Development of Natural Cyclic Peptide Inhibitors of XRCC4/XLF Interaction for Radio-Sensitization of Breast Tumor Cells

MOHAINI MOHAMMED AL
Virginia Commonwealth University

Follow this and additional works at: https://scholarscompass.vcu.edu/etd
Part of the Medical Pharmacology Commons

© The Author

Downloaded from
https://scholarscompass.vcu.edu/etd/384
Development of Natural Cyclic Peptide Inhibitors of XRCC4/XLF Interaction for Radio-
Sensitization of Breast Tumor Cells

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science
at Virginia Commonwealth University

By

Mohammed Al Mohaini
B.Sc. in Pharmaceutical Sciences, King Saud University, 2006

Director: Lawrence F. Povirk, Professor, Department of Pharmacology and Toxicology

Virginia Commonwealth University
Richmond, Virginia
May, 2012
ACKNOWLEDGEMENT

I would like to thank my parents, Ahmad Almohaini and Hejjiah Almutlaq, for their prayer, care, love and support throughout my life. High regards also go to my brothers and sisters, especially Abbas and Yousef, for their sincere encouragement and inspiration throughout my research work. I must thank my lovely wife Zainab Alramadhan for her love, understanding and taking care of me, my son Hasan and daughter Fatimah during my study.

I would like to express my deepest gratitude to my advisor, Dr. Lawrence Povirk for his excellent guidance, caring, patience, and advice that are necessary for me to proceed through the master program and complete my thesis. I warmly thank Dr. Matthew Hartman for his valuable advice and encouragement. His extensive knowledge and experience in mRNA display technique have a great impact on this project. Also, I would like to thank my committee member Dr. Richard Moran for his insightful comments and hard questions.

I would like to thank the current and former members of Dr. Povirk’s lab. Special thanks goes to Dr. Konstantin Akopiants for the training I received when I came to the lab in Fall 2010. His advice and help in cloning are highly appreciable. Also, I would like to thank the current and former members of Dr. Hartman’s lab for their help and advice in the selection process.

Finally, I would like to thank King Saud bin Abdulaziz University for Health Sciences (KSAU-HS) for awarding me a scholarship to study abroad. I am so thankful to have KSAU-HS support me financially in my goals of getting a master degree. I would like also to thank Ali Alkhamis, Yousef Alhashem, Abdulkhaliq Alsalman and Ali Alhammad for their support and being great friends.
TABLE OF CONTENT

List of Figures ........................................................................................................................................... vi
List of Abbreviations.................................................................................................................................... viii
Abstract.......................................................................................................................................................... x

I. Introduction ..................................................................................................................................................... 1
  1.1 Breast Cancer .............................................................................................................................................. 1
  1.2 Radiation Therapy ...................................................................................................................................... 2
  1.3 Homologous Recombination (HRR) ........................................................................................................ 3
  1.4 Non-Homologous End Joining (NHEJ) .................................................................................................... 3
  1.5 The Choice of DSBs Repair Method ......................................................................................................... 6
  1.6 The Choice of DSBs Repair Mechanism in Breast Cancer ....................................................................... 6
  1.7 XRCC4-XLF Interaction as a Target for Radiosensitization ................................................................... 7
  1.8 mRNA Display .......................................................................................................................................... 12

II. Methods ....................................................................................................................................................... 14
  2.1 Production of Wild Type GST- XRCC4\textsuperscript{1-157}, Double Mutant GST - XRCC4\textsuperscript{1-157} (M61R and F106E) and Double Mutant Trx - XRCC4\textsuperscript{1-157} Protein ........................................................................ 14
     2.1.1 Construction of Wild Type XRCC4\textsuperscript{1-157} .............................................................................. 14
     2.1.2 Insertion of XRCC4\textsuperscript{1-157} into the pGEX-4T-1 Expression Plasmid Vector .................. 15
     2.1.3 Construction of Double Mutants XRCC4\textsuperscript{1-157} (M61R and F106E) .............................. 16
     2.1.4 Insertion of Double-Mutants XRCC4\textsuperscript{1-157} (M61R and F106E) Gene into pET-32XT Expression Vector ...................................................................................................................... 16
     2.1.5 Production of Wild Type and Double-Mutant XRCC4\textsuperscript{1-157} Protein ................................ 17
     2.1.6 Ni-NTA Purification .......................................................................................................................... 18
     2.1.7 Thrombin Cleavage of GST- XRCC4\textsuperscript{157} DM Protein .................................................. 19
  2.2 mRNA Display and In Vitro Selection ....................................................................................................... 19
2.2.1 DNA library Synthesis .................................................................19
2.2.2 PCR Amplification of Library DNA ..............................................21
2.2.3 StrataClone Cloning and Sequencing ........................................21
2.2.4 Transcription and Purification of mRNA .......................................22
2.2.5 Psoralen Photo-Crosslinking .......................................................23
2.2.6 In vitro translation .................................................................24
2.2.7 Oligo(dT) Purification and Cyclization .......................................24
2.2.8 Reverse Transcription ............................................................25
2.2.9 Ni-NTA Purification ...............................................................26
2.2.10 In Vitro Selection .................................................................27
2.2.11 PCR Amplification of Selected Fusions .....................................28

