability of contact for the PR824 gating R824-D864 residue pair in the WT simulation. This makes sense since the Tryptophan in the R824W simulation cannot form a salt-bridge with D864, although, PR824 is still likely occluded due to the larger size of the

tryptophan. There are several other changes in residue-residue contact probability within and adjacent to PR824, including both higher and lower probabilities of contact within the WT simulation; this is likely due to the local effect the R824W mutant has on inter-residue interactions. One notable difference is the higher probability of the Y883-F844 contact in the WT simulation. This corresponds with the distance data for the Y883-F844 interaction previously described, and the lack of this interaction in the R824W simulation is likely due allosteric effects of the mutation. The R847-D848 contact probability is not included in this analysis, as only residues at least three away from one another are considered.

Another key area with a high density of contact probability difference is the α7-α8 bridge. There are several contacts involving residues within the α7-α8 bridge that have a lower probability of occurrence in the WT simulation. In fact, contacts between key residues involved in the formation of the 3_{10} helix have a lower probability of forming in the WT simulation. This could mean an increase in probability of a "ligand-ready" cap-binding pocket for R824W DM15 based on the trends associated with the secondary structure of the α7-α8 bridge discussed in Chapter 2[98]. Based on this observation we then calculated the percentage of time the α7-α8 bridge was a 3_{10} helix in both the WT and R824W simulations. Strikingly, the α7-α8 bridge was a 3_{10} helix for 44.9% of the R824W simulation but only 5.32% of the WT simulation.

The difference in residue community (a group of residues that are tightly connected) contacts, was used to investigate the macro effect of the R824W mutation on LARP1 DM15 dynamics. Ten residue communities were identified from the consensus contact network and suggest the modularity of the molecule (Figure 33B). α1, the α1- α2 loop, and part of α2 make up the first residue community. The second residue community is made of residues in the C-terminal half of α2. The third residue community is comprised of residues in the α2- α3 loop and part of α3.

The fourth residue community contains residues in α3, the α3- α4 bridge, and α4. The fifth residue community has residues at the C-terminal end of α4, and in the α4- α5 loop. The sixth residue community is made up of residues in α5. The seventh residue community comprises residues in α5, the α5- α6 loop, α6, and the α6- α7 loop. The eighth residue contains residues in the α6- α7 loop, α7, and residues in the α7- α8 bridge. The ninth residue community is made up of residues in the α7- α8 bridge, and α8. The tenth residue community is made of residues in the C-terminal part of α8.

Key residues in the cap-binding pocket are split between residue communities 6 and 8; key residues in the +1-binding pocket are split between residue communities 4 and 6. PR824 residues are in residue communities 2, 3, 4, 5, and 6. We do not see a change in the contact probability between residue communities 6 and 8, however we do see a higher probability of contact between residue communities 4 and 6 in the WT simulation. Residue community pairs 2-4, 2-4, 2-5, 3-4, 4-5, 4-6 all have a higher probability of contact in the WT simulation (or a lower probability in the R824W simulation), while residue community pairs 3-5, 5-6 have a lower probability of contact in the WT simulation. Residue community pair 3-6 does not change contact probability between the two simulations. The residue community pair 8-10 contact probability differences involve many residues at the artificial truncation of the construct, therefore it is difficult to infer any meaningful information from it. These changes in the community-community difference contact network illustrate the macro level dynamic changes caused by the R824W mutation, and that likely culminate in the observed alterations to the cap- and +1-binding pocket dynamics.

4.4 Discussion

Recent studies suggest allosteric regulation may be the key to reconciling conflicting models, supported by conflicting experimental results described in the literature, that indicate LARP1 may stimulate and/or repress translation and stabilize TOP mRNAs[2, 81-84]. Understanding the impact of LARP1 on TOP mRNA translation is key to successfully characterizing its role in translation regulation and how that role changes in disease[71]. Our previous work focused on elucidating the dynamics of the cap- and +1-binding pockets of LARP1 DM15[98]. We also previously identified a novel pocket in Chapter 3, PR824, that may have an allosteric relationship with the +1-binding pocket. However, the consequences of perturbing PR824 had not been investigated until now. The recorded R824W mutation associated with breast cancer was a prime motivation and candidate to explore the effects of perturbing PR824. I extended the existing WT 500 ns LARP1 DM15 4ZC4:B simulation to ~2 µs and carried out a ~2 µs simulation for the LARP1 DM15 4ZC4:B R824W mutant discussed in Chapter 3.

The tryptophan in the R824W mutant is larger than the WT arginine and decreases PR824 pocket volume or fully occludes it. I first measured any effect of the R824W mutation on the PR824 pocket volume. The probability distribution of the PR824 pocket volume did not vary greatly between the WT and R824W simulations (Figure 26), although there was a slight increase in the percentage of frames that had a PR824 pocket volume of 15 Å^3 or less. This is not a surprise, as the pocket already had a high probability of being closed in the WT simulation. In addition, as previously described in Chapter 3, R824W occludes and bifurcates the pocket; this means that even if there is some pocket volume present, it may not be as assessable as that in the WT simulation. Although R824W may occlude access to PR824, it seems to cause an allosteric effect that could be similar to that of a binding event in PR824 via altering the contact network.

4.4.1 Alternate +1-Binding Pockets May Bind an Alternate +1 Nucleotide

Clustering analysis with respect to the positions of key residues in the cap- and +1-binding pockets revealed alternate +1-binding pockets in both the WT and R824W simulations that may accommodate an alternate +1-nucleotide or could even bind a different type of molecule entirely (Figure 27, Figure 28). In both of these alternate +1-binding pockets, R847 flips out of the pocket and interacts with D848. Unique to the WT alternate +1-binding pocket is a largely persistent change of position for residue Y883, in which it flips into the pocket; this position also alters the canonical +1-binding pocket This Y883 position could potentially stabilize the observed alternative +1-binding pocket via a cation-π interaction with R847, and sterically occlude R847 from flipping back into the pocket. This Y883 position is not observed in the R824W simulation, which could partly explain why the canonical and alternate +1-binding pockets transition more rapidly in that context.

Unpublished studies have shown that LARP1 can bind transcripts starting with +1 nucleotides other than a C (E. Nguyen and Berman, unpublished), and the alternate +1-binding pockets identified in these simulations could potentially accommodate these +1 nucleotides. The WT and R824W simulation alternative +1-binding pockets provide different chemical environments from one another and from the canonical +1-binding pocket (Figure 27, Figure 28). These chemical environments could be more receptive to binding alternate +1 nucleotides, thereby altering the class of transcripts whose translation LARP1 regulates. In the WT simulation alternate +1-binding pocket, the nucleobase could π-π stack with F844, and/or engage in a t-shaped π-π stacking interaction with Y883. S843 could act as a hydrogen bond partner in this alternate +1-

binding pocket. In the R824W simulation alternate +1-binding pocket, the nucleobase has more space to align π-π stacking interactions with Y883 and F844, because R847 flips out of the pocket to and interacts with D848. Once again S843 could acts as a hydrogen bond partner.

The energetic hill that must be surmounted to transition between the canonical and alternate +1-binding pocket conformations is potentially much smaller with the R824W mutation present. The two states interconvert more rapidly in the R824W simulation and this could affect LARP1 DM15 affinity for TOP mRNAs, cause a change +1-binding pocket specificity, or alter the binding partner to a completely different molecule. Previous studies have also indicated that LARP1 can bind the 3' UTR of mRNAs as well, and it is possible the alternate +1-binding pockets could better accommodate the 3' UTR[84, 86]. We observe several transitions between the two pocket states in the R824W simulation, but only one in the WT simulation. Even in the extended 100 ns additionally simulated, the WT simulation did not revert back to the canonical +1-binding pocket. This could mean that the LARP1 DM15 R824W +1 position preference is more rapidly changing, leading to a shift in the composition of transcripts being regulated. Additionally, it may be possible that a different source of allosteric regulation, such as mTORC1 phosphorylation, could shift the balance from the canonical +1-binding pocket to the alternate form[2, 35]. If the alternate +1-binding pocket alters the LARP1 affinity for TOP mRNAs, then this could be the mechanism explaining the previously observed phosphorylation effects.[2, 35]

4.4.2 The E886-R879 Interaction May Help "Lock" the WT Alternative +1-Binding Pocket in Place

Another potential contributing factor to the stability and persistence of the WT alternative +1-binding pocket is the newly identified E886-R879 interaction (Figure 31). A salt-bridge

interaction between E886 and R879 could increase the stability of this alternate pocket state, and it sterically occludes Y883 from flipping back out of the pocket. In addition, although the cap-binding pocket volume probability distributions vary only slightly between the WT and R824W simulations, the E886-R879 interaction sequesters E886 to a new position. This could further alter the ability of LARP1 DM15 to bind capped transcripts; an alteration to the cap-binding pocket could also shift the binding partner of LARP1 DM15 to mRNA 3' UTRs as discussed above.

4.4.3 dCNA Further Reveals the Allosteric Effect of the R824W Mutation

The dCNA sheds light on residue-residue contact, community-community contact, and even secondary structure formation differences (Figure 33)[150]. These differences illustrate the small- and large-scale dynamics alterations caused by the R824W mutation, including the changes to the inter-residue contacts of the cap- and +1-binding pocket. This analysis further supports the hypothesis that Y883 moving into the +1 pocket is more likely in the WT simulation. There was also a cluster of residue-residue contact probability changes in the α7-α8 bridge, with several being less likely to form in the WT simulation. The α7-α8 bridge was more likely to form a 3_{10} helix in the R824W simulation (44.9%) than in the WT simulation (5.32%). Our analysis in Cassidy, Lahr et al. 2019 would suggest this means the LARP1 DM15 R824W mutant would have a higher probability of being in a "ligand-ready" state than WT LARP1 DM15[98]. The simplest interpretation of these data is that this is true. However, we did not previously sample several of the conformational states seen in the ~2 μs simulations, most notably the alternate +1 binding pockets and the E886-R879 interaction. The latter, in particular, would at least change the trends associated with K924 sequestering E886, because of the newly observed and mutually exclusive interaction (Figure 31). Further analysis will need to be done to determine the effects of the α7-α8

bridge secondary structure on the dynamics of the cap-binding pocket with this more representative conformational ensemble.

The ten residue communities identified were used to investigate community contact probability changes, the net change in contact probability of all edges between two communities, caused by the R824W mutation. We found no change in contact probability for the residue communities comprising the cap-binding pocket residues, but there was an increase in contact probability for the residue communities comprising the +1-binding pocket residues. These community contact probability changes illustrate the larger scale effects of the allosteric R824W mutation, which likely lead to the changes in dynamics for the+1-binding pocket. Although there were no changes in contact probability for the communities the cap-binding pocket residues belong to, there were residue contact probability changes in the cap-binding pocket that were likely caused by the R824W mutation. There were also several changes in the contact probabilities between the residue communities comprising PR824, unsurprising and likely due to the larger tryptophan altering contacts locally. This is the source of the perturbation to the system which induced the observed changes in the residue-residue and community-community contact probabilities.

These data provide insight into how WT LARP1 DM15 could recognize alternate +1 nucleotides or 3' UTRs, and how the R824W mutation impacts the dynamics of the cap- and +1-binding pockets. Further analysis will need to be done to determine if the R824W mutation affects the affinity of LARP1 DM15 for TOP mRNAs or 3' UTRs. In addition, the effects of other known modes of LARP1 allosteric regulation need to be investigated. The same dCNA approach could be used in the future to characterize these effects. For example, one could use this approach to characterize the allosteric effects of specific mTORC1 phosphorylation events[35]. Structures of

118

extended constructs of LARP1 DM15 containing the phosphorylated residues would be necessary

for this analysis.

5.0 Drugging LARP1 DM15: A Cancer and COVID-19 Relevant Protein

5.1 Introduction

The amount of protein synthesis needed by a cell can vary. This depends on its cell type, stage in the cell-cycle, and environment[10, 157-161]. A cell responds to stressors or the current availability of nutrients by altering protein production. This requires a regulatory mechanism that can react to these environmental cues, and dysregulation of these mechanisms can lead to disease phenotypes, such as cancer[10, 56, 59, 162]. This regulatory process includes the crucial step of ribosome biogenesis[1, 11, 56, 59].

The mTORC1 signaling pathway serves as a central node in cell growth[10, 33, 35, 43, 84]. Translation of the mRNAs that encode ribosomal proteins is regulated by mTORC1[28, 35, 163]. These terminal oligopyrimidine (TOP) mRNAs, contain characteristic untranslated nucleotides that respond to signals that call for accelerating cell growth or inhibiting it[28, 29, 164]. mTORC1 abnormal activation has been observed in several cancers and has been targeted pharmacologically, such as in kidney cancer[10, 31, 32, 50, 52, 53, 55, 56, 58, 59, 63, 165]. However, due to its role in many other cellular processes, there can be many side effects[64].

La- related protein 1 (LARP1) has recently been identified as a node downstream of mTROC1 that regulates the translation of ribosomal proteins, and high expression of this protein has been connected to several cancers[2, 71, 85-87]. However, its exact role in this regulation is still being investigated; there is seemingly contradictory evidence that LARP1 can stimulate or repress this translation[2, 49, 79-83]. Therefore, it is important to investigate both the therapeutic goals of

inhibiting or enhancing this interaction. Furthermore, small molecules that are identified to modulate this interaction could be used to elucidate this mechanism more fully.

LARP1 is an attractive pharmacological target downstream of mTORC1; inhibiting LARP1 instead of mTORC1 could potentially minimize side effects. Recently, LARP1 has also been shown to bind SARS-CoV-2, the virus causing COVID-19, mRNA[96, 97]. LARP1 associates with the 5'UTR of SARS-CoV-2 mRNA, and additionally has been shown to repress replication[96, 97]. Due to these interactions, LARP1 has been posited as a potential therapeutic target for treatment of COVID-19[96, 97].

Previous work has guided the effort to drug the DM15 region of LARP1. Structural work demonstrated that the DM15 region of LARP1 recognizes the 7-methylguanosine cap and invariant first cytosine (+1) of mRNAs encoding ribosomal proteins[79, 80]. We subsequently detailed the dynamics governing the mechanism of mRNA binding by DM15. We utilized molecular dynamics simulations of the LARP1 DM15 structure to investigate druggability[98]. We found evidence supporting that the cap- and +1-binding pockets are both druggable[98]. However, we had not investigated any potential compounds that could bind these or other pockets. Here we utilize druggability simulations, molecular dynamics simulations with small molecule probes, to identify pharmacophores (chemical blueprints to filter compounds by) targeting the cap- and +1-binding pockets. We also identify a pharmacophore targeting a potentially allosteric pocket. A virtual screen was used to further filter potential hits, followed by thermal shift assay validation for the top-ranking compounds. We identify five compounds that thermally stabilize LARP1 DM15, indicating binding.

5.2 Methods

5.2.1 Model Building, Parameterization, and Druggability Simulations

Fourteen microstates were used as starting points for the druggability simulations: Twelve representative conformational states previously identified from the 500ns LARP1 DM15 4ZC4:B simulation (discussed in Chapter 2), the previously identified largest cap-binding pocket microstate, and the LARP1 DM15 4ZC4:B structure[98]. Hydrogen atoms were added using the PDB2PQR 2.1.1 implementation of the PROPKA algorithm, with pH of 7.0[123, 124]. The hydrogen-bond network was also optimized with PDB2PQR[123, 124]. The protein structure file was generated using the autopsf psfgen plugin of VMD[133]. The topology files required by the VMD plugin DruGUI were used[166]. DruGUI was used for the simulation setup using the default settings, with the exception of the simulation box padding being set to 10 Å[166]. Each starting microstate was run for three 40ns separate production runs; and the NAMD 2.9 molecular dynamics simulation package was used[129, 130]. Probe grid calculation and druggability analysis were done utilizing the default settings. Druggable sites were visualized via DruGUI and categorized based off of the binding pockets they targeted[166]. Structure visualizations were made using VMD and DruGUI[133, 166].

5.2.2 Pharmacophore Building, and Pharmit Small Molecule Search

Pharmacophores were built utilizing the coordinates of druggable sites identified, and assigning properties to individual interaction spots based on the corresponding probe molecules. The chemical environment of the receptor was also taken into account. The receptor PDB fie and

pharmacophores were loaded into the Pharmit web server[167]. An exclusion receptor setting of 1.5 was utilized unless it was deemed too restrictive. These features and settings were used to search the ZINC database via Pharmit[167, 168].

