INVESTIGATING THE EFFECTS OF BRCA1 CONSTRUCT LENGTH ON ITS INTERACTION WITH PALB2

BY

Jaehyun Lee

Submitted in partial fulfillment of the

requirements for Departmental Honors in

the Department of Biology

Texas Christian University

Fort Worth, Texas

May 3, 2021

INVESTIGATING THE EFFECTS OF BRCA1 CONSTRUCT LENGTH ON ITS INTERACTION WITH PALB2

Project Approved:

Supervising Professor: Mikaela Stewart, Ph.D.

Department of Biology

Michael Chumley, Ph.D.

Department of Biology

Wendy Williams, Ph.D.

John V. Roach Honors College

ABSTRACT

Mutations in BReast CAncer 1 protein (BRCA1) play a crucial role in DNA damage control such as double-strand DNA break repair mechanisms. Mutations in BRCA1 increase the chance of disrupted genetic integrity by their contributions to the development of breast cancer. BRCA1 must bind to its partner protein, Partner and Localizer to BRCA2 (PALB2), in order to properly carry out its function in the repair mechanism pathway, but its conformation once bound to PALB2 is not clear. In its active state, PALB2 is known to remain in an alpha helical coiled-coil homodimer conformation. Through this observation, we hypothesized that the intrinsically disordered region of BRCA1 on its binding surface will undergo a conformational change into an alpha helical form. In order to test this hypothesis, we first created a truncated BRCA1, making it 50 amino acids long, then conducted nuclear magnetic resonance experiments (NMR). Through the NMR experiments, we found that the binding interface of BRCA1 does change its conformation into a helical state, forming a coiled-coil heterodimer upon binding with PALB2.

ACKNOWLEDGEMENTS

I would like to thank the Texas Christian University College of Science and Engineering for providing the resources needed for this research and my committee members, Drs. Michael Chumley and Wendy Williams, for their insights. Additionally, I would like to thank Davis, Russell, and other members of the Stewart Lab. Especially Christine for her unconditional support and encouragement. I thank my parents, Joonkyu Lee and Moonhee Woo, and my brother, Hakhyun Lee for shaping me into the person I am today and giving me the courage I need. Lastly, I would like to thank Dr. Mikaela Stewart for her knowledge, guidance, and wisdom she has shown me throughout this invaluable experience. There are no adequate words for me to express how much I admire her passion for her students.

TABLE OF CONTENTS

INTRODUCTION

Hundreds of thousands of lives around the world are affected by breast cancer, whether it be direct health issues or seeing loved ones suffer. Breast cancer is the second most common cancer in American women, after skin cancers. In 2021 alone, it is estimated that 281,550 new cases of invasive breast cancer are expected to be diagnosed in women in the U.S[1]. One in eight, or 13%, of U.S. women are projected to develop breast cancer in their life[1]. However, a woman's lifetime risk of developing breast and/or ovarian cancer markedly increases if she inherits a harmful variant in BReast CAncer gene 1 (*BRCA1*). 55-72% of women who inherit harmful mutations in *BRCA1* will develop breast cancer by 70-80 years of age, compared to 13% of those who do not inherit those muations[2]. The inherited mutations of *BRCA1* are detrimental to an individual due to the central role it plays in maintaining the genetic integrity.

DNA damage repair mechanisms are the most crucial aspect of maintaining the integrity of the genome. These mechanisms often involve proteins interacting with each other to prevent further accumulations of potentially harmful mutations. Although there are many forms of DNA damage, double-strand DNA breaks (DSB) are one of the most damaging types for genetic integrity. DSB can arise from DNA's exposure to toxic chemicals or excessive radiation. The tumor-suppressing protein BRCA1 is a key player in repairing DSB, via a process called homologous recombination (HR)[3]. One of the crucial traits of tumor-suppressing proteins is their ability to localize at the site of the DNA damage and repair it[4]. With the localization of tumor-suppressing proteins, the damage repair is further enhanced and aided. If proteins inherit mutations in BRCA1 that cause its loss of function, cellular mutations will accumulate, and cancerous cells are more likely to arise.