III. Results ..........................................................................................35

3.1 Cloning of Wild Type XRCC4<sup>157</sup> into pGEX-4T-1 Vector and Production of WT GST- XRCC4<sup>157</sup> Protein .................................................................35
3.2 Cloning of Double-mutant XRCC4<sup>157</sup> (M61R and F106E) into pGEX-4T-1 Vector and Production of DM GST-XRCC4<sup>157</sup> Protein ........................................43
3.3 Thrombin Cleavage of GST- XRCC4<sup>157</sup> DM Protein ....................47
3.4 Cloning of Double-mutant XRCC4<sup>157</sup> (M61R and F106E) into pET-32XT Vector and Production of DM Thioredoxin (Trx) - XRCC4<sup>157</sup> Protein ...............................49
3.5 mRNA Display and In Vitro Selection .............................................55
3.5.1 Preparation of the DNA libraries .............................................55
3.5.2 Small Scale of In Vitro Transcription and Translation .................59
3.5.3 In Vitro Selection .....................................................................62
3.5.3.1 First Round ........................................................................62
3.5.3.2 Second Round .....................................................................63
3.5.3.3 Third Round ........................................................................64

IV. Discussion and Future Direction ....................................................67

4.1 Therapeutic potential of suppressing DSB repair ............................67
4.2 Future Directions ........................................................................70
References..............................................................................................................................................75
Vita..............................................................................................................................................................83
**LIST OF FIGURES**

| Figure 1-1 | Nonhomologous end joining ................................................................. | 5 |
| Figure 1-2 | Crystal structures of XRCC4 and XLF ...................................................... | 9 |
| Figure 1-3 | Filaments of alternating XRCC4-XLF dimers ............................................. | 10 |
| Figure 1-4 | Interface of the XRCC4\(^{1-157}\)-XLF\(^{1-224}\) complex formed by the distal part of their N-terminal head domains ............................................................... | 11 |
| Figure 2-1 | *In vitro* transcription, photo crosslinking and *in vitro* translation .......... | 30 |
| Figure 2-2 | mRNA-Peptide Fusions Purification Steps .................................................. | 31 |
| Figure 2-3 | Pre-Clearing (Negative Selection) with GST Protein Only ........................ | 32 |
| Figure 2-4 | Pre-Clearing (Negative Selection) with GST-XRCC4\(^{157}\) Protein ............... | 33 |
| Figure 2-5 | *In vitro* Selection .................................................................................. | 34 |
| Figure 3-1 | The strategy of cloning the WT XRCC4\(^{157}\) gene into the pGEX-4T-1 vector and production of GST tagged XRCC4\(^{157}\) protein ................................................ | 38 |
| Figure 3-2 | The XRCC4\(^{157}\) fragment and pGEX-4T-1 Vector Cut ................................ | 39 |
| Figure 3-3 | Cloning of Wild Type XRCC4\(^{157}\) into pGEX-4T-1 Vector ........................ | 40 |
| Figure 3-4 | XRCC4\(^{157}\) sequence ........................................................................... | 41 |
| Figure 3-5 | Protein Composition of WT GST-XRCC4\(^{157}\) as determined by SDS-PAGE ...... | 42 |
| Figure 3-6 | The strategy of cloning the DM XRCC4\(^{157}\) gene into the pGEX-4T-1 vector and production of GST tagged XRCC4\(^{157}\) protein ................................................ | 44 |
| Figure 3-7 | Double-mutant XRCC4\(^{157}\) sequence ....................................................... | 45 |
| Figure 3-8 | Protein Composition of DM GST-XRCC4\(^{157}\) as determined by SDS-PAGE ...... | 46 |
| Figure 3-9 | Thrombin cleavage of GST-XRCC4\(^{157}\) DM as determined by SDS-PAGE ...... | 48 |
| Figure 3-10 | The strategy of cloning the DM XRCC4\(^{157}\) gene into the pET-32XT vector and production of Thioredoxin tagged XRCC4\(^{157}\) protein ................................................ | 51 |
| Figure 3-11 | DNA segment encoding DM XRCC4\(^{157}\) (M61R and F106E) was cut out from the pGEX-4T-1 vector and the pET-32XT was linearized ................................................ | 52 |
**LIST OF ABBREVIATIONS**