5.2.3 Further Filtering and Virtual Screen

The identified small molecules were then further filtered via the Lipinski rule of 5 with Open Babel 2.3.2[169]. Ligand 3D models were prepared for the virtual screen by utilizing Gypsum-DL 1.1.1 to create up to 5 variants[170]. Ligands and receptors (the protein structures) were further prepared for the virtual screen utilizing MGLTools 1.5.6[171, 172]. Ligands were then docked into the receptor that their matching pharmacophore was identified with. AutoDock Vina 1.1.2 was used for docking: size x = 25, size y = 25, size z = 25, energy range = 4, exhaustiveness = 100[173]. The center of each docking box was determined via getting the geometric center of the respective pocket with Scoria[174]. The docking box was extended to include both the cap- and +1-binding pockets. This was done due to the similar π-π stacking capabilities of both pockets; it could have been possible a ligand would have a higher affinity for the other pocket. The top ligands and respective poses were ranked via docking score (ΔG). Visual inspection was then done to further refine ranking based on interactions with the receptor. Jennifer Walker provided custom pipeline scripts that were modified and used to facilitate ligand preparation, receptor preparation, and docking.

5.2.4 Thermal Shift Assay

The top twenty small molecules based on the virtual screen and the m^{7}GpppC positive control were then tested for their ability to stabilize the protein fold of LARP1 DM15 by protein thermal shift assay. The protein thermal shift assay was performed using the protocol described for the protein melting temperature assay from Cassidy, Lahr et al. 2019[98]. Alterations were as follows: Ligand or DMSO was added to a concentration of 500 µM. The final concentrations of the other buffer components were 52 mM Tris pH 8; 63 mM NaCl; 103 µM DTT; and 11% glycerol. Ligands were purchased from MolPort. MolPort IDs: Compound 1 - MolPort-007-553-381, Compound 2 - MolPort-000-732-625, Compound 3 - MolPort-000-722-539, Compound 4 - MolPort-005-919-164, Compound 5 - MolPort-004-573-505. LARP1 DM15 was obtained from Elaine Nguyen of the Berman lab.

<div align="center">

5.3 Results

</div>

5.3.1 Druggability Simulations Reveal Exploitable Pharmacophores

Previously, we had identified druggable hotspots but had not identified any potential ligands[98]. We ran druggability simulations using the DruGUI plugin of VMD and NAMD 2.9 to identify pharmacophores targeting key LARP1 DM15 pockets[129, 130, 166]. The twelve previously identified representative LARP1 DM15 microstates from the 500ns simulation were used as starting points for these simulations. We also used the LARP1 DM15 4ZC4:B structure, and the previously identified microstate with the largest cap-binding pocket[80, 98]. Since the druggability

simulations are of short timescales, having several distinct starting structures is important to sample as much conformational space as possible. Each simulation was run for 40ns and three production simulations were run for each starting structure. We identified druggable sites targeting the cap-binding pocket, +1-binding pocket, cap/+1-binding pockets combined, and a pocket adjacent to the α7-α8 bridge.

These druggable sites were comprised of interaction spots, representing a receptor/probe interaction, that were linked together to form a drug like size. Predicted binding affinity was used to prioritize druggable sites for conversion into pharmacophores. We then created pharmacophores for these druggable sites by attributing the appropriate probe's chemical characteristics at the corresponding location. The chemical environment introduced by the receptor was also used to inform the creation of the pharmacophore. In total, we identified eight pharmacophores targeting the cap-binding pocket, sixteen targeting the +1-binding pocket, sixteen targeting the cap/+1-binding pockets combined, and ten targeting the α7-α8 bridge adjacent pocket.

5.3.2 Database Screening and Docking Reveal Cap- and +1- Binding Pocket Chemical Features Engaged by Compounds

Pharmacophores and the receptors they were identified from were loaded into the Pharmit online server to identify matching ligands[167]. We then used a virtual screen of the identified compounds as a second filter. Ligands were docked into the starting structure the corresponding pharmacophore was identified from. A ranked list of ligands, targeting the previously mentioned pockets, was created by sorting by docking score and visual inspection. From this list the top twenty compounds were then chosen for further characterization.

125
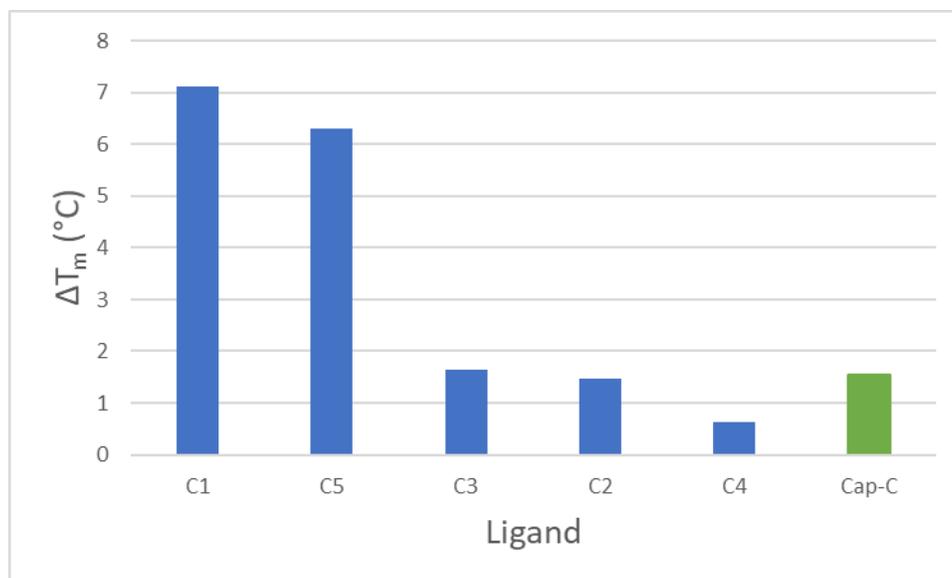
Figure 34. Potential compound hits increase thermal stability. Thermal shift assays reveal that five of the selected compounds increase the thermal stability of LARP1 DM15 (C1-C5, blue). Three cause a greater increase in protein stability than known binder m7GpppC (green), and two afford a slightly lower increase[79]. n = 1.

A protein thermal shift assay (TSA) was used to validate the *in silico* results and identify binders via an increase in LARP1 DM15 thermal stability (Figure 34). Out of the twenty compounds tested, five afforded an increase in thermal stability. Three caused a greater increase in protein stability than known binder m^7GpppC, and two afford a slightly lower increase (Figure 34)[79]. Docking reveals that four out of the five potential hits are predicted to bind at least partly in the cap-binding pocket (Figure 35 B-E). Two of the five bind the cap-binding pocket and extend over Y883 partly into the +1-binding pocket (35B, 35E). Two of the five bind the cap-binding pocket and extend towards the N-terminus of α5 and α7 (Figure 35C, 35D). Lastly, one of the five binds a pocket adjacent to the α7-α8 bridge (figure 35F).
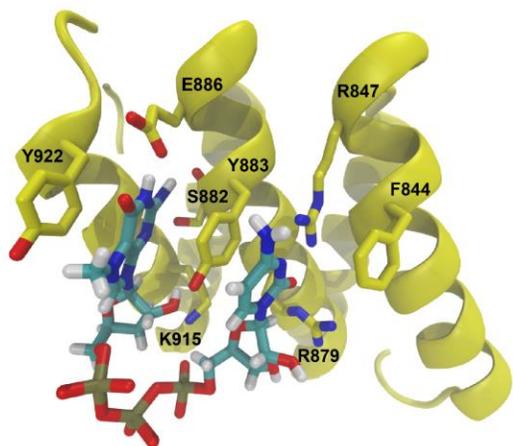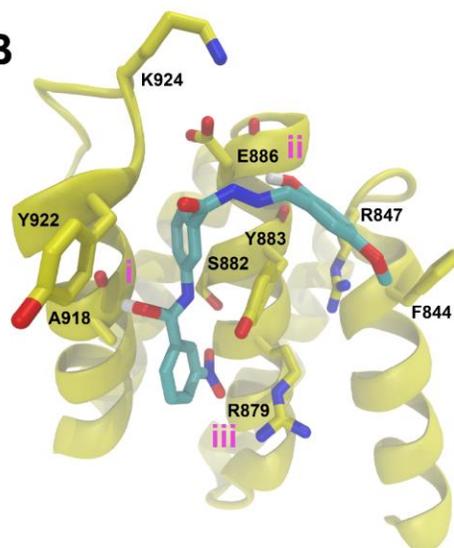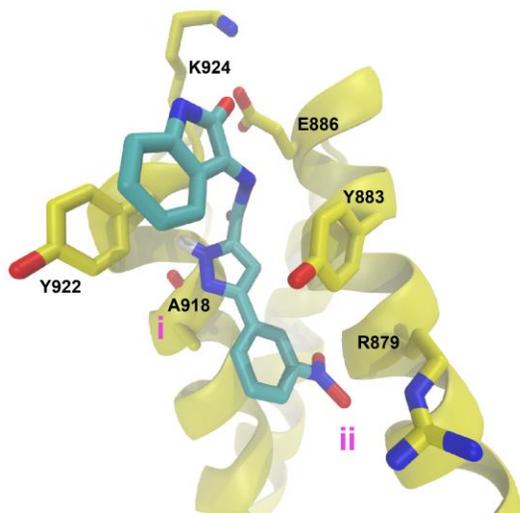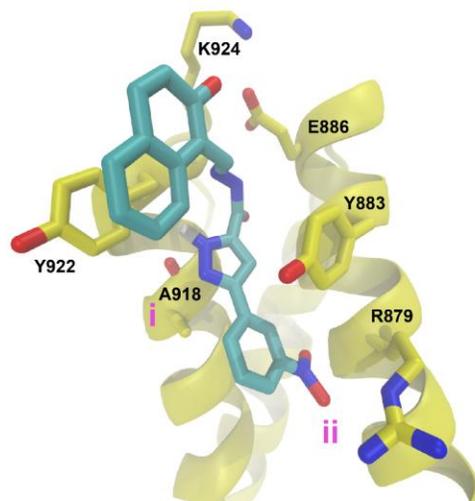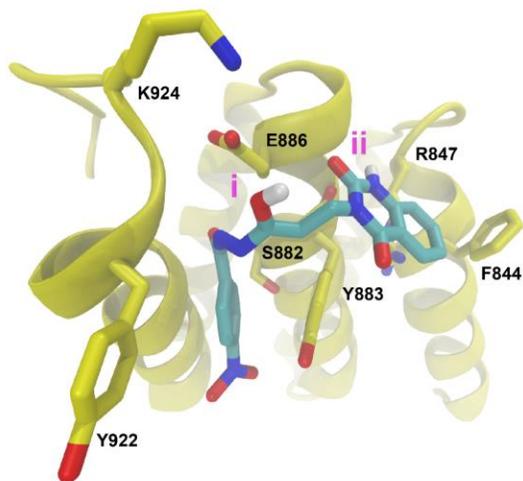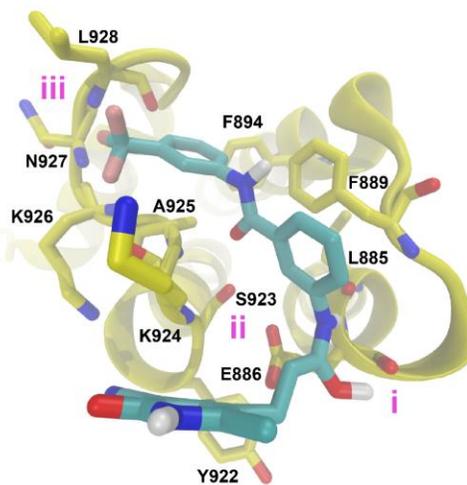
Figure 35. Small molecule putative hits docked into LARP1 DM15. (A)The 5V87:B cocrystal structure illustrating how known binder m7GppppC engages the cap- and +1- binding pockets[79, 80]. (B) Compound 1 docked into the cap- and part of the +1-binding pocket.(C) Compound 2 docked into the cap-binding pocket. (D) Compound 3 docked into the cap-binding pocket. (E) Compound 4 docked into the cap- and part of the +1-binding pocket. (F) Compound 5 docked into the α7-α8 bridge adjacent pocket.

These compounds will be referred to as compounds 1-5 for the remainder of this chapter, as designated in the caption of Figure 35. Compound 1 engages the cap-binding pocket via π-π stacking with Y922/Y883 of the receptor, and the center aromatic ring of the compound. There is also a potential t-shaped π-π stacking interaction with Y883 and one of the end aromatic groups of the compound. The OH group toward the one end of the compound engages in a hydrogen bond with the backbone O of A918 (Figure 35B, i). The other OH off the aromatic ring in the compound hydrogen bonds with the backbone O of Y883, and is in close proximity to the backbone O of E886 (Figure 35B, ii) The nitro group in the compound engages in hydrogen bonding with R879 (Figure 35B, iii).

Compounds 2 and 3 bind the cap-binding pocket in similar poses to each other. Both have an aromatic group that π-π stacks with Y922, and a second aromatic group that π-π stacks with Y883 (Figure 35C, 35D). They also both have an NH group that hydrogen bonds with the backbone O of A918 (Figure 35C, i, 35D, i). The nitro group in both compounds hydrogen bonds with R879 (Figure 35C, ii, 35D, ii)

Compound 4 engages the cap-binding pocket via π-π stacking with Y922/Y883 of the receptor. There is also potentially t-shaped π-π stacking with the compound and Y883, where the compound intrudes into the +1-binding pocket. Furthermore, there is hydrogen bonding between an OH group of the compound and E886 (Figure 35E, i); an NH group of the compound engages in hydrogen bonding with the backbone O of Y883.

Compound 5 was found to bind the pocket adjacent to the α7-α8 bridge. There are several instances of hydrogen bonding between the compound and receptor: an OH group of the compound and the backbone O of E886 (Figure 35F, i), an NH group of the compound and the backbone O of L885 (Figure 35F, i), an O of the compound and S923 (Figure 35F, ii), and three F in the compound and the backbone N of K926, N927, L928. Docking scores and SMILEs for all compounds can be found in Table 1.

Table 4. LARP1 DM15 putative hit characteristics, and *in silico* docking score.

Compound #	SMILES	Binding pocket	Docking score (ΔG)
1	COc1ccc(O)c(\C=N\NC(=O)c2cccc(NC(=O)c3cccc(c3)[N+]([O-])=O)c2)c1	Cap/+1	-8.6
2	[O-][N+](=O)c1cccc(c1)-c1cc(n[nH]1)C(=O)N\N=C1/C(=O)Nc2ccccc12	Cap	-8.6
3	Oc1ccc2ccccc2c1\C=N\NC(=O)c1cc([nH]n1)-c1cccc(c1)[N+]([O-])=O	Cap	-8.5
4	[O-][N+](=O)c1ccc(cc1)C(=O)NNC(=O)CCn1c(=O)[nH]c2ccccc2c1=O	Cap/+1	-8.4
5	Cc1[nH]c(=O)c(C#N)c(C)c1CCC(=O)Nc1cccc(c1)C(=O)Nc1cccc(c1)C(F)(F)F	α7-α8 bridge	-8.7

5.4 Discussion

5.4.1 Compounds That Thermally Stabilize LARP1 DM15 Have Potential Enhancing or

Inhibitory Roles on LARP1 DM15/TOP mRNA Binding

Due to the complex nature of the role of LARP1 in translational regulation, the therapeutic goal may switch between enhancement or inhibition based on the context[2, 97]. This means that identifying two sets of compounds, those with the ability to inhibit or enhance LARP1 DM15/mRNA binding, is paramount. Here I identified 4 compounds (Compounds 1-4) that bind in the cap- or cap/+1-binding pockets. Notably, all of these compounds extend beyond just the cap-binding pocket. This increases confidence in the specificity of the compounds. These compounds would likely exclude TOP mRNA access to these pockets due to steric occlusion. This would lead to a potential inhibitory effect on LARP1 DM15/mRNA binding. This therapeutic strategy could be attractive in several carcinoma contexts. LARP1 has been shown to be upregulated in several carcinomas, and its ability to sustain translation via stabilizing TOP mRNA could be diminished with this strategy[71].

Alternatively, a context in which one would potentially want to enhance LARP1 DM15/mRNA binding is COVID-19 treatment. LARP1 has been shown to associate with several viral RNAs, and notably has recently been shown to bind the 5'UTR of SARS-CoV-2 mRNA[96, 97]. Further, it has been shown that LARP1 represses SARS-CoV-2 replication[97]. This could be due to LARP1 DM15 outcompeting translation initiation factors for SARS-CoV-2 mRNA, and in this context the goal would be to enhance the LARP1 DM15/mRNA interaction to prevent translation

of viral proteins. An allosteric binder would be required in this case, and we previously showed how the α7-α8 bridge secondary structure influences the dynamics of the cap-binding pocket[98]. Compound 5 binds to residues in this bridge and could impart an allosteric effect via this influence. Further study would be required to identify the exact effect, inhibitory/enhancing, on LARP1/mRNA binding. However, this is a novel first step towards this therapeutic goal. In addition to the therapeutic uses, these compounds could be used as probes in experiments to further elucidate the role of LARP1 in the regulation of translation.