Along with BRCA1, PALB2 and BRCA2 are known to be other tumor suppressor proteins crucial for cellular DSB repair. BRCA1 plays an important role in targeting PALB2 and BRCA2 to the DNA damage site[3]. PALB2, in an inactive state, exists as an alpha helical homodimer. However, once DNA damage has incurred, the homodimer dissociates and binds with BRCA1.

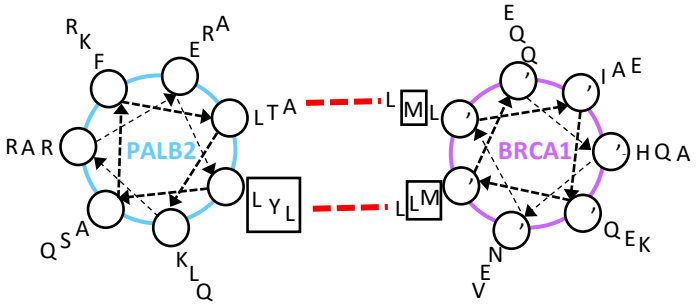


Figure 1. Predicted BRCA1-PALB2 binding interface. The top-down view of PALB2 (residue 9-42) and BRCA1 (residue 1393-1424) helices. The boxed amino acids, which are known to affect binding if mutated, are responsible for the hetero-oligomeric interaction between PALB2 and BRCA1.

The binding interface of BRCA1 to PALB2 is intrinsically disordered, meaning it lacks a definite shape or form. This region of unknown structure further complicates the prediction of risk from inherited BRCA1 mutations and hinders prevention of breast cancer. As shown in Figure 2, it can be predicted that the BRCA1 binding domain will change its shape into an alpha helix, forming a coiled-coil complex upon binding with PALB2 or it could remain intrinsically disordered once bound.

Studying the conformational change of BRCA1 upon binding to PALB2 will allow us to draw a clearer picture in developing a system where we can better project their *in vivo* interactions using an *in vitro* method. Being able to mimic the cellular reactions *in vitro* is highly important as it

allows us to study their interactions further to provide predictions to individuals whose family has a history of mutant BRCA1 inheritance.

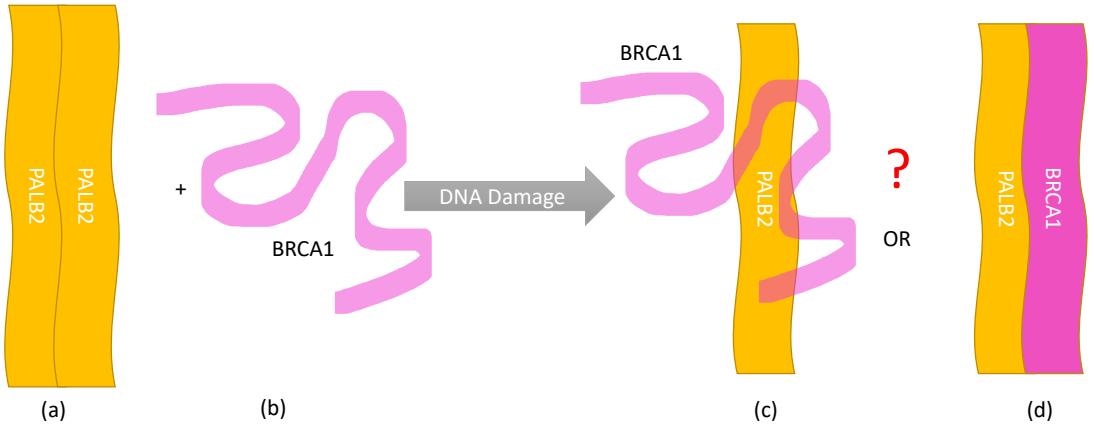


Figure 2. Predictions of BRCA1 conformational change upon binding to PALB2. (a) the homodimer of PALB2 before activation upon DNA damage. (b) the intrinsically disordered region of BRCA1. (c) shows the possibility of BRCA1 remaining intrinsically disordered after binding to PALB2. (d) shows the possibility of BRCA1 taking on an alpha-helix conformation when binding to PALB2.