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>Å</td>
<td>angstrom</td>
</tr>
<tr>
<td>APS</td>
<td>ammonium persulfate</td>
</tr>
<tr>
<td>ATM</td>
<td>Ataxia Telangiectasia Mutated</td>
</tr>
<tr>
<td>ATP</td>
<td>Adenosine triphosphate</td>
</tr>
<tr>
<td>BRCA</td>
<td>Breast cancer susceptibility protein</td>
</tr>
<tr>
<td>BSA</td>
<td>bovine serum albumin</td>
</tr>
<tr>
<td>ºC</td>
<td>Celsius</td>
</tr>
<tr>
<td>Ca$^2+$</td>
<td>Calcium</td>
</tr>
<tr>
<td>CD</td>
<td>Circular dichroism</td>
</tr>
<tr>
<td>CDK</td>
<td>cyclin-dependent kinase</td>
</tr>
<tr>
<td>cDNA</td>
<td>complementaryDNA</td>
</tr>
<tr>
<td>DNA</td>
<td>deoxyribonucleic acid</td>
</tr>
<tr>
<td>DNA-PK</td>
<td>DNA-dependent protein kinase</td>
</tr>
<tr>
<td>DNA-PKcs</td>
<td>DNA-dependent protein kinase catalytic subunit</td>
</tr>
<tr>
<td>DSBs</td>
<td>double-strand breaks</td>
</tr>
<tr>
<td>DTT</td>
<td>Dithiothreitol</td>
</tr>
<tr>
<td>EDTA</td>
<td>ethylenediaminetetraacetic acid</td>
</tr>
<tr>
<td>EtBr</td>
<td>Ethidium bromide</td>
</tr>
<tr>
<td>HRR</td>
<td>homologous recombination repair</td>
</tr>
<tr>
<td>HPLC</td>
<td>high performance liquid chromatography</td>
</tr>
<tr>
<td>IR</td>
<td>ionizing radiation</td>
</tr>
<tr>
<td>Kb</td>
<td>kilobase</td>
</tr>
<tr>
<td>kD</td>
<td>kilodalton</td>
</tr>
<tr>
<td>µl</td>
<td>microliter</td>
</tr>
<tr>
<td>Min</td>
<td>minute</td>
</tr>
<tr>
<td>Term</td>
<td>Description</td>
</tr>
<tr>
<td>----------</td>
<td>--------------------------------------------------</td>
</tr>
<tr>
<td>Mre11</td>
<td>meiotic recombination 11 homolog</td>
</tr>
<tr>
<td>MRN</td>
<td>Mre11/Rad50/Nbs1</td>
</tr>
<tr>
<td>NHEJ</td>
<td>nonhomologous end joining</td>
</tr>
<tr>
<td>PAGE</td>
<td>polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>PBS</td>
<td>phosphate buffered saline</td>
</tr>
<tr>
<td>PCR</td>
<td>polymerase chain reaction</td>
</tr>
<tr>
<td>RAD50</td>
<td>family of RADiation sensitive genes(50 homolog)</td>
</tr>
<tr>
<td>RNA</td>
<td>ribonucleic acid</td>
</tr>
<tr>
<td>Rpm</td>
<td>rotations per minute</td>
</tr>
<tr>
<td>RT</td>
<td>room temperature</td>
</tr>
<tr>
<td>SCID</td>
<td>severe combined immune-deficiency</td>
</tr>
<tr>
<td>Sec</td>
<td>seconds</td>
</tr>
<tr>
<td>SDS</td>
<td>sodium dodecyl sulfate</td>
</tr>
<tr>
<td>ssDNA</td>
<td>single-strand DNA</td>
</tr>
<tr>
<td>TBE</td>
<td>Tris/borate/EDTA</td>
</tr>
<tr>
<td>TEMED</td>
<td>N, N, N’, N’-tetramethylenediamine</td>
</tr>
<tr>
<td>UV</td>
<td>ultraviolet</td>
</tr>
<tr>
<td>V</td>
<td>volts</td>
</tr>
<tr>
<td>V(D)J</td>
<td>variable, diversity, joining</td>
</tr>
<tr>
<td>WT</td>
<td>wild type</td>
</tr>
<tr>
<td>XLF</td>
<td>XRCC4-like factor</td>
</tr>
<tr>
<td>XRCC</td>
<td>X-ray cross complement protein</td>
</tr>
</tbody>
</table>
ABSTRACT