These two strategies, direct inhibition or allosteric modulation, allow for a greater level of flexibility in therapeutic strategy than either one on its own. Here we describe the first tangible in drugging LARP1 DM15, via identification of five putative LARP1 DM15 binding compounds. Further study is required to validate binding affinity, site of binding, and effect on LARP1 DM15/mRNA binding.

6.0 Discussion

6.1 Allosteric Regulation of LARP1

LARP1 binding of TOP mRNA has been shown to be sensitive to allosteric modifications[2, 35, 81]. There are at least 26 mTORC1 phosphorylation sites in LARP1[35]. An extended LARP1 DM15 construct (665–947) was used by Jia et. al to characterize the effects of these modifications on TOP mRNA binding *in vitro*[35]. Phosphomimetic mutation of a cluster of residues N-terminal to the DM15 region has been shown to decrease LARP1 affinity for TOP mRNA[35]. Conversely, phosphomietic mutation of a cluster of residues C-terminal to the DM15 region has been shown to increase LARP1 affinity for TOP mRNA[35]. Given that there is data to support a repressive and a stimulatory role of LARP1 in TOP mRNA translation, allosteric regulation of LARP1 likely plays a part[2, 49, 82-84]. Further characterizing the effects of LARP1 phosphorylation on protein dynamics could elucidate the "molecular switch", flipped by phosphorylation, that reconciles the repressive/stimulatory/stabilizing narratives[84].

The LARP1 DM15 R824W point mutation found in breast cancer patients is also a putative source of an allosteric signal, as described in Chapter 4[113]. We showed that this mutation alters both the cap- and +1-binding pocket dynamics. However, the effect of the R824W mutation on LARP1 DM15 TOP mRNA affinity has not yet been characterized. This mutation could potentially alter the specificity of LARP1 DM15 for RNAs containing a different +1 nucleotide; this would change the class of transcripts that LARP1 binds and translationally regulates. The pathogenic effect of this mutation in breast cancer has also not yet be characterized, and determining the

mutation's effect on LARP1 RNA binding would shed light on its functional significance. Further, the effect of LARP1 R824W on human cell lines could be explored.

LARP1 DM15/La-Module binding may be the source of an autoregulatory form of allosteric regulation, as detailed in Chapter 3. We found evidence that residues within LARP1 DM15 coevolved with residues within the La-Module, indicating they may bind one another. LARP1 DM15 PQ810 and PR824 both contain residues that coevolve with residues in the La-Module. An interaction with the La-Module at either of these pockets may cause an allosteric change in protein dynamics, which could alter the ability of LARP1 to bind TOP mRNA. R824 is putatively allosterically linked to R847 in the +1-binding pocket, and PQ810 needs to be further explored for any potential allosteric relationship with the DM15 RNA binding pockets.

6.1.1 Future Direction: Investigate the Effect of Phosphorylation on LARP1 DM15 Dynamics

The effects of phosphorylation on LARP1 protein dynamics and the mechanism of RNA binding are not known. The existing crystal structures of LARP1 do not include any of the characterized mTORC1 phosphorylation sites[79, 80]. Crystal structures of extended constructs of LARP1 DM15 containing these phosphorylation sites are needed in order to investigate the effects of phosphorylation on protein dynamics. After these structures are solved, with and without phosphorylation, molecular dynamics simulations could be used to investigate any allosteric effects. The phosphorylation modifications could also be modelled *in silico*, if only the unphosphorylated structure can be solved.

The dynamics of the cap- and +1-binding pocket could be compared between simulations with phosphorylated and unphosphorylated LARP1. Difference contact network analysis could be

used to further probe the effect of phosphorylation on contacts between residues and between residue communities within LARP1[150]. Observed alterations in the dynamics of the cap- and +1-binding pockets could explain the mechanism behind the shifts in LARP1 TOP mRNA affinity upon phosphorylation at specific sites[35].

6.1.2 Future Direction: Investigate the Effect of R824W on LARP1 DM15 mRNA Binding and Translation Regulation

We provide evidence in Chapter 4 that R824W is an allosteric mutation given that it alters the dynamics of the RNA binding pockets, but we need to characterize its full biological effect *in vitro* and in cells. The R824W mutation would be introduced into LARP1 DM15 with site-directed mutagenesis[175]. Following recombinant protein expression in *E. coli* and purification, LARP1 DM15 R824W would be used in electrophoretic mobility shift assays (EMSAs) to determine the effect of the mutation on LARP1 DM15/TOP mRNA affinity[79, 80, 176]. A switch in specificity for another +1 nucleotide would also be investigated with EMSAs by determining the affinity of LARP1 DM15 R824W for RNAs containing alternative +1 nucleotides, followed up by competition assays[79, 80]. Alternatively, the R824W mutation could shift LARP1 DM15 affinity for 3' UTRs, as LARP1 has be previously shown to bind 3' UTRs of mRNAs[84, 86]. The previously described in vitro experiments could similarly be carried out to determine if there is a shift in affinity of LARP1 DM15 for specific mRNA 3' UTRs in the presence of the R824W mutation.

The effect of the mutation on global translation would be investigated by expressing LARP1 R824W in LARP1 knockout HEK293T cells; polysome profile analysis would then be used to look at the effect on polysome abundance to assess the global effect of the mutation as compared to matched HEK293T cells[83]. RNA would be extracted from fractions and RT-qPCR

would be used to determine the effect of the mutation specifically on TOP mRNA translation. If the *in vitro* experiments determine a shift in affinity to a different +1 nucleotide, the effect on translation for specific RNAs with the appropriate +1 nucleotide would also be checked with RT-qPCR. Concurrently, the effect of LARP1 R824W on cell viability would be explored; any observed effects would clarify any pathogenic function of the mutation.

If these the LARP1 R824W mutation does cause a shift in +1 nucleotide specificity for LARP1 R824W, functionally LARP1 would regulate the translation of an entirely different class of transcripts. Most transcripts start with a purine, not a pyrimidine like TOP mRNAs (+1C). A shift in specificity to a purine would suggest LARP1 R824W is likely directly regulating the translation of a larger percentage of transcripts in the cell, compared to WT LARP1. However, LARP1 specificity would shift away from transcripts encoding ribosomal proteins; this would likely decrease the impact that LARP1 has on global translation. Therefore, the exact effect of the LARP1 R824W mutation on the cell is difficult to predict, highlighting the importance of the aforementioned experiments.

6.1.3 Future Direction: Investigate LARP1 DM15/La-Module Autoregulation and its Effect on TOP mRNA Binding and Translation Regulation

LARP1 La-Module binding to the DM15 region could be a source of autoregulatory allostery affecting TOP mRNA binding. Understanding if and how the La-Module and DM15 region interact would be critical in identifying the mechanism behind any allosteric regulation imparted by the binding event. The potential LARP1 DM15/La-Module interaction would be explored *in vitro* with cross-linking followed up by mass spectrometry to characterize the potential binding interface, based on the cross-linked residues[177]. Information regarding intra-domain

crosslinks within the La-Module would also be used to inform the creation of a higher confidence La-Module homology model (if a crystal structure is still unavailable). Protein-protein docking could then be used to investigate the interaction interface at a higher resolution than provided by cross-linking mass spectrometry; the docking would be guided by the constraints identified in the cross-linking mass spectrometry but would also provide information on how the non-cross-linked residues interact[178]. To validate this model, LARP1 DM15 or La-Module mutants that obliterate this interaction would need to be cloned, expressed, purified, and subjected to cross-linking experiments.

If the aforementioned experiments support the DM15/La-Module interaction, then it would be crucial to determine if there is any effect of this binding event on LARP1 TOP mRNA binding. Both the La-Module and DM15 region can bind TOP mRNAs, and so the binding event's effect on each would be investigated. La-Module and DM15 mutants incapable of binding TOP mRNA would be needed to probe the effect on each component of the complex individually. The DM15 R840E/Y883A double mutation has been shown to obliterate TOP mRNA binding, but no mutation is known to do so for the La-Module. Scanning mutagenesis would be used in order to identify this mutant. EMSAs would then be used to determine any effect on TOP mRNA affinity with DM15/La-Module complex.

6.2 Drugging LARP1

High levels of LARP1 have been observed in several carcinomas, but the exact role that LARP1 plays in cancer is unknown[71, 85, 86, 89]. Although high expression of LARP1 has been correlated with metastasis, poor prognosis, chemotherapeutic resistance and invasion, LARP1

DM15 has been shown to repress translation of TOP mRNAs[2, 71, 85-88, 94]. This is seemingly counterintuitive, but studies have also shown that LARP1 DM15 can stabilize TOP mRNAs, and one study indicated that LARP1 stimulates TOP mRNA translation[2, 81-84]. A more thorough understanding of LARP1's role in TOP mRNA translation is required for a focused therapeutic strategy.

Given the complicated role of LARP1 in translation regulation and how much is still unknown, it is unclear whether the goal should be to inhibit or enhance RNA binding when drugging LARP1 DM15. It would be best to explore both of these strategies, inhibition and enhancement, because of the current gaps in knowledge regarding the role of LARP1 in translation regulation. Identified compounds could even be used to further our understanding of LARP1's function; they could be used in *in vitro* and in cell assays further characterizing the effect of LARP1 on translation regulation and TOP mRNA binding. LARP1 has also recently been found to repress SARS-CoV-2 replication, and was found to bind genomic and subgenomic RNAs[96, 97]. In this case it would be crucial to develop a small molecule that enhances LARP1 RNA binding.

It would be beneficial to explore drugging LARP1 DM15 by separately targeting several different pockets; this would increase the probability of successfully achieving the desired therapeutic effect, and offer a variety of small molecule probes for future experiments. We identified five compounds as putative hits in Chapter 5; these compounds target either the cap-, cap- and +1-, or α7-α8 bridge binding pockets. Only the compound targeting the α7-α8 bridge binding pocket could potentially enhance RNA binding via an allosteric effect; the other compounds would sterically occlude RNA binding. Since directly targeting the cap- or +1-binding pockets might not offer the desired therapeutic effect, PY880, PR824, and PQ810 could be viable allosteric alternatives to pharmacologically target.

Currently to our knowledge there is no central repository for LARP1 cancer mutations, nor a detailed accounting of the prevalence of mutations in patient populations for specific cancers. We showed the LARP1 R824W breast cancer-associated mutation alters the dynamics of the cap- and +1-binding pockets, but we do not know how pathogenically meaningful these findings are because we do not know how prevalent this mutation is in breast cancer patient populations. Collating the existing data, and potentially supplementing it with expanded studies is needed to know the impact. This is crucially important to know how to best allocate resources in drugging LARP1. For example, if you target a novel cancer mutant pocket environment, you would want to know approximately what percentage of the patient population the resulting compound would be relevant to.

The LARP1 La-Module which includes two RNA binding domains, a La motif and an RNA recognition motif, could also be pharmacologically targeted. The La-Module also binds to TOP mRNA and has been hypothesized to aid in TOP mRNA circularization and translation[90]. No structures of the LARP1 La-Module currently exist, however obtaining a structure would open a whole new path to drugging LARP1.

6.2.1 Future Direction: Determine the Prevalence of LARP1 Cancer Mutations in Patient Populations

Existing databases of mutations observed in various cancers (e.g. COMSIC, Bravo, TCGA) do not currently have a consensus on the prevalence of recorded LARP1 cancer mutations[113, 179, 180]. This makes it difficult to predict the pathogenic importance of any individual mutation, such as R824W; the accurate prevalence of a mutation would also allow for the prioritization of developing a targeted therapy. Therefore, I would collate all of the existing LARP1 cancer

139

mutation data in to a central repository. The repository would be a valuable resource for any investigation into therapeutic development or biological function of LARP1.

Some stumbling blocks may arise in the creation of the repository, such as the availability of patient IDs. For example, Bravo does not, to our knowledge, provide patient IDs for the data provided, and this could result in duplicate records of sequencing data[179]. It may also be discovered during the collating process that not enough data currently exists to provide a robust analysis; in this case, I would collaborate with other researchers to acquire more whole exome sequencing data of patients with carcinomas where LARP1 is highly expressed. This would be an ongoing effort to keep the databases up to date and as relevant as possible.

6.2.2 Future Direction: Drugging LARP1 DM15

6.2.2.1   Characterization of the five putative hits

We must further characterize the compounds identified as putative hits in chapter 5 to evaluate their therapeutic potential. I would determine the affinity of the LARP1 DM15 region for all five-compounds using isothermal titration calorimetry or microscale thermophoresis; this will allow us to validate that they bind LARP1 DM15 and prioritize them based on affinity[181, 182]. Concurrently, I would setup co-crystallization experiments to determine the structures of the compounds bound to LARP1 DM15, to validate that they are binding in the pockets we predicted. The co-crystal structures could also be used for further optimization of the compounds; the observed LARP1 DM15-compound interactions can be used to inform changes that increase affinity and specificity of the compounds. The appropriate EC_{50} or IC_{50} for the compounds with respect to LARP1 DM15 TOP mRNA binding would be determined via EMSA; this measure will be used to further prioritize compounds based on their functional effect. Once the compounds are

optimized to low µM or nM affinity, I would characterize their effect on global and TOP mRNA translation in HEK293T[82, 83]. Polysome profiling along with RNA extraction and RT-qPCR would be used after treating the cells with the compound, using DMSO and torin1 as controls[81, 83]. I would also determine the compound's effects on cancer cell lines, with the focus on viability and migration[87, 183, 184]. If the five putative hits cannot be optimized efficiently, bind regions of the DM15 region that are not relevant to mRNA binding, or fail for other reasons, I would revisit the docked ranked listed of compounds. I could also revisit the original pharmacophore ranked listed and prioritize compounds solely based on the pharmacophore, as an alternative approach.

6.2.2.2 Drugging PQ810, PR824, PY880 and further characterization of biological function

The biological functions of PQ810, PR824, and PY880 must be characterized as they could also be pharmacologically targeted. Residues in PQ810 coevolved with residues in the La-Module; if a DM15/La-Module interaction at this pocket causes an allosteric effect, then a compound could disrupt this interaction by binding to PQ810. I could also look to develop a compound that recapitulates the putative allosteric effect of La-Module binding to PQ810. We showed that PY880 and PR824 have putative allosteric relationships with the cap- and +1-binding pockets, respectively. Compounds targeting these pockets might cause an allosteric effect that could enhance TOP mRNA binding to DM15. If they inhibit TOP mRNA binding, they could be used if targeting the cap- and/or +1-binding pocket leads to non-specific binders.

I would run druggability simulations, using structures with the most voluminous states of PQ810, PR824, and PY880 as starting structures, as described in Chapter 5. Pharmacophores that map to each pocket would be created from the druggability simulations. I would then search drug databases using the pharmacophores as filters, and then run virtual screens with the remaining

compounds. I would then further characterize these compounds *in vitro* and in cells, the same way as stated for the current five putative hits. The assays used to determine the effects of the compounds would also help determine the biological function of PQ810, PR824, and PY880. Any effect on LARP1 DM15 TOP mRNA binding would be through an allosteric effect tied to the respective pocket (as long as the binding pockets for each compound is validated *in vitro*). Further experiments could be done to see if a compound targeting PR810 or PR824 could disrupt a potential DM15/La-Module interaction (assuming there is evidence for this interaction determined with experiments in 6.1.3).

6.2.2.3 A machine learning approach to a fragment screen hit optimization

It is possible the pharmacophore and virtual screen approaches yield no useful results, and in that case an *in vitro* fragment screen could be used. LARP1 DM15 has been crystalized several times and LARP1 DM15 crystals can be soaked with small molecule fragments, such as those in the X-Chem compound libraries. I would then resolve the structures of LARP1 DM15 bound to these fragments. Although these would just be small molecule fragments, it is possible to optimize compounds based off of them. One approach would be through a visual inspection and determining how one could extend the fragments to further exploit pockets of LARP1 DM15. A more recently developed approach would be to utilize machine learning implementations of small molecule optimization; DeepFrag is one such tool[185]. It takes a protein structure with a small molecule bound, and then one chooses the linking atom from which the fragment branches. Given its training set and the available chemical environment of your protein of interest, DeepFrag "guesses" which moiety would be the best compliment to the receptor[185]. These new compounds, based on the fragment screen and extended by DeepFrag, could then be tested *in vitro* and in cells.