BRCA1^{91} was truncated to BCRA1^{50} in order to predict the conformation BRCA1^{50} takes upon binding with PALB2. BRCA1 and PALB2 heterodimerize, but the binding domain conformation and the overall structure is unknown[5]. BRCA1's role in the DNA repair mechanism is further understood by studying its length and structure. We hypothesize that our BRCA1^{50} binds to form a coiled-coil alpha helix structure upon binding to the PALB2 region, which is crucial for the DNA repair mechanism.

METHODS

Mutagenesis

In order to obtain a shorter construct of BRCA1, mutagenesis was performed using a template plasmid containing a 91 amino-acid-long segment of human BRCA1 ($BRCA1^{91}$), which spans from Ser-1377 to Glu-1467. To model the shorter construct of this BRCA1 ($BRCA1^{50}$), the site-specific mutagenic technique used the DNA primers that would eventually replace Asn-1427 with a terminating codon. The new primers with a premature stop codon were used to produce the target plasmid, using the fidelity hot start Q5 hot-start polymerase (New English Biolabs) in the polymerase chain reaction thermal cycler with appropriate temperature and time settings. $BRCA1^{50}$ spanned from Ser-1377 to Ser-1426, becoming a 50 amino-acid long segment. During this process a temperature gradient was used to anneal the products of forward and reverse primers together.

Primer Name	Sequence
BRCA1 N1427stop_F	5’ GCA TGG GAG CCA GCC TTC TTA AAG CTA CCC TTC CAT CAT AAG 3’
BRCA1 N1427stop_R	5’ CTT ATG ATG GAA GGG TAG CTT TAA GAA GGC TGG CTC CCA TGC 3’

Table 1. Primer sequences

The obtained mutagenic products were then digested with the enzyme DpnI at 37 °C for two hours. DpnI enzyme, targeting only the methylated strands of the DNA, removed all template parent vectors. When digested, the DNA was then transformed into competent DH5α *E. coli* cells through a heat shock at 42 °C for 45 seconds. The transformed *E. coli* cells were placed in

300 µL of super optimal broth with catabolic repression (SOC medium) and transferred onto lysogeny broth (LB broth) agar plates, which contains 100 µg/mL kanamycin. To maximize colony formation, the agar plates were incubated overnight at 37 °C. A single colony from the agar plate was isolated and placed into the 5 mL of LB broth for 14-16 hours in an incubator at 37 °C and 250 rpm for optimal growth. Then, the plasmids from the *E. coli* cells were purified, using the Qiagen mini-prep kit by following the provided instructions. The plasmid was sequenced to ensure that the intended mutation was present.

Buffer	Purpose	Composition
Nickel (Ni) buffer A pH 7.4	Protein purification	5 mM imidazole, 0.5 M NaCl, 20 mM TRIS
H3C cleavage dialysis buffer pH 7.0	Protein purification	25 mM $NaPO_4$, 150 mM NaCl, 1 mM DTT
NMR Buffer pH 6.5	Protein purification	25 mM $NaPO_4$, 50 mM NaCl
BME loading buffer	Protein purification	0.002% bromophenol blue, 10% glycerol, 5% 2-mercaproethanol, 2.5% SDS, 62.5 mM TRIS-HCl
Coomassie blue	Gel staining	1 g Coomassie, 10% acetic acid, 50% methanol, 40% milli-Q water
Destaining buffer	Gel destaining	10% glacial acetic acid, 50% methanol, 40% milli-Q water