DEVELOPMENT OF NATURAL CYCLIC PEPTIDE INHIBITORS OF XRCC4/XLF INTERACTION FOR RADIO-SENSITIZATION OF BREAST TUMOR CELLS

By Mohammed Al Mohaini, B.Sc. Pharm

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science at Virginia Commonwealth University

Virginia Commonwealth University, 2012

Advisor: Lawrence F. Povirk, Professor, Department of Pharmacology and Toxicology

Breast cancer is the second leading cause of cancer death in women according to the American Cancer Society. The standard treatment regimen for breast cancer involves ionizing radiation combined with surgery and chemotherapy. Ionizing radiation induces a complex signaling response in cells resulting in either growth arrest, senescence or cell death, and the cell killing after exposure to radiation results largely from DNA double-strand breaks (DSBs). There are two main mechanisms in mammalian cells responsible for repairing the DSBs; the primary mechanism is non-homologous end joining (NHEJ) and the secondary mechanism is homologous recombination (HRR). Previous studies showed that breast tumor cells depend mainly on NHEJ for repairing induced DNA damage. XRCC4 and XLF are two essential proteins in the NHEJ process. The interaction between XRCC4 and XLF (also called Cernunnos) is responsible for stimulating ligase IV for rejoining DNA ends. A single mutation on the XLF-binding interface of XRCC4 at M61, F106, M59 or D58 has been shown to disrupt
its interaction with XLF and thus inhibiting NHEJ. Therefore, it is proposed that small natural cyclic peptides that bind to the XLF interface of XRCC4 near M61 and F106 can be identified through an mRNA display \textit{in vitro} selection, and these peptides will inhibit NHEJ and thereby radiosensitize breast tumor cells. We have synthesized five DNA libraries that produced mRNA-peptide fusions containing a trillion unique peptide sequences that will be used for the selection of peptide inhibitors of the XRCC4/XLF interaction, and we have verified their randomness. Tagged wild-type and mutant versions of the head domain of XRCC4 protein, containing the XLF binding site, were successfully purified, and the wild-type version was applied to initial stages of selection of inhibitory peptides by mRNA display.

The percentage of the mRNA-peptide fusions that bound to the XRCC4\textsuperscript{157} after the first round was 2.1\%. The recovery after the second and third rounds was 1.14\% and 2\%, respectively. Results obtained thus far, although preliminary, suggest that the mRNA display method can be successfully applied to the XLF/XRCC4 interaction.
I. INTRODUCTION

1.1 Breast Cancer

Breast cancer in the United States is one of the most common cancers among women. It is the second leading cause of cancer death in women, surpassed only by lung cancer (ACS, 2012). Also, breast cancer is the number one cancer killer in women aged 20-59 years old (Ahmedin et al., 2010). There is a chance of one out eight women that will have an invasive breast cancer some time during their lives. Also, one out of 36 women expected to die in 2012 will do so as a result of breast cancer death.

Breast cancer treatment requires multiple approaches. The recommended treatment for most patients with “invasive” breast cancer is surgical removal of primary tumor, systemic chemotherapy and/or hormonal therapy as well as radiotherapy (Buchholz, 2009). Breast irradiation is indicated for most patients who undergo breast-conserving surgery (lumpectomy). Also, radiation therapy is recommended after mastectomy for patients who are at high risk of developing recurrence (Shenkier et al., 2004). The current treatment available for breast cancer has shown to be quite effective in suppressing the breast cancer growth and improving patients’ quality of life (Buchholz, 2009). However, the risk of breast cancer recurrence is still not uncommon. Recurrent breast cancer can occur months or years after the initial treatment. Breast cancer recurrence can reappear as local recurrence at the original tumor site (invasive or noninvasive ipsilateral breast tumor recurrence), as regional recurrence in the axilla, chest muscle, internal mammary lymph nodes or supracavicular fossa lymph nodes, or as distant...
recurrence (metastasis) in which the tumor cells leave the breast tissue and spread to other sites including bone, lungs, brain and other organs (Millar et al., 2009; Clarke et al., 2005; Bruce et al., 2006). The prognosis of recurrence and metastatic breast cancer is generally poor and associated with high resistant to further treatment and reduced survival rate (Dean-Colomb & Esteva, 2008; Gonzalez-Angulo et al., 2007; Jameel et al., 2004).