6.2.3 Future Direction: Drugging LARP1 La-Module

The LARP1 La-Module which includes two RNA binding domains, a La motif and an RNA recognition motif, also binds to TOP mRNA and has been hypothesized to aid in TOP mRNA circularization and translation[90]. Specifically, the La-Module has been shown to bind poly(A) RNA and TOP motif containing RNA[90]. Given this separate function from DM15, the La-Module could also be an attractive pharmacological target in carcinomas in which LARP1 is highly expressed. Unfortunately, no crystal structure for the LARP1 La-Module currently exists and previous attempts to create homology models have not given rise to models with high confidence levels[8]. Further exploring new crystallization conditions to obtain a crystal structure is imperative in pursuing the La-Module as a pharmacological target. Other human proteins also contain La-motifs and RRMs, and so one would need to take care in ensuring compound specificity[70].

A co-crystal structure of the La-Module bound to poly(A) RNA and/or TOP motif containing RNA would allow for the characterization of the RNA binding surface(s). After the RNA binding surfaces are identified, druggability simulations would be used to identify pharmacophores targeting that surface. A similar approach of pharmacophore development as used in Chapter 5 and a virtual screen would be used to rank compounds. The same *in vitro* and in-cell compound optimization and experiments from 6.2.1 would be used to further characterize the effects of any compounds.

It could also be possible that a drug cocktail containing compounds targeting both the La-Module and the DM15 region is required for the desired therapeutic effect; one could also test this *in vitro* an in cells. Full length LARP1 would be incubated simultaneously with compounds targeting both the La-Module and the DM15 region; EMSAs would then be used to obtain an EC_{50}

or IC$_{50}$ with TOP mRNA. The in-cell experiments would also be done after simultaneously treating with both La-Module and DM15 region targeting compounds.

6.3 Concluding Remarks

This thesis covers the investigation into LARP1 DM15 dynamics, novel pocket characterization, allosteric regulation, and small molecule ligand identification. We found that the LARP1 DM15 cap-binding pocket transitions between a ligand-ready and collapsed state; these two states were correlated with the secondary structure of the adjacent α7-α8 bridge. Pockets PQ810, PR824, and PY880 were characterized, and the latter two were found to have a putative allosteric connection to the +1- and cap-binding pockets, respectively. Coevolution analysis found potential LARP1 DM15/La-Module binding sites at PQ810 and PR824. The LARP1 DM15 R824W breast cancer-associated mutation was simulated and found to affect the dynamics of both the cap- and +1-binding pockets; R824W increased the probability of an alternative +1-binding pocket conformation, potentially altering the +1 specificity of the mutant. We also identified small molecules that putatively bind the cap-, cap- and +1-, or α7-α8 bridge binding pockets; these small molecules would likely alter the ability of LARP1 DM15 to bind TOP mRNA and could potentially be optimized for therapeutic use in carcinomas that have high expression of LARP1.

Further exploration of each of these findings is required to more fully understand the biological relevance and impact of LARP1's role in translational regulation. Crystal structures of extended constructs beyond just LARP1 DM15 are required to further investigate the dynamics of all of LARP1 and to elucidate the impact of allosteric regulation, such as the mTORC1 phosphorylation of LARP1. Small molecules that bind PQ810, PR824, and PY880 could be used

to further characterize the pockets' biological function, and any allosteric relationships with the cap- and +1-binding pockets. The potential LARP1 DM15/La-Module interaction needs to be probed *in vitro* for validation and any functional effect on TOP mRNA binding. Similarly, the LARP1 DM15 R824W mutant also needs *in vitro* and in cell characterization and any effect on +1 nucleotide binding preference or cell viability must be examined. Lastly, the effect of the identified small molecules binding to LARP1 DM15 needs to be interrogated *in vitro* and in cells; the effects would provide insight into the therapeutic viability of the small molecules and how they could be further optimized.

Appendix A ProteinVR: Web-Based Molecular Visualization in Virtual Reality

Adapted from: "ProteinVR: Web-based molecular visualization in virtual reality." Cassidy, K. C.; Šefčík, J.; Raghav, Y.; Chang, A.; Durrant, J. D., ProteinVR: Web-based molecular visualization in virtual reality. PLOS Computational Biology 2020, 16 (3), e1007747. https://doi.org/10.1371/journal.pcbi.1007747[186]

Comparisons with other programs performed by K.C.C. and J.D.D. 3D environments created by K.C.C, J.D.D, and J.S. Testing performed by K.C.C, J.S, Y.R, A.C, and J.D.D. Writing (original draft) performed by K.C.C and J.D.D. Writing (review and editing) performed by K.C.C, J.S, Y.R, A.C, and J.D.D. Software development performed by J.D.D.

Appendix A.1 Design and Basic Usage

Popular solutions for visualizing biological structures include 2D pictures and pseudo 3D representations (projecting 3D models onto 2D screens) with computer graphics[133, 187-190]. While these solutions allow for a great deal of manipulation and visual analysis, they have notable drawbacks. For example, immersion and depth are lacking. On the other hand, virtual reality (VR) stereoscopic 3D addresses both these limitations[191, 192]. The 3D virtual environments immerse the viewer into the visualization and can add context to educational lessons[193-196]. True stereoscopic depth perception aids in the discernment of key features of biological structures such as distances and interactions.

Although virtual reality provides these benefits, it can be hard to access. Virtual reality headsets and accompanying computers (if needed) can be cost prohibitive; creating the VR scenes and importing the biological structures can require several pieces of software. ProteinVR, a web browser-based approach to VR, is accessible to anyone with a smartphone, and a low-cost accompanying VR headset such as Google Cardboard. In addition, the scenes can be viewed in a pseudo 3D implementation via a web browser.

At the ProteinVR webpage, one types in the PDB ID of interest and selects the desired 3D environment from the premade list. It is also possible to upload a PDB file from a computer. After these selections are made the scene is loaded onto the screen, and all the information is stored in the URL (when not using local files). To navigate to the same scene or share it, one need only copy and paste the URL. There are also representation and selection tools that allow you to choose how the biological structure is visualized. Any changes made to the visualization is updated and encoded in the URL.

Appendix A.2 Examples of Use
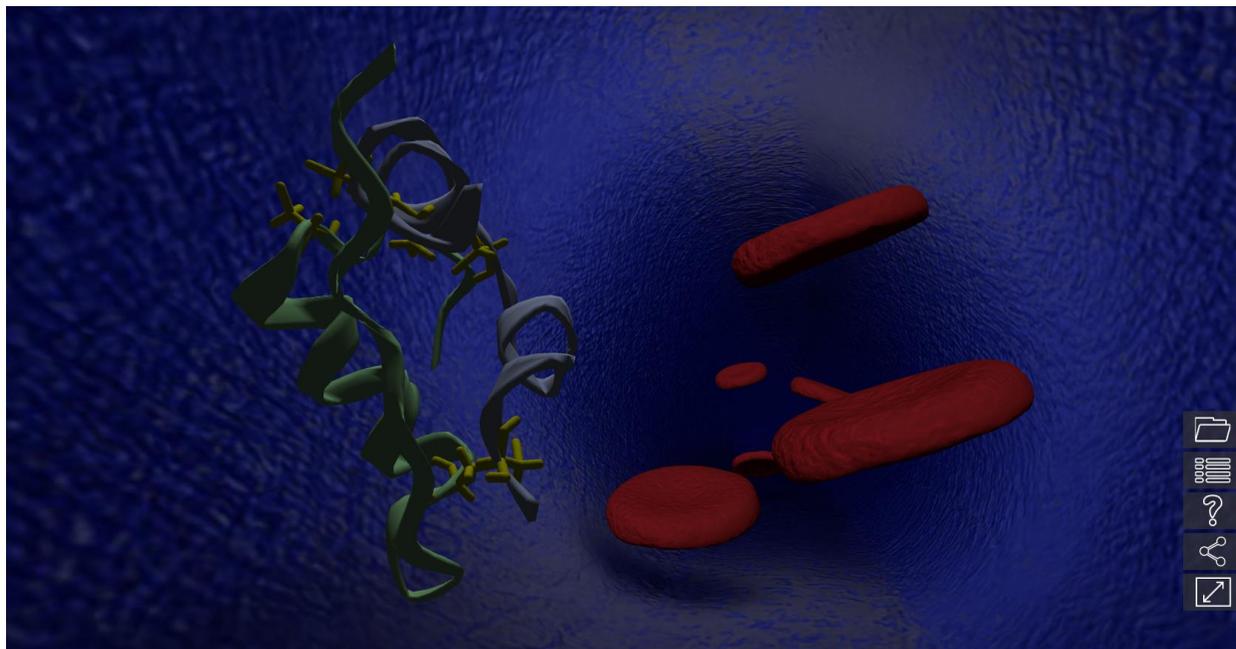
Appendix Figure 1. Example of an educational use for ProteinVR: visualization of insulin in the bloodstream. High school and undergraduate biology classes cover how insulin works in the body. A key step of insulin transport is its release into the bloodstream. ProteinVR can be used to create a visualization of this event and to allow students to investigate the structure of insulin. The representation and color selection allow one to highlight the interacting A and B chains, blue and green, respectively. Cysteine residues are shown in stick representation and colored yellow to highlight the key disulfide bonds (PDB ID: 2HIU)[197].

Potential use cases for ProteinVR include education, research analysis, and research collaboration. High school and undergraduate biology classes cover protein structure and function as part of the curriculum. 2D images and pseudo 3D computer graphic representations of protein structures are commonly used to present these topics to students; however, a stereoscopic 3D representation offers more informative perspective and depth. These advantages allow one to investigate characteristics, such as key interacting residues, with a greater understanding of their implications in 3D space.

For example, many high school and undergraduate classes go over how insulin works in the body. ProteinVR can be used to visualize a key part of this topic: the release of insulin into the blood stream (Appendix Figure 1). With ProteinVR's representation and color-selection tools, one can highlight the A and B chains along with the residues involved in the key disulfide bonds between the two chains. One can then virtually "walk around" the protein and interrogate these features in stereoscopic 3D, allowing for a detailed and accurate understanding of how the A and B chain are "glued" together through these disulfide bonds. Another option would be to focus the representation on secondary structure. The representation and color-selection tools allow you to highlight the atoms involved in the alpha helix hydrogen bonds. Due to the prevalence of smart phone ownership, an educator would only need to acquire the accompanying low-cost VR headsets (such as the Google Cardboard).

Researchers would also find a use for ProteinVR. Stereoscopic 3D visualization aids in investigating protein structures and/or protein-ligand interactions. A sense of depth is crucial for identifying key interactions. One can easily miss an interaction due to view angle or a poor implementation of simulated depth (such as fog) in pseudo 3D representations generated by popular programs such as Pymol or VMD. ProteinVR avoids these pitfalls with the afforded depth perception of stereoscopic 3D, which could allow for a more expedient and thorough analysis.

Finally, ProteinVR is perfect for research collaborations with its easily shared visualizations via URLs. Since the state of the visualization is all encoded in the URL, one can communicate discoveries and insights with collaborators without any software downloads. This makes the work in a collaboration less cumbersome and expedites sharing. ProteinVR has a low barrier to entry because one only needs a low-cost VR headset (beyond the smart phone), and no

computer workstation is required. This means that collaborators who are not in a computational field can easily access the visualization, without the need for vast computational resources.

Appendix A.3 Comparison with Other Programs

Molecule-to-Mesh Pipelines

ProteinVR makes setting up VR molecular visualizations particularly easy. In contrast, some other VR programs rely on complex software pipelines that require users to install (and master) third-party modeling programs such as Blender (Blender Foundation) and Unity (Unity Technologies). Users must setup molecular representations (e.g., ribbon, stick, surface) during the initial modeling stage, making it impossible to change the representation in real-time VR.

The open-source BlendMol plugin for Blender is one example of this effective but difficult-to-manage approach[131, 198]. BlendMol/Blender can produce photorealistic images of protein structures that are well suited for scientific publication and educational outreach. Third-party Blender plugins can also export BlendMol models to VR-compatible formats. But the Blend- Mol method for preparing VR models is far from automated and requires some expertise in 3D modeling.

RealityConvert, like BlendMol, provides a molecule-to-mesh pipeline that generates molecular meshes for VR and AR scenes[199]. An easy-to-use web app helps overcome some barriers to use. But the web app only accepts very small molecules (< 200 lines). Processing larger molecules requires the command-line version and its four dependencies: PyMOL, Blender, Open Babel, and Molconvert (ChemAxon)[169, 187, 198]. Many other VR and AR approaches for molecular

visualization involve similarly challenging software pipelines [19, 21, 41][200-202]. In contrast, ProteinVR requires no download or dependencies and so is more accessible.

Desktop Applications

A number of desktop applications enable VR molecular visualiza- tion directly, without requiring a complex pipeline. These desktop programs often limit their compatibility to high-end VR devices[191, 201, 203, 204]. In contrast, ProteinVR is generally more accessible because it supports a broad range of VR headsets as well as non-VR fallback approaches such as device-orientation-based viewing. This broad support is possible because ProteinVR relies on the WebVR API, which standardizes the way VR-enabled websites inter- act with various devices, as well as the BabylonJS JavaScript game engine, which provides a broad range of video-game-style navigation schemes. As a web-based app, ProteinVR also requires no download or installation, further improving accessibility.

That having been said, desktop programs that cater to high-end VR headsets are able to implement useful features that ProteinVR currently lacks. Molecular Rift is a good example of such a desktop program[191]. This innovative, open-source VR application allows users to navigate molecular structures without VR controllers, using hand gestures. The commercial program Nanome (Nanome Inc.) is a second notable example. Nanome's easy-to-use and detailed user interface permits not only molecular visualization, but also molecular manipulation (e.g., in silico mutagenesis). The free version of Nanome does come with some important limitations, however. For example, VR molecular scenes created with the free version are entirely public. In fact, as we were testing Nanome, another user joined our room and was able to observe our activities.

Other Web Applications

Recognizing the advantages of the web-based approach, others have also explored online VR molecular-visualization systems [205, 206]. One example is iview[206]. Though the iview website includes a "virtual reality" button, this button was not functional on any of the browsers we tested. The iview source code does make reference to WebVR, so perhaps it is the user interface, rather than the underlying codebase, that is broken. We note also that the iview server went offline after our initial tests, though the connectivity problem may be temporary. Regardless, ProteinVR provides additional features—including 3D environments and device-orientation mode—that iview and other programs currently lack.

Unity Game Engine

The Unity game engine (Unity Technologies) warrants specific men- tion because it powers several desktop VR applications, including Molecular Rift, described above, and Molecular Zoo, a program for teaching young students about biomolecules[191, 207]. The open-source library UnityMol even enables on-the-fly molecular-mesh generation in Unity apps, much as 3Dmol.js does for ProteinVR[208].

Unity has several advantages over the BabylonJS game engine behind ProteinVR. Its advanced editor greatly simplifies development, and its online community has developed many add-ons (both free and commercial) that allow developers to easily add the specific features that their application requires. If desktop Unity applications are properly optimized, their performance also surpasses that of any web-based app because most browsers cap graphics updates at 60 frames per second. Finally, because the WebVR standard is still evolving, Unity applications are arguably more stable, at least for the time being.

Despite these advantages, we built ProteinVR using BabylonJS because it is particularly well suited for web apps. Unity applications can be compiled to run in the browser, but they are

almost always far larger than the equivalent BabylonJS app, requiring more time and band- width to download. Unity also lacks official support for browser-based apps on mobile devices, and the BabylonJS approach to WebVR is much more straightforward than its Unity counter- part. Because BabylonJS is itself written in JavaScript, integration with web technologies such as WebRTC and the HTML5 DOM (e.g., buttons, popups, menus, etc.) is also much easier. Finally, the Unity engine is closed source, and free use requires Unity-specific branding.

Bibliography

1.   Baßler, J.; Hurt, E., Eukaryotic Ribosome Assembly. *Annual Review of Biochemistry* **2019,** *88* (1), 281-306.

2.   Fonseca, B. D.; Lahr, R. M.; Damgaard, C. K.; Alain, T.; Berman, A. J., LARP1 on TOP of ribosome production. *Wiley Interdiscip Rev RNA* **2018**, e1480.

3.   Warner, J. R., The economics of ribosome biosynthesis in yeast. *Trends in Biochemical Sciences* **1999,** *24* (11), 437-440.