Table 2. Buffers used in experiments

Transformation and Protein Purification

The plasmid for BRCA1[50] was transformed into BL21(DE3) component *E. coli* cells. Similar to mutagenesis, cells were treated with a heat shock at 42 °C for 30 seconds for optimal outcomes. The transformed cells were left to recover and grow and were then plated on LB agar plates, containing 50 µg/mL kanamycin. After incubating the plates overnight at 37 °C, the suspended colonies were harvested off the plate and grown in 1 L of autoclaved LB with 50 µg/mL

kanamycin. Once reaching 0.6 optical density (OD), the cells were induced with 0.200 mM IPTG, shaking at 250 rpm overnight at 16 °C. Following induction, the cell cultures from the shaker were centrifuged at 3500 rpm, 4 °C, and 20 minutes. The pellets formed were resuspended in a 25 mL mixture of Ni Buffer A (Table 2), lysozyme, bovine DNaseI (Goldbio), and general use protease inhibitor cocktail (AMRESCO). The resuspended cells were then lysed with the Vibra Cell sonicator by Sonics (15 seconds pulse, 30 seconds rest, 85% amplitude for 10 minutes of pulse time) while making sure that the resuspended cells were kept cold on ice. The sonicated solution was centrifuged at 14000 RCF, 4 °C for 25 minutes in order to separate the soluble proteins from the insoluble cellular debris in the pellet. After separating the protein supernatant, a sample of the protein solution was mixed in a one-to-one ratio with the BME loading buffer (Table 2) to prepare samples to load onto a gel to confirm the presence of the proteins. This sampling process was repeated after resuspending the collected pellet.

The cobalt affinity column purification, using the Histidine (His)-Trap TALON column (GE healthcare), was used to isolate $BRCA1^{50}$ from the cell lysate loaded into the Äkta Start system, following the manufacturer instructions. The His-tag found on $BRCA1^{50}$ has a strong affinity to the cobalt column and makes the protein isolation possible. Consequently, $BRCA1^{50}$ binds to the column while other proteins potentially present in the lysate do not. The desired protein was eventually collected in fractions through imidazole washed in increasing concentrations, which displaces the proteins from the cobalt column as imidazole has a stronger affinity to the cobalt column than the His-tags do.

The $BRCA1^{50}$-containing fractions were collected, and each fraction sample was taken and mixed with BME loading dye (Table 2). The fractions were then combined altogether and

dialyzed overnight in the H3C cleavage dialysis buffer (Table 2). The protein was treated with the human rhinovirus type 14 protease (H3C) for an hour to promote cleavage of the His-SUMO tag from the BRCA1^{50} construct after the overnight dialysis. The H3C-treated protein solution was admitted through both the glutathione and nickel affinity columns in sequence to remove glutathione-tagged H3C protease and His-tagged SUMO.

Protein Concentration

The dialyzed protein solution was retrieved out of the buffer and was concentrated in a 3,000 Dalton cut-off concentrator (PALL) in a centrifuge at 3000 rpm and 4 °C in ten-minute increments until the protein concentrate measured 250 μL. PALB2 has a non-zero extinction coefficient, therefore its concentration was measured at its absorbance at 280 nm on the nano drop. Serial dilutions of BRCA1^{50} were loaded onto the SDS PAGE gel, and varying band thickness were compared with the PALB2 band thickness with the known concentration from the nano drop. Comparing band thickness allowed qualitative measurement of the protein concentration.

Secondary Structure Analysis

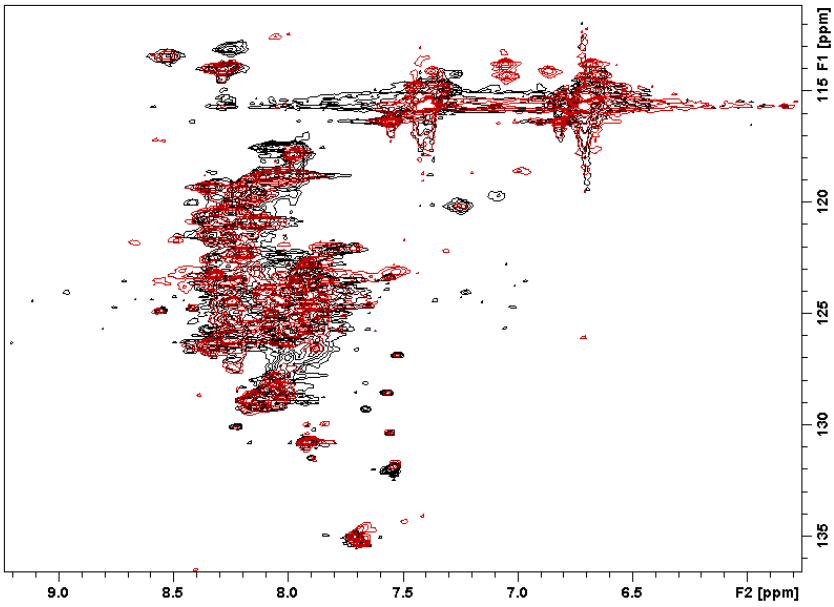
HSQC spectra were collected on 600 MHz Avance NEO 600 (Bruker) spectrometer at Texas A&M University's Biomedical NMR facility.