1.2 Radiation Therapy

Radiation therapy is an essential part for the management of all stages of breast cancer (Tahernia et al., 2010). In general, ionizing radiation causes chain reactions in cells leading to formation of multiple reactive free radicals causing damage in the DNA. This stimulates a DNA damage response (DDR) that promotes multiple signal transduction cascades, which coordinate DNA repair, cell cycle progression and cell death modes to facilitate the accurate transmission of genetic material after DNA damage (Kesari et al., 2011). The initiation of DDR starts by the activation of Mre11-Rad50-Nbs1 (MRN) complex and recruitment of ataxia-telangiectasia mutated (ATM) protein (D’Amous & Jackson, 2002; Shiloh, 2003). Exposure of cells to therapeutic radiation can lead to different consequences including growth arrest, senescence, or one of multiple modes of cell death such as apoptosis, mitotic catastrophe, autophagy and necrosis (Abbott & Holt, 1999; K. Chu et al., 2004; Lehmann et al., 2007; Ullrich et al., 2008). Nevertheless, the tumoricidal activity of IR is found to be largely related to the formation of DSBs (Iliakis, 1991).

There are two major mechanisms to repair DSBs; that are activated when DSBs occur including homologous recombination (HR) and non-homologous end joining (NHEJ) (Valerie & Povirk, 2003; Lieber, 2010).
1.3 Homologous Recombination (HRR)

HR requires a sister chromatid to act as a template for repair; therefore, HR is only active in late S and G2 phases of the cell cycle and it is an error-free repair pathway (Helleday et al., 2007). HRR is highly efficient in repairing DSBs due to replication forks collapse, IR and inter-strand cross-links (Sung & Klein, 2006). In HRR, the DSB is first detected by the MRN complex, and the DNA ends subjected to 5'→3' resection by CtIP, to generate a 3'-single strand DNA (ssDNA) overhang (Li and Heyer, 2008). Also, CtIP is known to be regulated by CDK and ATM (Chen et al, 2008; Li et al, 2000). After that, replication protein A (RPA) binds to the 3'-ssDNA overhang to prevent premature strand invasion. Then, different proteins including Rad52, BRCA2 and the Rad51 paralogues (Rad51B, Rad51C, Rad51D, XRCC2 and XRCC3) are recruited to replace RPA by RAD51 (San et al., 2008). This allows Rad51 to form a filament to invade a homologous sequence. After that, DNA polymerase η is recruited to extend the 3'-DNA end which leads to capturing a second DSB end to form Holliday junctions; these junctions are resolved in a process involving different proteins resulting in crossover or non-crossover products (Helleday et al., 2007; Li & Heyer, 2008).

1.4 Non-Homologous End Joining (NHEJ):

NHEJ is the primary repair mechanism of DSBs, and it is the main repair mechanism in G0 and G1; even though it can function throughout the cell cycle (Rothkamm et al., 2003). Also, NHEJ is essential mechanism in repairing DSBs due to V(D)J recombination (Jankovic et al., 2007). There are three main steps in the NHEJ pathway including the detection of DSBs, removal of the non-ligatable end groups and religation of the processed DNA ends (Williamson et al., 2009). In other words, repairing DSBs by NHEJ pathway involves nuclease, polymerase, and ligase activities (Lieber, 2010). Several studies have shown the importance of some proteins
in NHEJ pathway in human cells. These proteins are Ku70/80 heterodimer (Ku), DNA-dependent protein kinase catalytic subunit (DNA-PKcs), DNA ligase IV, XRCC4 and XLF (Cernunnos) (Lieber, 2010). Once DBSs occur, the first step in NHEJ is the detection of the breaks, which is achieved by Ku in which Ku heterodimer is recruited to the end of the DNA damage (Weterings & Chen, 2008). When Ku binds, it acts as a scaffold to facilitate the binding of other NHEJ proteins. Also, Ku recruits DNA-PKcs, a member of the phosphatidylinositol-3 kinase-like family of serine-threonine protein kinases (PIKKs), to form the DNA-PK complex which tethers the DNA ends together (Williamson et al., 2009). Upon synopsis of the two DNA ends, DNA-PKcs phosphorylates itself in trans leading to its dissociation from the DNA ends which facilitates the access of downstream proteins, to the ends of the DSB (Weterings, 2007). After that, polymerases μ and λ are involved in gap filling (Mahaney et al., 2009), and the ends are ligated by the XRCC4-DNA ligase IV complex (X4L4), which is stimulated by an interaction between XLF and XRCC4. Also, additional proteins may be required in a more complex subpathway of NHEJ including MRN complex, 53BP1, Artemis and the ATM kinase (reviewed in Valerie and Povirk, 2003). This subpathway may be necessary for a subpopulation of more-difficult-to-repair DSBs, such as DSBs with heavily damaged termini, DSBs in heterochromatin, or DSB whose ends have become physically separated.
Figure 1-1: Nonhomologous end joining. The Ku 70/80 heterodimer binds to DNA DSB ends and recruits DNA-PKcs. This is followed by the recruitment of XLF:XRCC4:DNA ligase IV to the DNA ends which in turn leads to the synapsis of the two ends. DNA-PK autophosphorylation causes a conformational change followed by its dissociation from the DNA ends. This facilitates the access of downstream endonucleases proteins to process the DNA ends before gap filling by pol µ and λ and ligation that mediated by DNA ligase IV. Adapted from (Povirk, Valerie 2003).
1.5 The Choice of DSBs Repair Method