4.   Dever, N. S. a. T. E., Eukaryotic translation initiation factors and regulators *Current Opinion in Structural Biology* **2003,** *13*, 56-63.

5.   Hinnebusch, A. G.; Lorsch, J. R., The Mechanism of Eukaryotic Translation Initiation: New Insights and Challenges. *Cold Spring Harbor Perspectives in Biology* **2012,** *4* (10), a011544-a011544.

6.   Dever, T. E.; Green, R., The Elongation, Termination, and Recycling Phases of Translation in Eukaryotes. *Cold Spring Harbor Perspectives in Biology* **2012,** *4* (7), a013706-a013706.

7.   Lahr, R., Elucidating the molecular mechanism of LARP1 in post-transcriptional regulation of TOP transcripts. *University of Pittsburgh* **2019**.

8.   Al-Ashtal, H., Making Ends Meet: Role of the LARP1 La-Module in TOP mRNA Recognition. *University of Pittsburgh* **2020**.

9.   Panda, A. C.; Martindale, J. L.; Gorospe, M., Polysome Fractionation to Analyze mRNA Distribution Profiles. *Bio Protoc* **2017,** *7* (3).

10.   Saxton, R. A.; Sabatini, D. M., mTOR Signaling in Growth, Metabolism, and Disease. *Cell* **2017,** *168* (6), 960-976.

11.   Pelletier, J.; Thomas, G.; Volarević, S., Ribosome biogenesis in cancer: new players and therapeutic avenues. *Nature Reviews Cancer* **2018,** *18* (1), 51-63.

12.   Ramanathan, A.; Robb, G. B.; Chan, S.-H., mRNA capping: biological functions and applications. *Nucleic Acids Research* **2016,** *44* (16), 7511-7526.

13.   Shimotohno, K.; Kodama, Y.; Hashimoto, J.; Miura, K. I., Importance of 5'-terminal blocking structure to stabilize mRNA in eukaryotic protein synthesis. *Proceedings of the National Academy of Sciences* **1977,** *74* (7), 2734-2738.

14.   Shatkin, A., Capping of eucaryotic mRNAs. *Cell* **1976,** *9* (4), 645-653.

15.   Dreyfus, M.; Régnier, P., The Poly(A) Tail of mRNAs. *Cell* **2002,** *111* (5), 611-613.

16.   Zhao, J.; Hyman, L.; Moore, C., Formation of mRNA 3' ends in eukaryotes: mechanism, regulation, and interrelationships with other steps in mRNA synthesis. *Microbiol Mol Biol Rev* **1999,** *63* (2), 405-45.

17.   Keller, W.; Bienroth, S.; Lang, K. M.; Christofori, G., Cleavage and polyadenylation factor CPF specifically interacts with the pre-mRNA 3' processing signal AAUAAA. *EMBO J* **1991,** *10* (13), 4241-4249.

18.   Wells, S. E.; Hillner, P. E.; Vale, R. D.; Sachs, A. B., Circularization of mRNA by eukaryotic translation initiation factors. *Mol Cell* **1998,** *2* (1), 135-40.

19.   Dong, Z.; Zhang, J. T., Initiation factor eIF3 and regulation of mRNA translation, cell growth, and cancer. *Crit Rev Oncol Hematol* **2006,** *59* (3), 169-80.

20.   Xie, J.; Kozlov, G.; Gehring, K., The "tale" of poly(A) binding protein: the MLLE domain and PAM2-containing proteins. *Biochim Biophys Acta* **2014,** *1839* (11), 1062-8.

21.   Alekhina, O. M.; Terenin, I. M.; Dmitriev, S. E.; Vassilenko, K. S., Functional Cyclization of Eukaryotic mRNAs. *Int J Mol Sci* **2020,** *21* (5).

22.   Jackson, R. J.; Hellen, C. U. T.; Pestova, T. V., The mechanism of eukaryotic translation initiation and principles of its regulation. *Nature Reviews Molecular Cell Biology* **2010,** *11* (2), 113-127.

23.   Villa, N.; Do, A.; Hershey, J. W. B.; Fraser, C. S., Human Eukaryotic Initiation Factor 4G (eIF4G) Protein Binds to eIF3c, -d, and -e to Promote mRNA Recruitment to the Ribosome. *Journal of Biological Chemistry* **2013,** *288* (46), 32932-32940.

24.   Pratt, G.; Galpine, A.; Sharp, N.; Palmer, S.; Clemens, M. J., Regulation of in vitro translation by double-stranded RNA in mammalian cell mRNA preparations. *Nucleic Acids Res* **1988,** *16* (8), 3497-510.

25.   Hellen, C. U. T., Translation Termination and Ribosome Recycling in Eukaryotes. *Cold Spring Harbor Perspectives in Biology* **2018,** *10* (10), a032656.

26.   Prakash, V.; Carson, B. B.; Feenstra, J. M.; Dass, R. A.; Sekyrova, P.; Hoshino, A.; Petersen, J.; Guo, Y.; Parks, M. M.; Kurylo, C. M.; Batchelder, J. E.; Haller, K.; Hashimoto, A.; Rundqivst, H.; Condeelis, J. S.; Allis, C. D.; Drygin, D.; Nieto, M. A.; Andäng, M.; Percipalle, P.; Bergh, J.; Adameyko, I.; Farrants, A.-K. Ö.; Hartman, J.; Lyden, D.; Pietras, K.; Blanchard, S. C.; Vincent, C. T., Ribosome biogenesis during cell cycle arrest fuels EMT in development and disease. *Nature Communications* **2019,** *10* (1).

27.   Penzo, M.; Montanaro, L.; Treré, D.; Derenzini, M., The Ribosome Biogenesis—Cancer Connection. *Cells* **2019,** *8* (1), 55.

28. Meyuhas, O.; Kahan, T., The race to decipher the top secrets of TOP mRNAs. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms* **2015**, *1849* (7), 801-811.

29. Levy, S.; Avni, D.; Hariharan, N.; Perry, R. P.; Meyuhas, O., Oligopyrimidine tract at the 5' end of mammalian ribosomal protein mRNAs is required for their translational control. **1991**, *88* (8), 3319-3323.

30. Fonseca, B. D.; Smith, E. M.; Yelle, N.; Alain, T.; Bushell, M.; Pause, A., The ever-evolving role of mTOR in translation. *Semin Cell Dev Biol* **2014**.

31. Hare, S. H.; Harvey, A. J., mTOR function and therapeutic targeting in breast cancer. *Am J Cancer Res* **2017,** *7* (3), 383-404.

32. Zoncu, R.; Efeyan, A.; Sabatini, D. M., mTOR: from growth signal integration to cancer, diabetes and ageing. *Nat Rev Mol Cell Biol* **2011,** *12* (1), 21-35.

33. De La Cruz López, K. G.; Toledo Guzmán, M. E.; Sánchez, E. O.; García Carrancá, A., mTORC1 as a Regulator of Mitochondrial Functions and a Therapeutic Target in Cancer. *Frontiers in Oncology* **2019,** *9*.

34. Kang, S. A.; Pacold, M. E.; Cervantes, C. L.; Lim, D.; Lou, H. J.; Ottina, K.; Gray, N. S.; Turk, B. E.; Yaffe, M. B.; Sabatini, D. M., mTORC1 Phosphorylation Sites Encode Their Sensitivity to Starvation and Rapamycin. *Science* **2013,** *341* (6144), 1236566-1236566.

35. Jia, J.-J.; Lahr, R. M.; Solgaard, M. T.; Moraes, B. J.; Pointet, R.; Yang, A.-D.; Celucci, G.; Graber, T. E.; Hoang, H.-D.; Marius; Pena, I. A.; Hollensen, A. K.; Smith, E. M.; Chaker-Margot, M.; Anton, L.; Dajadian, C.; Livingstone, M.; Hearnden, J.; Wang, X.-D.; Yu, Y.; Maier, T.; Damgaard, C. K.; Berman, A. J.; Alain, T.; Fonseca, B. D., mTORC1 promotes TOP mRNA translation through site-specific phosphorylation of LARP1. *Nucleic Acids Research* **2021,** *49* (6), 3461-3489.

36. Proud, X. W. a. C. G., mTORC1 signaling: what we still don't know. *Journal of Molecular Cell Biology* **2010**, 1-16.

37. Hsu, P. P.; Kang, S. A.; Rameseder, J.; Zhang, Y.; Ottina, K. A.; Lim, D.; Peterson, T. R.; Choi, Y.; Gray, N. S.; Yaffe, M. B.; Marto, J. A.; Sabatini, D. M., The mTOR-regulated phosphoproteome reveals a mechanism of mTORC1-mediated inhibition of growth factor signaling. *Science* **2011,** *332* (6035), 1317-22.

38. Bond, P., Regulation of mTORC1 by growth factors, energy status, amino acids and mechanical stimuli at a glance. *Journal of the International Society of Sports Nutrition* **2016,** *13* (8).

39. Dibble, C. C.; Cantley, L. C., Regulation of mTORC1 by PI3K signaling. *Trends in Cell Biology* **2015,** *25* (9), 545-555.

40. Yip, C. K.; Murata, K.; Walz, T.; Sabatini, D. M.; Kang, S. A., Structure of the human mTOR complex I and its implications for rapamycin inhibition. *Mol Cell* **2010,** *38* (5), 768-74.

41. Feldman, M. E.; Apsel, B.; Uotila, A.; Loewith, R.; Knight, Z. A.; Ruggero, D.; Shokat, K. M., Active-Site Inhibitors of mTOR Target Rapamycin-Resistant Outputs of mTORC1 and mTORC2. *PLoS Biology* **2009,** *7* (2), e1000038.

42. Weber, J. D.; Gutmann, D. H., Deconvoluting mTOR biology. *Cell Cycle* **2012,** *11* (2), 236-48.

43. Yang, H.; Jiang, X.; Li, B.; Yang, H. J.; Miller, M.; Yang, A.; Dhar, A.; Pavletich, N. P., Mechanisms of mTORC1 activation by RHEB and inhibition by PRAS40. *Nature* **2017,** *552* (7685), 368-373.

44. Carroll, B.; Maetzel, D.; Maddocks, O. D.; Otten, G.; Ratcliff, M.; Smith, G. R.; Dunlop, E. A.; Passos, J. F.; Davies, O. R.; Jaenisch, R.; Tee, A. R.; Sarkar, S.; Korolchuk, V. I., Control of TSC2-Rheb signaling axis by arginine regulates mTORC1 activity. *Elife* **2016,** *5*.

45. Huang, J.; Brendan, A complex interplay between Akt, TSC2 and the two mTOR complexes. *Biochemical Society Transactions* **2009,** *37* (1), 217-222.

46. Bruno D. Fonseca, R. M. L., Christian K. Damgaard, Tommy Alain, Andrea J. Berman, LARP1 on TOP of ribosome production. *WIREs RNA* **2018,** *9* (5).

47. Ghosh, J.; Kapur, R., Role of mTORC1-S6K1 signaling pathway in regulation of hematopoietic stem cell and acute myeloid leukemia. *Exp Hematol* **2017,** *50*, 13-21.

48. Qin, X.; Jiang, B.; Zhang, Y., 4E-BP1, a multifactor regulated multifunctional protein. *Cell Cycle* **2016,** *15* (6), 781-6.

49. Berman, A. J.; Thoreen, C. C.; Dedeic, Z.; Chettle, J.; Roux, P. P.; Blagden, S. P., Controversies around the function of LARP1. *RNA Biology* **2021,** *18* (2), 207-217.

50. Yu, K.; Shi, C.; Toral-Barza, L.; Lucas, J.; Shor, B.; Kim, J. E.; Zhang, W.-G.; Mahoney, R.; Gaydos, C.; Tardio, L.; Kim, S. K.; Conant, R.; Curran, K.; Kaplan, J.; Verheijen, J.; Ayral-Kaloustian, S.; Mansour, T. S.; Abraham, R. T.; Zask, A.; Gibbons, J. J., Beyond Rapalog Therapy: Preclinical Pharmacology and Antitumor Activity of WYE-125132, an ATP-Competitive and Specific Inhibitor of mTORC1 and mTORC2. *Cancer Research* **2010,** *70* (2), 621-631.

51. He, J.; Wang, M.-Y.; Qiu, L.-X.; Zhu, M.-L.; Shi, T.-Y.; Zhou, X.-Y.; Sun, M.-H.; Yang, Y.-J.; Wang, J.-C.; Jin, L.; Wang, Y.-N.; Li, J.; Yu, H.-P.; Wei, Q.-Y., Genetic variations of mTORC1 genes and risk of gastric cancer in an eastern chinese population. *Molecular Carcinogenesis* **2013,** *52* (S1), 70-79.

52. Carracedo, A.; Ma, L.; Teruya-Feldstein, J.; Rojo, F.; Salmena, L.; Alimonti, A.; Egia, A.; Sasaki, A. T.; Thomas, G.; Kozma, S. C.; Papa, A.; Nardella, C.; Cantley, L. C.; Baselga, J.; Pandolfi, P. P., Inhibition of mTORC1 leads to MAPK pathway activation through a PI3K-dependent feedback loop in human cancer. *J Clin Invest* **2008,** *118* (9), 3065-74.

53. Wander, S. A.; Hennessy, B. T.; Slingerland, J. M., Next-generation mTOR inhibitors in clinical oncology: how pathway complexity informs therapeutic strategy. *J Clin Invest* **2011,** *121* (4), 1231-41.

54. Zou, Z.; Tao, T.; Li, H.; Zhu, X., mTOR signaling pathway and mTOR inhibitors in cancer: progress and challenges. *Cell & Bioscience* **2020,** *10* (1).

55. Grabiner, B. C.; Nardi, V.; Birsoy, K.; Possemato, R.; Shen, K.; Sinha, S.; Jordan, A.; Beck, A. H.; Sabatini, D. M., A Diverse Array of Cancer-Associated MTOR Mutations Are Hyperactivating and Can Predict Rapamycin Sensitivity. *Cancer Discovery* **2014,** *4* (5), 554-563.

56. Tian, T.; Li, X.; Zhang, J., mTOR Signaling in Cancer and mTOR Inhibitors in Solid Tumor Targeting Therapy. *International Journal of Molecular Sciences* **2019,** *20* (3), 755.

57. Guo, H.; German, P.; Bai, S.; Barnes, S.; Guo, W.; Qi, X.; Lou, H.; Liang, J.; Jonasch, E.; Mills, G. B.; Ding, Z., The PI3K/AKT Pathway and Renal Cell Carcinoma. *Journal of Genetics and Genomics* **2015,** *42* (7), 343-353.

58. Xu, J.; Pham, C. G.; Albanese, S. K.; Dong, Y.; Oyama, T.; Lee, C.-H.; Rodrik-Outmezguine, V.; Yao, Z.; Han, S.; Chen, D.; Parton, D. L.; Chodera, J. D.; Rosen, N.; Cheng, E. H.; Hsieh, J. J., Mechanistically distinct cancer-associated mTOR activation clusters predict sensitivity to rapamycin. *Journal of Clinical Investigation* **2016,** *126* (9), 3526-3540.

59. Hua, H.; Kong, Q.; Zhang, H.; Wang, J.; Luo, T.; Jiang, Y., Targeting mTOR for cancer therapy. *Journal of Hematology & Oncology* **2019,** *12* (1).

60. Hausch, F.; Kozany, C.; Theodoropoulou, M.; Fabian, A. K., FKBPs and the Akt/mTOR pathway. *Cell Cycle* **2013,** *12* (15), 2366-70.

61. Yoshida-Ichikawa, Y.; Tanabe, M.; Tokuda, E.; Shimizu, H.; Horimoto, Y.; Miura, K.; Saito, M., Overcoming the Adverse Effects of Everolimus to Achieve Maximum Efficacy in the Treatment of Inoperable Breast Cancer: A Review of 11 Cases at Our Hospital. *Case Rep Oncol* **2018,** *11* (2), 511-520.

62. Bukowski, R. M., Temsirolimus: a safety and efficacy review. *Expert Opin Drug Saf* **2012,** *11* (5), 861-79.

63. Xie, J.; Wang, X.; Proud, C. G., mTOR inhibitors in cancer therapy. *F1000Research* **2016,** *5*, 2078.

64. Pallet, N.; Legendre, C., Adverse events associated with mTOR inhibitors. *Expert Opin Drug Saf* **2013,** *12* (2), 177-86.