RESULTS

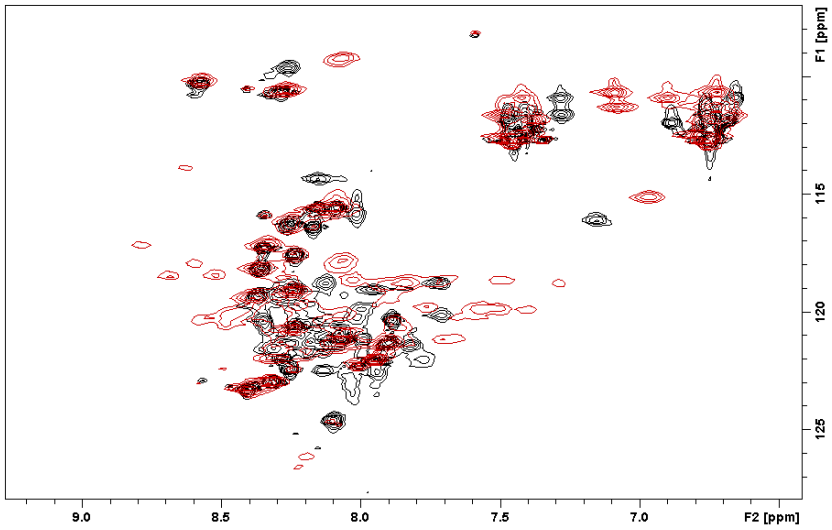
Protein structure was investigated using nuclear magnetic resonance (NMR). When studying the patterns of NMR spectra, it is important to understand the fundamental structures of proteins. Proteins are assembled from amino acids, all of which contain an amino functional group on the N-terminus and a carboxylic acid functional group on the C-terminus. Between these functional groups is the carbon where the sidechain of the amino acid is located, which is often referred to as the R group. To form protein, amino acids are joined with each other through a dehydration reaction where the amino group attacks the electrophilic carbon on the C-terminus of another amino acid. This is referred to as a peptide bond and has a partial double bond characteristic, which gives the protein stability.

The position of each peak in the NMR spectrum represents the chemical environment of the nitrogen and hydrogen atoms in the backbone of the amino acid it originates from. The peak position is referred to as its "chemical shift." Proline is the only amino acid that is not detected through this method due to its lack of an amide proton when bound to another amino acid through peptide bonds. The y-axes of Figures 3, 4, and 5 represent the chemical shift of nitrogen atoms whereas the x-axes show the chemical shift of hydrogen atoms in parts per million. Shown in the black spectrum of Figure 3 is the chemical shifts of the backbone nitrogen and hydrogen in the ^{15}N BRCA1^{91} whereas the red spectrum shows the chemical shifts of the backbone of nitrogen and hydrogen in the ^{15}N BRCA1^{91} construct upon the addition of 2x PALB2 that is not isotopically labeled. This makes PALB2 undetectable in the spectrum but allows us to see the effects of PALB2 binding on BRCA1. Therefore, the patterns observed on the NMR spectra are solely that of BRCA1 hydrogen and nitrogen atoms. The number of contour circles within a peak

represents the peak intensity; intensity is inversely proportional to protein flexibility and size. We can see that without the addition of PALB2 (black peak in Figure 3) the contour rings are generally more abundant, which means that BRCA1 is relatively flexible and small in size. Upon the introduction of PALB2 (red peaks in Figure 3), the contours of the peaks decrease in number, indicating that the chemical environments of the hydrogen and nitrogen atoms in BRCA1 changed most likely due to the interaction between two proteins. Further evidence of the BRCA1 and PALB2 interaction, can be observed by expanding two regions of the NMR spectra from Figure 3. The expanded regions of the spectra are presented in Figures 4 and 5.