Both DSB structure and chromatin complexities can affect the speed and choice of DSB repair pathway (Shibata et al., 2011). When the DSBs occur in euchromatin DNA (EC-DNA) without DNA complexity, DSBs would be efficiently repaired with faster kinetics by NHEJ pathway independent to the cell cycle phase. However, if DSBs occur in EC-DNA with high complexity, these DSBs would be repaired with slow kinetics in G1 and G2 phases by NHEJ, but NHEJ will stall in G2 phase allowing the DSBs to undergo resection to be repaired by HR. However, in case of chromatin complexity, the DSBs would be repaired by NHEJ with slow kinetics in G1 phase. Again, in G2 phase, NHEJ is inefficient and will stall allowing resection of the DSBs that would be repaired by HR. However, HR pathway can be switched to NHEJ in a heterochromatin DSBs if the CtIP dependent DSB end resection does not occur (Shibata et al., 2011).

1.6 The Choice of DSBs Repair Mechanism in Breast Cancer

Breast cancers can be divided into two types; the first type known as sporadic breast cancer develops after conception and accounts for 90-95% of breast cancer cases. The other type of breast cancers is familial breast cancer due to mutated genes inherited from one’s parents. It accounts for 5-10% of all breast cancers, and about 35 - 45% of familial breast cancer can be attributed to mutation in two different tumor suppressor genes known as BRCA1 and BRCA2 (Rosen et al., 2003; Miki et al., 1994; Wooster et al., 1995). Nevertheless, mutations in either of BRCA1 or BRCA2 occur occasionally in sporadic breast cancer (Futreal et al, 1994; Lancaster et al, 1996). In addition to the role of mutated BRCA1 and 2 in predisposing individuals to breast cancer, these two proteins play an important role in protecting genome stability by responding to
DNA damage. Particularly, these genes have been shown to be involved in DSBs repair and are mainly critical for efficient HR but appear to play only a minor role in NHEJ, although BRCA1 may influence NHEJ fidelity (Gudmundsdottir & Ashworth, 2006; Nagaraju & Scully, 2007; H. Wang et al., 2001; Zhuang et al., 2006). Many sporadic breast tumors show allelic loss (Johnson et al., 2002) and/or reduced expression (Birgisoittir et al, 2006; Jaspers et al., 2009; Yshikawa et al., 1999; Wilson et al., 1999) suggesting that those tumors would be more dependant on NHEJ. Conversely, breast tumor cells often show upregulation of receptors of the epidermal growth factor (EGFR) family (Johnson et al., 2006), which have been implicated in radioresistance (Contessa, Abell, Mikkelsen et al., 2006; Contessa, Abell, Valerie et al., 2006) and stimulation of NHEJ (Das et al., 2007). Therefore, targeting NHEJ may selectively increase radiosensitivity of many breast tumor cells.

1.7 XRCC4-XLF Interaction as a Target for Radiosensitization

The XRCC4 gene was isolated by complementation of radiosensitivity, V(D)J recombination deficiency and DSB repair deficiency of the Chinese Hamster ovary derivative XR-1 (Li et al., 1995). XRCC4 found to be associated with DNA ligase IV, and in vitro, the X4L4 complex is critical for broken DNA ends ligation through the NHEJ mechanism (Critchlow et al., 1997). Also, XRCC4 in cells was found to stabilize the DNA Ligase IV (Bryans et al., 1999). The ligation activity of X4L4 is stimulated by an interaction with XLF; the most recently discovered core protein of NHEJ (Ahnesorg et al., 2006; Buck et al., 2006). XLF was identified initially as missing gene in a subset of severe combined immune deficiency (SCID) patients with characteristics of growth retardation, microcephaly, immunodeficiency, increased cellular radiosensitivity and a defective V(D)J recombination
(Buck et al., 2006). Both proteins have no known enzymatic function, yet in mammalian cells, XRCC4<sup>−/−</sup> and XLF<sup>−/−</sup> genotype show increased sensitivity to ionizing radiation associated with severe defects in DSB repair (Giaccia et al., 1990; Zha et al., 2007).