65. Creel, P. A., Management of mTOR Inhibitor Side Effects. **2009,** *13* (0), 19-23.

66. Thoreen, C. C.; Kang, S. A.; Chang, J. W.; Liu, Q.; Zhang, J.; Gao, Y.; Reichling, L. J.; Sim, T.; Sabatini, D. M.; Gray, N. S., An ATP-competitive Mammalian Target of Rapamycin Inhibitor Reveals Rapamycin-resistant Functions of mTORC1. *Journal of Biological Chemistry* **2009,** *284* (12), 8023-8032.

67. Zeng, Z.; Shi, Y. X.; Tsao, T.; Qiu, Y.; Kornblau, S. M.; Baggerly, K. A.; Liu, W.; Jessen, K.; Liu, Y.; Kantarjian, H.; Rommel, C.; Fruman, D. A.; Andreeff, M.; Konopleva, M., Targeting of mTORC1/2 by the mTOR kinase inhibitor PP242 induces apoptosis in AML cells under conditions mimicking the bone marrow microenvironment. *Blood* **2012,** *120* (13), 2679-2689.

68. Pacey, S.; Shah, N.; Davies, B.; Bratt, O.; Warren, A.; Baird, R. D.; Gnanapragasam, V.; Ingle, S.; Stearn, S.; Machin, A.; Qian, W.; Zhao, W.; Harrington, E.; Oelmann, E. M.; Hategan, M.; Kumar, S. S.; Corbacho, J. G., A pharmacodynamic biomarker study of vistusertib (AZD2014), an mTORC1/2 inhibitor, given prior to radical prostatectomy (CANCAP02). *Journal of Clinical Oncology* **2018,** *36* (15_suppl), 5081-5081.

69. Nunnery, S. E.; Mayer, I. A., Management of toxicity to isoform α-specific PI3K inhibitors. *Annals of Oncology* **2019,** *30*, x21-x26.

70. Bousquet-Antonelli, C.; Deragon, J. M., A comprehensive analysis of the La-motif protein superfamily. *RNA* **2009,** *15* (5), 750-64.

71. Stavraka, C.; Blagden, S., The La-Related Proteins, a Family with Connections to Cancer. *Biomolecules* **2015,** *5* (4), 2701-2722.

72. Maraia, R. J.; Mattijssen, S.; Cruz-Gallardo, I.; Conte, M. R., The La and related RNA-binding proteins (LARPs): structures, functions, and evolving perspectives. *Wiley Interdiscip Rev RNA* **2017,** *8* (6).

73. Bayfield, M. A.; Yang, R.; Maraia, R. J., Conserved and divergent features of the structure and function of La and La-related proteins (LARPs). *Biochim Biophys Acta* **2010,** *1799* (5-6), 365-78.

74. Mattijssen, S.; Kozlov, G.; Fonseca, B. D.; Gehring, K.; Maraia, R. J., LARP1 and LARP4: up close with PABP for mRNA 3' poly(A) protection and stabilization. *RNA Biology* **2021,** *18* (2), 259-274.

75. Küspert, M.; Murakawa, Y.; Schäffler, K.; Vanselow, J. T.; Wolf, E.; Juranek, S.; Schlosser, A.; Landthaler, M.; Fischer, U., LARP4B is an AU-rich sequence associated factor that promotes mRNA accumulation and translation. *Rna* **2015,** *21* (7), 1294-305.

76.   Cai, L.; Fritz, D.; Stefanovic, L.; Stefanovic, B., Binding of LARP6 to the conserved 5'
        stem-loop regulates translation of mRNAs encoding type I collagen. *J Mol Biol* **2010,** *395*
        (2), 309-26.

77.   Stefanovic, L.; Longo, L.; Zhang, Y.; Stefanovic, B., Characterization of binding of
        LARP6 to the 5' stem-loop of collagen mRNAs: implications for synthesis of type I
        collagen. *RNA Biol* **2014,** *11* (11), 1386-401.

78.   Markert, A.; Grimm, M.; Martinez, J.; Wiesner, J.; Meyerhans, A.; Meyuhas, O.;
        Sickmann, A.; Fischer, U., The La-related protein LARP7 is a component of the 7SK
        ribonucleoprotein and affects transcription of cellular and viral polymerase II genes.
        EMBO reports **2008,** *9* (6), 569-575.

79.   Lahr, R. M.; Fonseca, B. D.; Ciotti, G. E.; Al-Ashtal, H. A.; Jia, J. J.; Niklaus, M. R.;
        Blagden, S. P.; Alain, T.; Berman, A. J., La-related protein 1 (LARP1) binds the mRNA
        cap, blocking eIF4F assembly on TOP mRNAs. *Elife* **2017,** *6*.

80.   Lahr, R. M.; Mack, S. M.; Heroux, A.; Blagden, S. P.; Bousquet-Antonelli, C.; Deragon,
        J. M.; Berman, A. J., The La-related protein 1-specific domain repurposes HEAT-like
        repeats to directly bind a 5'TOP sequence. *Nucleic Acids Res* **2015,** *43* (16), 8077-88.

81.   Philippe, L.; Vasseur, J. J.; Debart, F.; Thoreen, C. C., La-related protein 1 (LARP1)
        repression of TOP mRNA translation is mediated through its cap-binding domain and
        controlled by an adjacent regulatory region. *Nucleic Acids Res* **2018,** *46* (3), 1457-1469.

82.   Tcherkezian, J.; Cargnello, M.; Romeo, Y.; Huttlin, E. L.; Lavoie, G.; Gygi, S. P.; Roux,
        P. P., Proteomic analysis of cap-dependent translation identifies LARP1 as a key regulator
        of 5'TOP mRNA translation. *Genes Dev* **2014,** *28* (4), 357-71.

83.   Fonseca, B. D.; Zakaria, C.; Jia, J. J.; Graber, T. E.; Svitkin, Y.; Tahmasebi, S.; Healy,
        D.; Hoang, H. D.; Jensen, J. M.; Diao, I. T.; Lussier, A.; Dajadian, C.; Padmanabhan,
        N.; Wang, W.; Matta-Camacho, E.; Hearnden, J.; Smith, E. M.; Tsukumo, Y.; Yanagiya,
        A.; Morita, M.; Petroulakis, E.; Gonzalez, J. L.; Hernandez, G.; Alain, T.; Damgaard,
        C. K., La-related Protein 1 (LARP1) Represses Terminal Oligopyrimidine (TOP) mRNA
        Translation Downstream of mTOR Complex 1 (mTORC1). *J Biol Chem* **2015,** *290* (26),
        15996-6020.

84.   Hong, S.; Freeberg, M. A.; Han, T.; Kamath, A.; Yao, Y.; Fukuda, T.; Suzuki, T.; Kim,
        J. K.; Inoki, K., LARP1 functions as a molecular switch for mTORC1-mediated translation
        of an essential class of mRNAs. *eLife* **2017,** *6*.

85.   Hopkins, T. G.; Mura, M.; Abd Latip, N.; Weir, J.; Sweeney, K.; Berman, A. J.; Ghaem-
        Maghami, S.; Aboagye, E. O.; Blagden, S. P., The RNA-binding protein LARP1 regulates
        survival and tumorigenesis in ovarian cancer. Imperial College London: London, UK,
        2015.

86.   Hopkins, T. G.; Mura, M.; Al-Ashtal, H. A.; Lahr, R. M.; Abd-Latip, N.; Sweeney, K.;
        Lu, H.; Weir, J.; El-Bahrawy, M.; Steel, J. H.; Ghaem-Maghami, S.; Aboagye, E. O.;

Berman, A. J.; Blagden, S. P., The RNA-binding protein LARP1 is a post-transcriptional regulator of survival and tumorigenesis in ovarian cancer. *Nucleic Acids Res* **2016,** *44* (3), 1227-46.

87.   Mura, M.; Hopkins, T. G.; Michael, T.; Abd-Latip, N.; Weir, J.; Aboagye, E.; Mauri, F.; Jameson, C.; Sturge, J.; Gabra, H.; Bushell, M.; Willis, A. E.; Curry, E.; Blagden, S. P., LARP1 post-transcriptionally regulates mTOR and contributes to cancer progression. *Oncogene* **2015,** *34* (39), 5025-5036.

88.   Xie, C.; Huang, L.; Xie, S.; Xie, D.; Zhang, G.; Wang, P.; Peng, L.; Gao, Z., LARP1 predict the prognosis for early-stage and AFP-normal hepatocellular carcinoma. *J Transl Med* **2013,** *11*, 272.

89.   Ye, L.; Lin, S.-T.; Mi, Y.-S.; Liu, Y.; Ma, Y.; Sun, H.-M.; Peng, Z.-H.; Fan, J.-W., Overexpression of LARP1 predicts poor prognosis of colorectal cancer and is expected to be a potential therapeutic target. *Tumor Biology* **2016,** *37* (11), 14585-14594.

90.   Al-Ashtal, H. A.; Rubottom, C. M.; Leeper, T. C.; Berman, A. J., The LARP1 La-Module recognizes both ends of TOP mRNAs. *RNA Biology* **2019**, 1-11.

91.   Philippe, L.; Van Den Elzen, A. M. G.; Watson, M. J.; Thoreen, C. C., Global analysis of LARP1 translation targets reveals tunable and dynamic features of 5′ TOP motifs. *Proceedings of the National Academy of Sciences* **2020,** *117* (10), 5319-5328.

92.   Burrows, C.; Abd Latip, N.; Lam, S. J.; Carpenter, L.; Sawicka, K.; Tzolovsky, G.; Gabra, H.; Bushell, M.; Glover, D. M.; Willis, A. E.; Blagden, S. P., The RNA binding protein Larp1 regulates cell division, apoptosis and cell migration. *Nucleic Acids Res* **2010,** *38* (16), 5542-53.

93.   Selcuklu, S. D.; Donoghue, M. T.; Rehmet, K.; de Souza Gomes, M.; Fort, A.; Kovvuru, P.; Muniyappa, M. K.; Kerin, M. J.; Enright, A. J.; Spillane, C., MicroRNA-9 inhibition of cell proliferation and identification of novel miR-9 targets by transcriptome profiling in breast cancer cells. *J Biol Chem* **2012,** *287* (35), 29516-28.

94.   Kato, M.; Goto, Y.; Matsushita, R.; Kurozumi, A.; Fukumoto, I.; Nishikawa, R.; Sakamoto, S.; Enokida, H.; Nakagawa, M.; Ichikawa, T.; Seki, N., MicroRNA-26a/b directly regulate La-related protein 1 and inhibit cancer cell invasion in prostate cancer. *Int J Oncol* **2015,** *47* (2), 710-8.

95.   Xu, Z.; Xu, J.; Lu, H.; Lin, B.; Cai, S.; Guo, J.; Zang, F.; Chen, R., LARP1 is regulated by the XIST/miR-374a axis and functions as an oncogene in non-small cell lung carcinoma. *Oncology Reports* **2017**.

96.   Gordon, D. E.; Jang, G. M.; Bouhaddou, M.; Xu, J.; Obernier, K.; White, K. M.; O'Meara, M. J.; Rezelj, V. V.; Guo, J. Z.; Swaney, D. L.; Tummino, T. A.; Hüttenhain, R.; Kaake, R. M.; Richards, A. L.; Tutuncuoglu, B.; Foussard, H.; Batra, J.; Haas, K.; Modak, M.; Kim, M.; Haas, P.; Polacco, B. J.; Braberg, H.; Fabius, J. M.; Eckhardt, M.; Soucheray, M.; Bennett, M. J.; Cakir, M.; McGregor, M. J.; Li, Q.; Meyer, B.;

Roesch, F.; Vallet, T.; Mac Kain, A.; Miorin, L.; Moreno, E.; Naing, Z. Z. C.; Zhou, Y.; Peng, S.; Shi, Y.; Zhang, Z.; Shen, W.; Kirby, I. T.; Melnyk, J. E.; Chorba, J. S.; Lou, K.; Dai, S. A.; Barrio-Hernandez, I.; Memon, D.; Hernandez-Armenta, C.; Lyu, J.; Mathy, C. J. P.; Perica, T.; Pilla, K. B.; Ganesan, S. J.; Saltzberg, D. J.; Rakesh, R.; Liu, X.; Rosenthal, S. B.; Calviello, L.; Venkataramanan, S.; Liboy-Lugo, J.; Lin, Y.; Huang, X.-P.; Liu, Y.; Wankowicz, S. A.; Bohn, M.; Safari, M.; Ugur, F. S.; Koh, C.; Savar, N. S.; Tran, Q. D.; Shengjuler, D.; Fletcher, S. J.; O'Neal, M. C.; Cai, Y.; Chang, J. C. J.; Broadhurst, D. J.; Klippsten, S.; Sharp, P. P.; Wenzell, N. A.; Kuzuoglu-Ozturk, D.; Wang, H.-Y.; Trenker, R.; Young, J. M.; Cavero, D. A.; Hiatt, J.; Roth, T. L.; Rathore, U.; Subramanian, A.; Noack, J.; Hubert, M.; Stroud, R. M.; Frankel, A. D.; Rosenberg, O. S.; Verba, K. A.; Agard, D. A.; Ott, M.; Emerman, M.; Jura, N.; Von Zastrow, M.; Verdin, E.; Ashworth, A.; Schwartz, O.; D'Enfert, C.; Mukherjee, S.; Jacobson, M.; Malik, H. S.; Fujimori, D. G.; Ideker, T.; Craik, C. S.; Floor, S. N.; Fraser, J. S.; Gross, J. D.; Sali, A.; Roth, B. L.; Ruggero, D.; Taunton, J.; Kortemme, T.; Beltrao, P.; Vignuzzi, M.; García-Sastre, A.; Shokat, K. M.; Shoichet, B. K.; Krogan, N. J., A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. *Nature* **2020,** *583* (7816), 459-468.

97.   Schmidt, N.; Lareau, C. A.; Keshishian, H.; Ganskih, S.; Schneider, C.; Hennig, T.; Melanson, R.; Werner, S.; Wei, Y.; Zimmer, M.; Ade, J.; Kirschner, L.; Zielinski, S.; Dölken, L.; Lander, E. S.; Caliskan, N.; Fischer, U.; Vogel, J.; Carr, S. A.; Bodem, J.; Munschauer, M., The SARS-CoV-2 RNA–protein interactome in infected human cells. *Nature Microbiology* **2021,** *6* (3), 339-353.

98.   Cassidy, K. C.; Lahr, R. M.; Kaminsky, J. C.; Mack, S.; Fonseca, B. D.; Das, S. R.; Berman, A. J.; Durrant, J. D., Capturing the Mechanism Underlying TOP mRNA Binding to LARP1. *Structure* **2019,** *27* (12), 1771-1781.e5.

99.   Gingras, A. C., Regulation of translation initiation by FRAP/mTOR. *Genes & Development* **2001,** *15* (7), 807-826.

100.   Brunn, G. J.; Hudson, C. C.; Sekulić, A.; Williams, J. M.; Hosoi, H.; Houghton, P. J.; Lawrence, J. C.; Abraham, R. T., Phosphorylation of the Translational Repressor PHAS-I by the Mammalian Target of Rapamycin. *Science* **1997,** *277* (5322), 99-101.

101.   Burnett, P. E.; Barrow, R. K.; Cohen, N. A.; Snyder, S. H.; Sabatini, D. M., RAFT1 phosphorylation of the translational regulators p70 S6 kinase and 4E-BP1. *Proceedings of the National Academy of Sciences* **1998,** *95* (4), 1432-1437.

102.   Filipowicz, W.; Furuichi, Y.; Sierra, J. M.; Muthukrishnan, S.; Shatkin, A. J.; Ochoa, S., A protein binding the methylated 5'-terminal sequence, m7GpppN, of eukaryotic messenger RNA. *Proceedings of the National Academy of Sciences* **1976,** *73* (5), 1559-1563.

103.   Sonenberg, N.; Morgan, M. A.; Merrick, W. C.; Shatkin, A. J., A polypeptide in eukaryotic initiation factors that crosslinks specifically to the 5'-terminal cap in mRNA. *Proceedings of the National Academy of Sciences* **1978,** *75* (10), 4843-4847.

104. Sonenberg, N.; Hinnebusch, A. G., Regulation of Translation Initiation in Eukaryotes: Mechanisms and Biological Targets. *Cell* **2009,** *136* (4), 731-745.

105. Aoki, K.; Adachi, S.; Homoto, M.; Kusano, H.; Koike, K.; Natsume, T., LARP1 specifically recognizes the 3' terminus of poly(A) mRNA. *FEBS Lett* **2013,** *587* (14), 2173-8.

106. Wilbertz, J. H.; Voigt, F.; Horvathova, I.; Roth, G.; Zhan, Y.; Chao, J. A., Single-Molecule Imaging of mRNA Localization and Regulation during the Integrated Stress Response. *Molecular Cell* **2019,** *73* (5), 946-958.e7.