(a)



(b)

Figure 3. BRCA1 construct displayed in NMR spectra. (a) ^{15}N BRCA1^{91} shown in black and ^{15}N BRCA1^{91}+ 2x PALB2 shown in red. (b) ^{15}N BRCA1^{50} shown in black and ^{15}N BRCA1^{50} + 2x PALB2 shown in red

Figure 4 is an excerpt of the glycine region of the spectra from the upper left corner of Figure 3. Glycine is known to be a helical breaker in protein folding, meaning that it is usually found on the ends of the alpha helices or in unstructured regions rather than the middle of alpha helices. If BRCA1^{50} undergoes a conformational change into an alpha helix upon binding to PALB2 as predicted, it can be speculated that glycine is not likely to be directly involved in the protein-protein binding due to their helix-breaking nature. However, the specific structure of the binding interface is not clearly presented, and this is speculation based on scientific reasoning. Therefore, it is unlikely that this glycine is directly interacting with PALB2.

In Figure 4, each peak in the spectra is randomly labeled to conveniently reference them. The contours of the black peaks are more abundant (more rings) than the peaks originating from the same amino acids when PALB2 is present. Along with the changes in the contour, GlyC in Figue

4 (a) and (b) show shifts in the peaks, indicating that the chemical environments change upon the introduction of PALB2. This indicates that GlyC is either in close proximity to PALB2 in the bound state, or that this region of BRCA2 takes on a different structure upon binding to PALB2. However, GlyA and GlyB in ^{15}N BRCA1^{91} and ^{15}N BRCA1^{50} do not show a significant shift in the peaks, indicating these amino acids are distant from the PALB2 binding region. This region of the spectrum allows us to see that there is some similarity in PALB2 binding from the different lengths of BRCA1, as the same peak shifts and in a similar direction regardless of length.

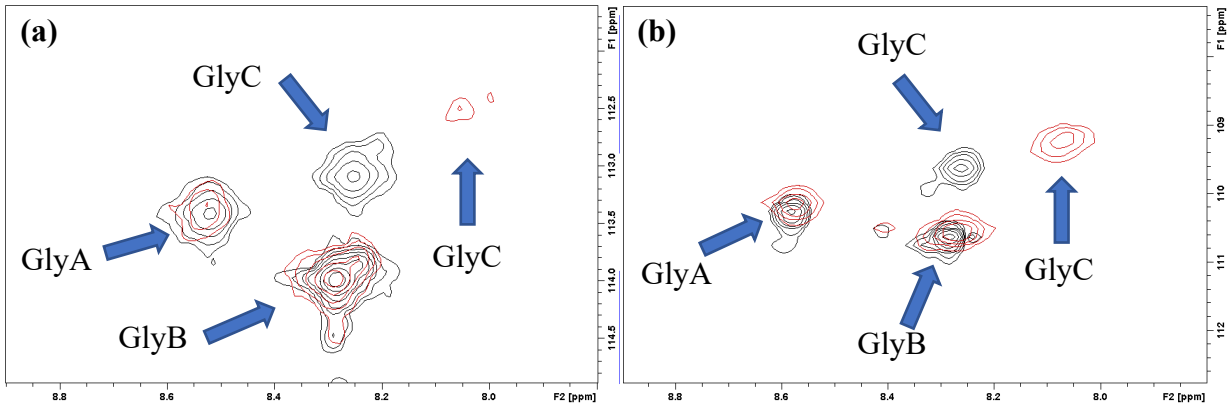


Figure 4. (a) Glycine region of ^{15}N BRCA1^{91} +/- 2x PALB2. (a) Glycine region of ^{15}N BRCA1^{50}+/- 2x PALB2