Crystal structure studies of XRCC4 and XLF show structural similarity of both proteins in which both of them form homodimers and have similar N-terminal head domains and long α helical tails (Li et al., 2008; Andres et al., 2007; Junop et al., 2000). Mutagenesis studies of the head domains of XRCC4 and XLF found residues E55, D58, M61 and F106 of XRCC4, and R64, L65 and L115 of XLF as critical to XRCC4-XLF interaction (Ropars et al., 2011; Malivert et al., 2010). Based on the crystal structures of XRCC4 and XLF, and based on mutagenesis analysis, it has been suggested that these two proteins could form filaments of alternating XRCC4-XLF dimers which could twist a DSB and help to align the ends (Andres et al., 2007). Regardless of the accuracy of the predicted filament formation between XRCC4 and XLF, the interaction between these proteins is essential for NHEJ. Thus, this critical interaction should be susceptible to disruption by small peptides that bind this XLF-interacting region of XRCC4, and such disruption should severely suppress NHEJ.
**Figure 1-2: Crystal structures of XRCC4 and XLF.** Both XRCC4 and XLF have two head domains and long α helix (Andres et al., 2007)
Figure 1-3: Filaments of alternating XRCC4-XLF dimers. Proposed filaments that show the formation of alternating filaments between XLF (orange) and XRCC4 (blue). The binding of XRCC4 to Ligase IV (yellow) is also indicated (Andres et al., 2007).
Figure 1-4: Interface of the XRCC4\textsuperscript{1–157}-XLF\textsuperscript{1–224} complex formed by the distal part of their N-terminal head domains. Residues involved in the interface between XRCC44 (cyan) and XLF (green) are shown in a stick representation. Residues are labeled with “X” and “C” superscripts for X4 and Cernunnos (XLF), respectively. Dashed lines indicate inter- and intra-molecular hydrogen bonds and salt bridges between side-chain atoms proposed from Rosetta modeling. (Taken from Ropars et al., 2011)
1.8 mRNA Display

mRNA display is an *in vitro* selection technology that can be used to synthesize libraries of mRNA-peptide fusions containing up to 10 trillion unique sequences (Roberts & Szostak, 1997). The main idea of this technology is the covalent attachment of the peptide chain to the 3′ end of its own mRNA template. The process starts when the mRNA first becomes covalently attached at its 3′ end to a puromycin containing short DNA linker via a psoralen photo-crosslink. During an *in vitro* translation, when the ribosome reaches the crosslinked region and translation pauses, puromycin mimics the 3′ end of an incoming tRNA by entering the ribosome A-site and accepting the nascent peptide forming the covalent bond between the mRNA and the peptide it encodes. Using a library of mRNAs transcribed from a synthetic DNA template, this process will create large collections of mRNA-peptide fusions. The RNA portion of the mRNA-peptide fusions can be reverse-transcribed and PCR-amplified, which allows for the identification of the functional peptide by DNA sequencing. Several rounds of selection and amplification can be carried out in order to allow for an enrichment of unique sequences with the required properties.

mRNA display has many applications. For example, this method has been used to identify molecules that bound to TNF-α with a dissociation constant of 20 pM (Xu et al., 2002). Also, this technique was used to select molecules that bind to phosphorylated IkBα with an affinity of 18 nM (Olson et al., 2008). Moreover, mRNA display was helpful in selecting Bcl-xL inhibitor with an IC50 of 0.9 µM (Matsumura et al., 2010). In addition, an inhibitor with an IC50 of 2-5 nM has been identified by mRNA display to inhibit a protein-ligand interaction between VEGF and its receptor Flt-1 with 800 Å² of buried interfacial surface area (Getmanova et al., 2006; Wiesmann et al., 1997). Also, the mRNA display has been used to develop two unnatural and one natural peptide inhibitors of thrombin with $K_d$ in nanomolar range (Schilpp et al.,
2012). The unnatural peptide inhibitors have $K_d$ of 4.5 and 20 nM while the natural one has a $K_d$ of 1.5 nM.
II. METHODS

2.1 Production of Wild Type GST- XRCC4<sup>1</sup>-<sup>157</sup>, Double Mutant GST - XRCC4<sup>1</sup>-<sup>157</sup> (M61R and F106E) and Double Mutant Trx - XRCC4<sup>1</sup>-<sup>157</sup> Protein