107. Merret, R.; Descombin, J.; Juan, Y. T.; Favory, J. J.; Carpentier, M. C.; Chaparro, C.; Charng, Y. Y.; Deragon, J. M.; Bousquet-Antonelli, C., XRN4 and LARP1 are required for a heat-triggered mRNA decay pathway involved in plant acclimation and survival during thermal stress. *Cell Rep* **2013,** *5* (5), 1279-93.

108. Gentilella, A.; Moron-Duran, F. D.; Fuentes, P.; Zweig-Rocha, G.; Riano-Canalias, F.; Pelletier, J.; Ruiz, M.; Turon, G.; Castano, J.; Tauler, A.; Bueno, C.; Menendez, P.; Kozma, S. C.; Thomas, G., Autogenous Control of 5'TOP mRNA Stability by 40S Ribosomes. *Mol Cell* **2017,** *67* (1), 55-70 e4.

109. Damgaard, C. K.; Lykke-Andersen, J., Translational coregulation of 5'TOP mRNAs by TIA-1 and TIAR. *Genes Dev* **2011,** *25* (19), 2057-68.

110. Durrant, J. D.; Votapka, L.; Sørensen, J.; Amaro, R. E., POVME 2.0: An Enhanced Tool for Determining Pocket Shape and Volume Characteristics. **2014,** *10* (11), 5047-5056.

111. Frey, B. J.; Dueck, D., Clustering by Passing Messages Between Data Points. *Science* **2007,** *315* (5814), 972-976.

112. Brenke, R.; Kozakov, D.; Chuang, G. Y.; Beglov, D.; Hall, D.; Landon, M. R.; Mattos, C.; Vajda, S., Fragment-based identification of druggable 'hot spots' of proteins using Fourier domain correlation techniques. *Bioinformatics* **2009,** *25* (5), 621-7.

113. Forbes, S. A.; Beare, D.; Boutselakis, H.; Bamford, S.; Bindal, N.; Tate, J.; Cole, C. G.; Ward, S.; Dawson, E.; Ponting, L.; Stefancsik, R.; Harsha, B.; Kok, C. Y.; Jia, M.; Jubb, H.; Sondka, Z.; Thompson, S.; De, T.; Campbell, P. J., COSMIC: somatic cancer genetics at high-resolution. *Nucleic Acids Research* **2017,** *45* (D1), D777-D783.

114. Hornbeck, P. V.; Zhang, B.; Murray, B.; Kornhauser, J. M.; Latham, V.; Skrzypek, E., PhosphoSitePlus, 2014: mutations, PTMs and recalibrations. *Nucleic Acids Res* **2015,** *43* (Database issue), D512-20.

115. Udeshi, N. D.; Svinkina, T.; Mertins, P.; Kuhn, E.; Mani, D. R.; Qiao, J. W.; Carr, S. A., Refined Preparation and Use of Anti-diglycine Remnant (K- -GG) Antibody Enables Routine Quantification of 10,000s of Ubiquitination Sites in Single Proteomics Experiments. **2013,** *12* (3), 825-831.

116. Mukhopadhyay, D.; Riezman, H., Proteasome-independent functions of ubiquitin in endocytosis and signaling. *Science* **2007,** *315* (5809), 201-5.

117. Schnell, J. D.; Hicke, L., Non-traditional Functions of Ubiquitin and Ubiquitin-binding Proteins. **2003,** *278* (38), 35857-35860.

118. Zhao, J.; Goldberg, A. L., Coordinate regulation of autophagy and the ubiquitin proteasome system by MTOR. *Autophagy* **2016,** *12* (10), 1967-1970.

119. Zhao, J.; Zhai, B.; Gygi, S. P.; Goldberg, A. L., mTOR inhibition activates overall protein degradation by the ubiquitin proteasome system as well as by autophagy. *Proceedings of the National Academy of Sciences* **2015,** *112* (52), 15790-15797.

120. Deng, L.; Jiang, C.; Chen, L.; Jin, J.; Wei, J.; Zhao, L.; Chen, M.; Pan, W.; Xu, Y.; Chu, H.; Wang, X.; Ge, X.; Li, D.; Liao, L.; Liu, M.; Li, L.; Wang, P., The ubiquitination of rag A GTPase by RNF152 negatively regulates mTORC1 activation. *Mol Cell* **2015,** *58* (5), 804-18.

121. Kozakov, D.; Grove, L. E.; Hall, D. R.; Bohnuud, T.; Mottarella, S. E.; Luo, L.; Xia, B.; Beglov, D.; Vajda, S., The FTMap family of web servers for determining and characterizing ligand-binding hot spots of proteins. *Nat Protoc* **2015,** *10* (5), 733-55.

122. Ngan, C. H.; Bohnuud, T.; Mottarella, S. E.; Beglov, D.; Villar, E. A.; Hall, D. R.; Kozakov, D.; Vajda, S., FTMAP: extended protein mapping with user-selected probe molecules. *Nucleic Acids Res* **2012,** *40* (Web Server issue), W271-5.

123. Dolinsky, T. J.; Czodrowski, P.; Li, H.; Nielsen, J. E.; Jensen, J. H.; Klebe, G.; Baker, N. A., PDB2PQR: expanding and upgrading automated preparation of biomolecular structures for molecular simulations. **2007,** *35* (Web Server), W522-W525.

124. Dolinsky, T. J.; Nielsen, J. E.; McCammon, J. A.; Baker, N. A., PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. *Nucleic Acids Res* **2004,** *32* (Web Server issue), W665-7.

125. Olsson, M. H.; Søndergaard, C. R.; Rostkowski, M.; Jensen, J. H., PROPKA3: Consistent Treatment of Internal and Surface Residues in Empirical pKa Predictions. *J Chem Theory Comput* **2011,** *7* (2), 525-37.

126. D.A. Case, D. S. C., T.E. Cheatham, III, T.A. Darden, R.E. Duke, T.J. Giese, H. Gohlke, A.W. Goetz, D. Greene, N. Homeyer, S. Izadi, A. Kovalenko, T.S. Lee, S. LeGrand, P. Li, C. Lin, . Liu, T. Luchko, R. Luo, D. Mermelstein, K.M. Merz, G. Monard, H. Nguyen, I. Omelyan, A. Onufriev, F. Pan, R. Qi, D.R. Roe, A. Roitberg, C. Sagui, C.L. Simmerling, W.M. Botello-Smith, J. Swails, R.C. Walker, J. Wang, R.M. Wolf, X. Wu, L. Xiao, D.M. York and P.A. Kollman, Amber 2017. *University of California, San Francisco* **2017**.

127. Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C., ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *Journal of Chemical Theory and Computation* **2015,** *11* (8), 3696-3713.

128. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L., Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics* **1983,** *79* (2), 926-935.

129. Kalé, L.; Skeel, R.; Bhandarkar, M.; Brunner, R.; Gursoy, A.; Krawetz, N.; Phillips, J.; Shinozaki, A.; Varadarajan, K.; Schulten, K., NAMD2: Greater Scalability for Parallel Molecular Dynamics. *Journal of Computational Physics* **1999,** *151* (1), 283-312.

130. Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K., Scalable molecular dynamics with NAMD. *J Comput Chem* **2005,** *26* (16), 1781-802.

131. Durrant, J. D., BlendMol: advanced macromolecular visualization in Blender. *Bioinformatics* **2019,** *35* (13), 2323-2325.

132. Michaud-Agrawal, N.; Denning, E. J.; Woolf, T. B.; Beckstein, O., MDAnalysis: a toolkit for the analysis of molecular dynamics simulations. *J Comput Chem* **2011,** *32* (10), 2319-27.

133. Humphrey, W.; Dalke, A.; Schulten, K., VMD: Visual molecular dynamics. *Journal of Molecular Graphics* **1996,** *14* (1), 33-38.

134. Studier, F. W., Protein production by auto-induction in high-density shaking cultures. *Protein Expression and Purification* **2005,** *41* (1), 207-234.

135. Adams, P. D.; Afonine, P. V.; Bunkoczi, G.; Chen, V. B.; Davis, I. W.; Echols, N.; Headd, J. J.; Hung, L. W.; Kapral, G. J.; Grosse-Kunstleve, R. W.; McCoy, A. J.; Moriarty, N. W.; Oeffner, R.; Read, R. J.; Richardson, D. C.; Richardson, J. S.; Terwilliger, T. C.; Zwart, P. H., PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **2010,** *66* (Pt 2), 213-21.

136. Winn, M. D.; Ballard, C. C.; Cowtan, K. D.; Dodson, E. J.; Emsley, P.; Evans, P. R.; Keegan, R. M.; Krissinel, E. B.; Leslie, A. G.; McCoy, A.; McNicholas, S. J.; Murshudov, G. N.; Pannu, N. S.; Potterton, E. A.; Powell, H. R.; Read, R. J.; Vagin, A.; Wilson, K. S., Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* **2011,** *67* (Pt 4), 235-42.

137. Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K., Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **2010,** *66* (Pt 4), 486-501.

138. Zlatev, I.; Manoharan, M.; Vasseur, J. J.; Morvan, F., Solid-phase chemical synthesis of 5'-triphosphate DNA, RNA, and chemically modified oligonucleotides. *Curr Protoc Nucleic Acid Chem* **2012,** *Chapter 1*, Unit1.28.

139. Chan, S., Targeting the mammalian target of rapamycin (mTOR): a new approach to treating cancer. *British Journal of Cancer* **2004,** *91* (8), 1420-1424.

140. Liu, Q.; Kirubakaran, S.; Hur, W.; Niepel, M.; Westover, K.; Thoreen, C. C.; Wang, J.; Ni, J.; Patricelli, M. P.; Vogel, K.; Riddle, S.; Waller, D. L.; Traynor, R.; Sanda, T.; Zhao, Z.; Kang, S. A.; Zhao, J.; Look, A. T.; Sorger, P. K.; Sabatini, D. M.; Gray, N. S., Kinome-wide Selectivity Profiling of ATP-competitive Mammalian Target of Rapamycin (mTOR) Inhibitors and Characterization of Their Binding Kinetics. **2012,** *287* (13), 9742-9752.

141. Brown, D. K.; Penkler, D. L.; Sheik Amamuddy, O.; Ross, C.; Atilgan, A. R.; Atilgan, C.; Tastan Bishop, Ö., MD-TASK: a software suite for analyzing molecular dynamics trajectories. *Bioinformatics* **2017,** *33* (17), 2768-2771.

142. Van Wart, A. T.; Durrant, J.; Votapka, L.; Amaro, R. E., Weighted Implementation of Suboptimal Paths (WISP): An Optimized Algorithm and Tool for Dynamical Network Analysis. **2014,** *10* (2), 511-517.

143. Baker, N. A.; Sept, D.; Joseph, S.; Holst, M. J.; McCammon, J. A., Electrostatics of nanosystems: Application to microtubules and the ribosome. *Proceedings of the National Academy of Sciences* **2001,** *98* (18), 10037-10041.

144. Ashkenazy, H.; Abadi, S.; Martz, E.; Chay, O.; Mayrose, I.; Pupko, T.; Ben-Tal, N., ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic Acids Research* **2016,** *44* (W1), W344-W350.

145. Landau, M.; Mayrose, I.; Rosenberg, Y.; Glaser, F.; Martz, E.; Pupko, T.; Ben-Tal, N., ConSurf 2005: the projection of evolutionary conservation scores of residues on protein structures. **2005,** *33* (Web Server), W299-W302.

146. Bakan, A.; Dutta, A.; Mao, W.; Liu, Y.; Chennubhotla, C.; Lezon, T. R.; Bahar, I., Evol and ProDy for bridging protein sequence evolution and structural dynamics. *Bioinformatics* **2014,** *30* (18), 2681-2683.

147. Zerbino, D. R.; Achuthan, P.; Akanni, W.; Barrell, D.; Bhai, J.; Billis, K.; Cummins, C.; Gall, A.; Girón, C. G.; Gil, L.; Gordon, L.; Haggerty, L.; Haskell, E.; Hourlier, T.; Izuogu, O. G.; Janacek, S. H.; Juettemann, T.; To, J. K.; Laird, M. R.; Lavidas, I.; Liu, Z.; Loveland, J. E.; Maurel, T.; McLaren, W.; Moore, B.; Mudge, J.; Murphy, D. N.; Newman, V.; Nuhn, M.; Ogeh, D.; Ong, C. K.; Parker, A.; Patricio, M.; Riat, H. S.; Schuilenburg, H.; Sheppard, D.; Sparrow, H.; Taylor, K.; Thormann, A.; Vullo, A.; Walts, B.; Zadissa, A.; Frankish, A.; Hunt, S. E.; Kostadima, M.; Langridge, N.; Martin, F. J.; Muffato, M.; Perry, E.; Ruffier, M.; Staines, D. M.; Trevanion, S. J.; Aken, B. L.; Cunningham, F.; Yates, A.; Flicek, P., Ensembl 2018. *Nucleic Acids Research* **2018,** *46* (D1), D754-D761.

148. Thoreen, C. C.; Chantranupong, L.; Keys, H. R.; Wang, T.; Gray, N. S.; Sabatini, D. M., A unifying model for mTORC1-mediated regulation of mRNA translation. *Nature* **2012,** *485* (7396), 109-13.

149. O'Reilly, K. E.; Rojo, F.; She, Q. B.; Solit, D.; Mills, G. B.; Smith, D.; Lane, H.; Hofmann, F.; Hicklin, D. J.; Ludwig, D. L.; Baselga, J.; Rosen, N., mTOR inhibition

induces upstream receptor tyrosine kinase signaling and activates Akt. *Cancer Res* **2006,** *66* (3), 1500-8.

150. Yao, X. Q.; Momin, M.; Hamelberg, D., Elucidating Allosteric Communications in Proteins with Difference Contact Network Analysis. *J Chem Inf Model* **2018,** *58* (7), 1325-1330.

151. D.A. Case, H. M. A., K. Belfon, I.Y. Ben-Shalom, S.R. Brozell, D.S. Cerutti, T.E. Cheatham, III, V.W.D. Cruzeiro, T.A. Darden, R.E. Duke, G. Giambasu, M.K. Gilson, H. Gohlke, A.W. Goetz, R. Harris, S. Izadi, S.A. Izmailov, C. Jin, K. Kasavajhala, M.C. Kaymak, E. King, A. Kovalenko, T. Kurtzman, T.S. Lee, S. LeGrand, P. Li, C. Lin, J. Liu, T. Luchko, R. Luo, M. Machado, V. Man, M. Manathunga, K.M. Merz, Y. Miao, O. Mikhailovskii, G. Monard, H. Nguyen, K.A. O'Hearn, A. Onufriev, F. Pan, S. Pantano, R. Qi, A. Rahnamoun, D.R. Roe, A. Roitberg, C. Sagui, S. Schott-Verdugo, J. Shen, C.L. Simmerling, N.R. Skrynnikov, J. Smith, J. Swails, R.C. Walker, J. Wang, H. Wei, R.M. Wolf, X. Wu, Y. Xue, D.M. York, S. Zhao, and P.A. Kollman, Amber 2021. *University of California, San Francisco* **2021**.

152. Roe, D. R.; Cheatham, T. E., PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *Journal of Chemical Theory and Computation* **2013,** *9* (7), 3084-3095.

153. Murtagh, F.; Contreras, P., Algorithms for hierarchical clustering: an overview. *WIREs Data Mining and Knowledge Discovery* **2012,** *2* (1), 86-97.

154. Rodriguez-Bussey, I.; Yao, X. Q.; Shouaib, A. D.; Lopez, J.; Hamelberg, D., Decoding Allosteric Communication Pathways in Cyclophilin A with a Comparative Analysis of Perturbed Conformational Ensembles. *J Phys Chem B* **2018,** *122* (25), 6528-6535.

155. Souffrant, M. G.; Yao, X.-Q.; Momin, M.; Hamelberg, D., N-Glycosylation and Gaucher Disease Mutation Allosterically Alter Active-Site Dynamics of Acid-β-Glucosidase. *ACS Catalysis* **2020,** *10* (3), 1810-1820.

156. Yao, X.-Q.; Hamelberg, D., Detecting Functional Dynamics in Proteins with Comparative Perturbed-Ensembles Analysis. *Accounts of Chemical Research* **2019,** *52* (12), 3455-3464.

157. Kim, J.; Guan, K.-L., mTOR as a central hub of nutrient signalling and cell growth. *Nature Cell Biology* **2019,** *21* (1), 63-71.