Additionally, Figure 5 is an excerpt of Figure 3, specifically focusing on the non-glycine backbone region of BRCA1. The spread of ^{1}H range in Figure 5 can be analyzed to study the interactions between proteins. The range of the hydrogen atoms in ^{15}N BRCA1^{50} is much broader than that of ^{15}N BRCA1^{91}, indicating that the hydrogen atoms in ^{15}N BRCA1^{50} experience a greater degree of chemical environment change. Similar to the glycine region discussed and shown in Figure 4, the contour change upon binding with PALB2 in ^{15}N BRCA1^{50} is more prominent than that of ^{15}N BRCA1^{91}.

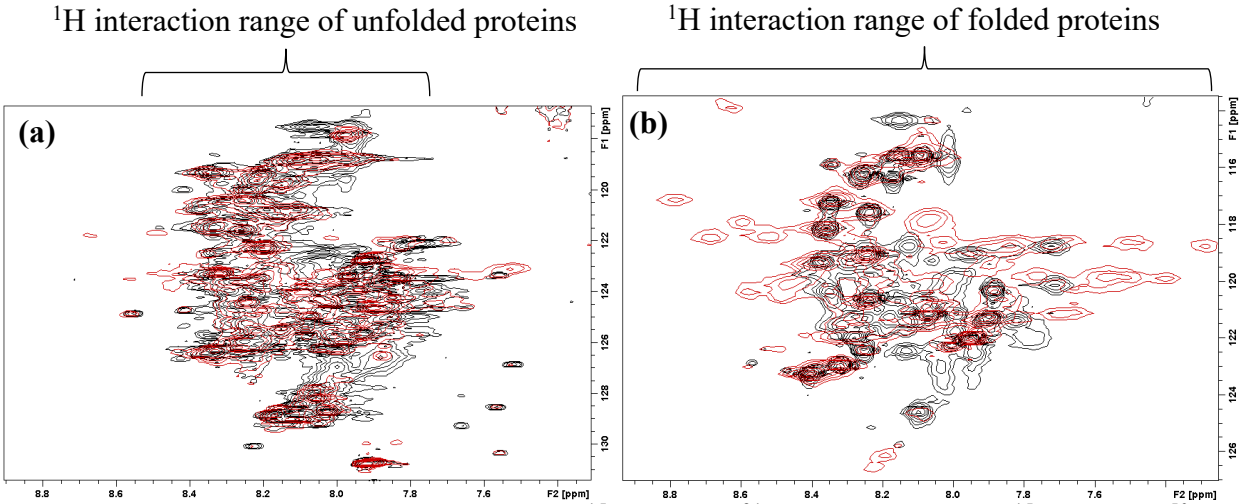


Figure 5. Backbone region of proteins. (a) ^{15}N BRCA1^{91} +/- 5x PALB2. (b) ^{15}N BRCA1^{50} +/- 2x PALB2

DISCUSSION

Our project focused on studying the secondary structure of BRCA1 upon binding with PALB2. BRCA1 plays an integral role in DNA damage repair mechanisms by recruiting and interacting with PALB2[3]. Together, they help to maintain the genomic integrity, and cancer risk increase correlates to the increase in lack of or hindered BRCA1 and PALB2 interaction[6]. Studying the intrinsically disordered binding interface of BRCA1 is important as it may provide additional information of its secondary structures when bound to different partners in different mechanisms[7].

The overall PALB2 and BRCA1 heterodimer structure and the conformation of the binding domain on BRCA1 lacks practical biochemical evidence. However, it is hypothesized to be a coiled-coil structure as PALB2 exists in a homodimer conformation in its inactive state[5]. This prediction is better displayed in Figure 1 as the binding affinity is known to decrease when the boxed amino acids are modified or removed[5]. Figure 1 supports the alpha helical model, but there was no confirmed data to support this helical model. Our findings provide sufficient evidence to support our initial hypothesis that the intrinsically disordered region of BRCA1 takes up an alpha helical structure upon binding with PALB2. The supporting evidence was successfully obtained by truncating $BRCA1^{91}$ into $BRCA1^{50}$, isotopically labeling the nitrogen atoms in the amino acid backbones (^{15}N) and studying the secondary structures of BRCA1 through NMR experiments. The experimental data of these protein interactions could play a crucial role in genetic counseling when informing patients of their possible risk due to inherited mutations as well as performing preventative surgeries only when necessary, as those procedures are too expensive and invasive to undergo if they do not have detrimental mutations. Our results

could be a foundation in determining whether the patients' inherited mutations are detrimental or not. With the structural information, physicians can provide a clearer picture to the patient regarding the risk level from the mutations.

Although we do not have the exact atoms that are involved in the binding interface of BRCA1 to PALB2, we can induce by the known characteristics of the amino acids common to the binding region. Figure 4 shows the glycine region excerpt of Figure 3, and these glycine peaks were chosen as they show consistency in their shift patterns. Between $BRCA1^{91}$ and $BRCA1^{50}$, the peaks shift in the similar direction, confirming their involvement in protein binding. From our data, there is not a clear way to confirm whether this region is directly involved in the protein-protein binding interaction between BRCA1 and PALB2. Glycine functions as a helix breaker in protein structures, so it is our speculation based on scientific reasoning that glycine is not present in the binding interface of the protein, but this region may undergo a conformational change of some sort upon binding.

In order for BRCA1 to carry out its functions in DSB, it must properly bind to PALB2. However, the conformation of the intrinsically disordered region of BRCA1 is not yet clear. "Fuzzy binding" refers to the conformation and interaction change of an intrinsically disordered region when bound to a partner protein[7]. The protein conformation in fuzzy binding not only depends on the conformation the protein it takes in its bound state but also the context of the binding interaction[7]. The chemical shifts of hydrogen and nitrogen atoms in Figure 5 show that the environment underwent significant changes. The shifts also display lower intensity contour and wider peaks, which allows us to induce that the proteins form secondary structures, although the

structure is unlikely to be a beta sheet. Beta sheet peaks tend to be more diverse, and the chemical shift range of the NMR data is relatively narrow.

REFERENCES

1. "How Common is Breast Cancer?". https://www.cancer.org/cancer/breast-cancer/about/how-common-is-breast-cancer.html.
2. "BRCA Gene Mutations: Cancer Risk and Genetic Testing.". https://www.cancer.gov/about-cancer.
3. Zhang, F.; Ma, J.; Wu, J.; Ye, L.; Cai, H.; Xia, B.; Yu, X., PALB2 links BRCA1 and BRCA2 in the DNA-damage response. *Curr Biol* **2009,** *19* (6), 524-9.
4. Sy, S. M.; Huen, M. S.; Chen, J., PALB2 is an integral component of the BRCA complex required for homologous recombination repair. *Proc Natl Acad Sci U S A* **2009,** *106* (17), 7155-60.
5. Song, F.; Li, M.; Liu, G.; Swapna, G. V. T.; Daigham, N. S.; Xia, B.; Montelione, G. T.; Bunting, S. F., Antiparallel Coiled-Coil Interactions Mediate the Homodimerization of the DNA Damage-Repair Protein PALB2. *Biochemistry* **2018,** *57* (47), 6581-6591.
6. Foo, T. K.; Tischkowitz, M.; Simhadri, S.; Boshari, T.; Zayed, N.; Burke, K. A.; Berman, S. H.; Blecua, P.; Riaz, N.; Huo, Y.; Ding, Y. C.; Neuhausen, S. L.; Weigelt, B.; Reis-Filho, J. S.; Foulkes, W. D.; Xia, B., Compromised BRCA1-PALB2 interaction is associated with breast cancer risk. *Oncogene* **2017,** *36* (29), 4161-4170.
7. Fuxreiter, M., Classifying the Binding Modes of Disordered Proteins. *Int J Mol Sci* **2020,** *21* (22).