158. Wengrod, J. C.; Gardner, L. B., Cellular adaptation to nutrient deprivation: crosstalk between the mTORC1 and eIF2α signaling pathways and implications for autophagy. **2015,** *14* (16), 2571-2577.

159. Mehta, R.; Chandler-Brown, D.; Ramos, F. J.; Shamieh, L. S.; Kaeberlein, M., Regulation of mRNA translation as a conserved mechanism of longevity control. *Adv Exp Med Biol* **2010,** *694*, 14-29.

160. Sengupta, S.; Peterson, T. R.; Sabatini, D. M., Regulation of the mTOR Complex 1 Pathway by Nutrients, Growth Factors, and Stress. *Molecular Cell* **2010,** *40* (2), 310-322.

161. Tesseraud, S.; Abbas, M.; Duchene, S.; Bigot, K.; Vaudin, P.; Dupont, J., Mechanisms involved in the nutritional regulation of mRNA translation: features of the avian model. *Nutr Res Rev* **2006,** *19* (1), 104-16.

162. Pelletier, J.; Thomas, G.; Volarević, S., Ribosome biogenesis in cancer: new players and therapeutic avenues. *Nature Reviews Cancer* **2017,** *18* (1), 51-63.

163. Cockman, E.; Anderson, P.; Ivanov, P., TOP mRNPs: Molecular Mechanisms and Principles of Regulation. *Biomolecules* **2020,** *10* (7), 969.

164. Zongaro, S.; Hukema, R.; D'Antoni, S.; Davidovic, L.; Barbry, P.; Catania, M. V.; Willemsen, R.; Mari, B.; Bardoni, B., The 3' UTR of FMR1 mRNA is a target of miR-101, miR-129-5p and miR-221: implications for the molecular pathology of FXTAS at the synapse. *Hum Mol Genet* **2013,** *22* (10), 1971-82.

165. Conciatori, F.; Ciuffreda, L.; Bazzichetto, C.; Falcone, I.; Pilotto, S.; Bria, E.; Cognetti, F.; Milella, M., mTOR Cross-Talk in Cancer and Potential for Combination Therapy. *Cancers* **2018,** *10* (1), 23.

166. Bakan, A.; Nevins, N.; Lakdawala, A. S.; Bahar, I., Druggability Assessment of Allosteric Proteins by Dynamics Simulations in the Presence of Probe Molecules. *J Chem Theory Comput* **2012,** *8* (7), 2435-2447.

167. Sunseri, J.; Koes, D. R., Pharmit: interactive exploration of chemical space. **2016,** *44* (W1), W442-W448.

168. Irwin, J. J.; Sterling, T.; Mysinger, M. M.; Bolstad, E. S.; Coleman, R. G., ZINC: A Free Tool to Discover Chemistry for Biology. *Journal of Chemical Information and Modeling* **2012,** *52* (7), 1757-1768.

169. O'Boyle, N. M.; Banck, M.; James, C. A.; Morley, C.; Vandermeersch, T.; Hutchison, G. R., Open Babel: An open chemical toolbox. *Journal of Cheminformatics* **2011,** *3* (1), 33.

170. Ropp, P. J.; Spiegel, J. O.; Walker, J. L.; Green, H.; Morales, G. A.; Milliken, K. A.; Ringe, J. J.; Durrant, J. D., Gypsum-DL: an open-source program for preparing small-molecule libraries for structure-based virtual screening. *Journal of Cheminformatics* **2019,** *11* (1).

171. Morris, G. M.; Huey, R.; Lindstrom, W.; Sanner, M. F.; Belew, R. K.; Goodsell, D. S.; Olson, A. J., AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J Comput Chem* **2009,** *30* (16), 2785-91.

172. Sanner, M. F., Python: a programming language for software integration and development. *J Mol Graph Model* **1999,** *17* (1), 57-61.

173. Trott, O.; Olson, A. J., AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem* **2010,** *31* (2), 455-61.

174. Ropp, P.; Friedman, A.; Durrant, J. D., Scoria: a Python module for manipulating 3D molecular data. *Journal of Cheminformatics* **2017,** *9* (1).

175. Carey, M. F.; Peterson, C. L.; Smale, S. T., PCR-Mediated Site-Directed Mutagenesis. *Cold Spring Harbor Protocols* **2013,** *2013* (8), pdb.prot076505.

176. Hellman, L. M.; Fried, M. G., Electrophoretic mobility shift assay (EMSA) for detecting protein–nucleic acid interactions. *Nature Protocols* **2007,** *2* (8), 1849-1861.

177. O'Reilly, F. J.; Rappsilber, J., Cross-linking mass spectrometry: methods and applications in structural, molecular and systems biology. *Nature Structural & Molecular Biology* **2018,** *25* (11), 1000-1008.

178. Lyskov, S.; Gray, J. J., The RosettaDock server for local protein-protein docking. *Nucleic Acids Res* **2008,** *36* (Web Server issue), W233-8.

179. Taliun, D.; Harris, D. N.; Kessler, M. D.; Carlson, J.; Szpiech, Z. A.; Torres, R.; Taliun, S. A. G.; Corvelo, A.; Gogarten, S. M.; Kang, H. M.; Pitsillides, A. N.; Lefaive, J.; Lee, S.-B.; Tian, X.; Browning, B. L.; Das, S.; Emde, A.-K.; Clarke, W. E.; Loesch, D. P.; Shetty, A. C.; Blackwell, T. W.; Smith, A. V.; Wong, Q.; Liu, X.; Conomos, M. P.; Bobo, D. M.; Aguet, F.; Albert, C.; Alonso, A.; Ardlie, K. G.; Arking, D. E.; Aslibekyan, S.; Auer, P. L.; Barnard, J.; Barr, R. G.; Barwick, L.; Becker, L. C.; Beer, R. L.; Benjamin, E. J.; Bielak, L. F.; Blangero, J.; Boehnke, M.; Bowden, D. W.; Brody, J. A.; Burchard, E. G.; Cade, B. E.; Casella, J. F.; Chalazan, B.; Chasman, D. I.; Chen, Y.-D. I.; Cho, M. H.; Choi, S. H.; Chung, M. K.; Clish, C. B.; Correa, A.; Curran, J. E.; Custer, B.; Darbar, D.; Daya, M.; De Andrade, M.; Demeo, D. L.; Dutcher, S. K.; Ellinor, P. T.; Emery, L. S.; Eng, C.; Fatkin, D.; Fingerlin, T.; Forer, L.; Fornage, M.; Franceschini, N.; Fuchsberger, C.; Fullerton, S. M.; Germer, S.; Gladwin, M. T.; Gottlieb, D. J.; Guo, X.; Hall, M. E.; He, J.; Heard-Costa, N. L.; Heckbert, S. R.; Irvin, M. R.; Johnsen, J. M.; Johnson, A. D.; Kaplan, R.; Kardia, S. L. R.; Kelly, T.; Kelly, S.; Kenny, E. E.; Kiel, D. P.; Klemmer, R.; Konkle, B. A.; Kooperberg, C.; Köttgen, A.; Lange, L. A.; Lasky-Su, J.; Levy, D.; Lin, X.; Lin, K.-H.; Liu, C.; Loos, R. J. F.; Garman, L.; Gerszten, R.; Lubitz, S. A.; Lunetta, K. L.; Mak, A. C. Y.; Manichaikul, A.; Manning, A. K.; Mathias, R. A.; McManus, D. D.; McGarvey, S. T.; Meigs, J. B.; Meyers, D. A.; Mikulla, J. L.; Minear, M. A.; Mitchell, B. D.; Mohanty, S.; Montasser, M. E.; Montgomery, C.; Morrison, A. C.; Murabito, J. M.; Natale, A.; Natarajan, P.; Nelson, S. C.; North, K. E.; O'Connell, J. R.; Palmer, N. D.; Pankratz, N.; Peloso, G. M.; Peyser, P. A.; Pleiness, J.; Post, W. S.; Psaty, B. M.; Rao, D. C.; Redline, S.; Reiner, A. P.; Roden, D.; Rotter, J. I.; Ruczinski, I.; Sarnowski, C.; Schoenherr, S.; Schwartz, D. A.; Seo, J.-S.; Seshadri, S.; Sheehan, V. A.; Sheu, W. H.; Shoemaker, M. B.; Smith, N. L.; Smith, J. A.; Sotoodehnia, N.; Stilp, A. M.; Tang, W.; Taylor, K. D.; Telen, M.; Thornton, T. A.; Tracy, R. P.; Van Den Berg, D. J.; Vasan, R. S.; Viaud-Martinez, K. A.; Vrieze, S.; Weeks, D. E.; Weir, B. S.; Weiss, S. T.; Weng, L.-C.;

Willer, C. J.; Zhang, Y.; Zhao, X.; Arnett, D. K.; Ashley-Koch, A. E.; Barnes, K. C.; Boerwinkle, E.; Gabriel, S.; Gibbs, R.; Rice, K. M.; Rich, S. S.; Silverman, E. K.; Qasba, P.; Gan, W.; Papanicolaou, G. J.; Nickerson, D. A.; Browning, S. R.; Zody, M. C.; Zöllner, S.; Wilson, J. G.; Cupples, L. A.; Laurie, C. C.; Jaquish, C. E.; Hernandez, R. D.; O'Connor, T. D.; Abecasis, G. R., Sequencing of 53,831 diverse genomes from the NHLBI TOPMed Program. *Nature* **2021,** *590* (7845), 290-299.

180.   Comprehensive genomic characterization defines human glioblastoma genes and core pathways. *Nature* **2008,** *455* (7216), 1061-1068.

181.   Vander Meulen, K. A.; Horowitz, S.; Trievel, R. C.; Butcher, S. E., Measuring the Kinetics of Molecular Association by Isothermal Titration Calorimetry. Elsevier: 2016; pp 181-213.

182.   Jerabek-Willemsen, M.; André, T.; Wanner, R.; Roth, H. M.; Duhr, S.; Baaske, P.; Breitsprecher, D., MicroScale Thermophoresis: Interaction analysis and beyond. *Journal of Molecular Structure* **2014,** *1077*, 101-113.

183.   Xia, M.; Huang, R.; Witt, K. L.; Southall, N.; Fostel, J.; Cho, M. H.; Jadhav, A.; Smith, C. S.; Inglese, J.; Portier, C. J.; Tice, R. R.; Austin, C. P., Compound cytotoxicity profiling using quantitative high-throughput screening. *Environ Health Perspect* **2008,** *116* (3), 284-91.

184.   Xiao, J.; Huang, Z.; Chen, C. Z.; Agoulnik, I. U.; Southall, N.; Hu, X.; Jones, R. E.; Ferrer, M.; Zheng, W.; Agoulnik, A. I.; Marugan, J. J., Identification and optimization of small-molecule agonists of the human relaxin hormone receptor RXFP1. *Nat Commun* **2013,** *4*, 1953.

185.   Green, H.; Koes, D. R.; Durrant, J. D., DeepFrag: a deep convolutional neural network for fragment-based lead optimization. *Chemical Science* **2021**.

186.   Cassidy, K. C.; Šefčík, J.; Raghav, Y.; Chang, A.; Durrant, J. D., ProteinVR: Web-based molecular visualization in virtual reality. *PLOS Computational Biology* **2020,** *16* (3), e1007747.

187.   DeLano, W. L., PyMOL: An Open-Source Molecular Graphics Tool. *CCP4 Newsletter On Protein Crystallography* **2002,** (40), 82-92.

188.   Nakano, C. M.; Moen, E.; Byun, H. S.; Ma, H.; Newman, B.; McDowell, A.; Wei, T.; El-Naggar, M. Y., iBET: Immersive visualization of biological electron-transfer dynamics. *Journal of Molecular Graphics and Modelling* **2016,** *65*, 94-99.

189.   Rose, A. S.; Hildebrand, P. W., NGL Viewer: a web application for molecular visualization. *Nucleic Acids Research* **2015,** *43* (W1), W576-W579.

190.   Rego, N.; Koes, D., 3Dmol.js: molecular visualization with WebGL. *Bioinformatics* **2015,** *31* (8), 1322-1324.

191.   Norrby, M.; Grebner, C.; Eriksson, J.; Boström, J., Molecular Rift: Virtual Reality for Drug Designers. *Journal of Chemical Information and Modeling* **2015,** *55* (11), 2475-2484.

192.   Ratamero, E. M.; Bellini, D.; Dowson, C. G.; Römer, R. A., Touching proteins with virtual bare hands. *Journal of Computer-Aided Molecular Design* **2018,** *32* (6), 703-709.

193.   Ferrell, J. B.; Campbell, J. P.; McCarthy, D. R.; McKay, K. T.; Hensinger, M.; Srinivasan, R.; Zhao, X.; Wurthmann, A.; Li, J.; Schneebeli, S. T., Chemical Exploration with Virtual Reality in Organic Teaching Laboratories. *Journal of Chemical Education* **2019,** *96* (9), 1961-1966.

194.   Trindade, J.; Fiolhais, C.; Almeida, L., Science learning in virtual environments: a descriptive study. *British Journal of Educational Technology* **2002,** *33* (4), 471-488.

195.   Mikropoulos, T. A.; Natsis, A., Educational virtual environments: A ten-year review of empirical research (1999–2009). *Computers & Education* **2011,** *56* (3), 769-780.

196.   Bennie, S. J.; Ranaghan, K. E.; Deeks, H.; Goldsmith, H. E.; O'Connor, M. B.; Mulholland, A. J.; Glowacki, D. R., Teaching Enzyme Catalysis Using Interactive Molecular Dynamics in Virtual Reality. *Journal of Chemical Education* **2019,** *96* (11), 2488-2496.

197.   Hua, Q.-X.; Gozani, S. N.; Chance, R. E.; Hoffmann, J. A.; Frank, B. H.; Weiss, M. A., Structure of a protein in a kinetic trap. *Nature Structural & Molecular Biology* **1995,** *2* (2), 129-138.

198.   Kent, B. R., 3D Scientific Visualization with Blender®. Morgan & Claypool Publishers: 2015. http://dx.doi.org/10.1088/978-1-6270-5612-0.

199.   Borrel, A.; Fourches, D., RealityConvert: a tool for preparing 3D models of biochemical structures for augmented and virtual reality. *Bioinformatics* **2017,** *33* (23), 3816-3818.

200.   Hoffman, M. A.; Provance, J. B., Visualization of molecular structures using HoloLens-based augmented reality. *AMIA Jt Summits Transl Sci Proc* **2017,** *2017*, 68-74.

201.   Zhang, J. F.; Paciorkowski, A. R.; Craig, P. A.; Cui, F., BioVR: a platform for virtual reality assisted biological data integration and visualization. *BMC Bioinformatics* **2019,** *20* (1).

202.   Müller, C.; Krone, M.; Huber, M.; Biener, V.; Herr, D.; Koch, S.; Reina, G.; Weiskopf, D.; Ertl, T., Interactive Molecular Graphics for Augmented Reality Using HoloLens. *Journal of Integrative Bioinformatics* **2018,** *15* (2).

203.   O'Connor, M.; Deeks, H. M.; Dawn, E.; Metatla, O.; Roudaut, A.; Sutton, M.; Thomas, L. M.; Glowacki, B. R.; Sage, R.; Tew, P.; Wonnacott, M.; Bates, P.; Mulholland, A. J.; Glowacki, D. R., Sampling molecular conformations and dynamics in a multiuser virtual reality framework. *Science Advances* **2018,** *4* (6), eaat2731.

204.   Goddard, T. D.; Huang, C. C.; Meng, E. C.; Pettersen, E. F.; Couch, G. S.; Morris, J. H.; Ferrin, T. E., UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Science* **2018,** *27* (1), 14-25.

205.   Balo, A. R.; Wang, M.; Ernst, O. P., Accessible virtual reality of biomolecular structural models using the Autodesk Molecule Viewer. *Nature Methods* **2017,** *14* (12), 1122-1123.

206.   Li, H.; Leung, K.-S.; Nakane, T.; Wong, M.-H., iview: an interactive WebGL visualizer for protein-ligand complex. *BMC Bioinformatics* **2014,** *15* (1), 56.

207.   Goddard, T. D.; Brilliant, A. A.; Skillman, T. L.; Vergenz, S.; Tyrwhitt-Drake, J.; Meng, E. C.; Ferrin, T. E., Molecular Visualization on the Holodeck. *Journal of Molecular Biology* **2018,** *430* (21), 3982-3996.

208.   Lv, Z.; Tek, A.; Da Silva, F.; Empereur-Mot, C.; Chavent, M.; Baaden, M., Game On, Science - How Video Game Technology May Help Biologists Tackle Visualization Challenges. *PLoS ONE* **2013,** *8* (3), e57990.