2.1.1 Construction of Wild Type XRCC4<sup>1</sup>-<sup>157</sup>

Wild type truncated XRCC4<sup>1</sup>-<sup>157</sup> was amplified by PCR from the full-length XRCC4 sequence (366 amino acids) on pFLAG-CMV-2 (K. Jones, 2005) using FastStart PCR Master mix (Roche) and the primers 5ʹ-GGATCCATGGAGAAGCAGAATCCACCCAC-3ʹ and 5ʹ-CTCGAGTCAGTGATGATGATGATGATGATGATGTCATTTCAATCTCTCAGAAGCCTTC-3ʹ with insertion of BamHI restriction site at the N terminus and an XhoI site following the 6xHis tag at the C terminus, immediately following residue 157. The PCR reaction contained (5 ng pFLAG-CMV-2 - XRCC4 plasmid), 1X PCR master mix (2X: 50 units/mL of Taq DNA polymerase supplied in a proprietary reaction buffer (pH 8.5), 400µM dATP, 400µM dGTP, 400µM dCTP, 400µM dTTP, 3mM MgCl₂) and 0.5 µM of each primer. The PCR reaction carried out with an initial heating of 94°C for 5 minutes followed by 5 cycles of 94°C for 10 seconds (denature), 50°C for 20 seconds (annealing) and 72°C for 1 minute (extension) followed by 35 cycles of 94°C for 10 seconds (denature), 58°C for 20 seconds (annealing) and 72°C for 1 minute (extension) with a final extension of 72°C for 7 minutes. The XRCC4<sup>1</sup>-<sup>157</sup> PCR product amplified was mixed with sample buffer and resolved on a 1% agarose gel containing ethidium
bromide for visualization using UV. The XRCC4^{1-157} PCR product was cut from the gel and purified by QIAquick Gel Extraction Kit from QIAGEN.

### 2.1.2 Insertion of XRCC4^{1-157} into the pGEX-4T-1 Expression Plasmid Vector

First, the pGEX-4T-1 expression vector (GE Healthcare) for expression of N-terminal GST protein was digested with *Bam*HI for 2 hours at 37° C; followed by 2 hours digestion at 37° C with *Xho*I and 30 minutes at 37° C with antarctic phosphatase (New England BioLabs Inc.). After that, the digested pGEX-4T-1 was resolved on a 1% agarose gel and purified by QIAquick Gel Extraction Kit. Also, XRCC4^{1-157} fragment was digested with *Bam*HI and *Xho*I for 2 hours at 37° C and purified. The ligation reaction was performed using a 10-fold molar excess of XRCC4 insert DNA over vector. Approximately 50 ng of pGEX-4T-1 was used and the amount of insert DNA needed was 48 ng. XRCC4^{1-157} and pGEX-4T-1 were ligated together at 16° C overnight using T4 DNA ligase and its supplied buffer (Roche, Indianapolis, IN) in a final ligation volume of 10 µL. Fifty µL of *E. coli* competent (DH5 α) bacterial cells (Strategene) were transformed by heat shock with 5 µL out of 10 µL of the ligated DNA. The DNA and *E. coli* cells were mixed and placed on ice for 30 minutes, placed in a 42° C water bath for 45 seconds and placed on ice for 2 minutes. Following 2 minutes on ice, 1000 µL of LB broth was added followed by 1 hour incubation with agitation at 37° C. In order to grow a mix of colonies, following incubation, 10 µL, 100 µL, 200 µL of the transformation mix were spread on LB agar plates containing 100 µg/ml ampicillin and incubated at 37° C overnight. pGEX-4T-1 contains an ampicillin resistance gene. Therefore, *E. coli* colonies that contain pGEX-4T-1 should grow in the presence of ampicillin. Individual colonies that grew on agar plates were selected using a sterile toothpick and each colony was added to 5 mL of LB medium containing 100 µg/ml
ampicillin and incubated at 37°C overnight with gentle agitation. Plasmid DNA from individual colonies were isolated using mini-prep plasmid DNA isolation kit from QIAGEN. Confirmation that the plasmid DNA, isolated from the colonies, contained XRCC4\textsuperscript{1-157} and pGEX-4T-1 was achieved by digesting the purified colonies with BamHI and XhoI restriction enzymes and compared the size on a 1% agarose gel. Furthermore, confirmation was achieved by DNA sequencing.

2.1.3 Construction of Double Mutants XRCC4\textsuperscript{1-157} (M61R and F106E)

Double-mutants of XRCC4\textsuperscript{1-157} (M61R and F106E) were generated directly from the wild type XRCC4\textsuperscript{1-157}-pGEX-4T-1 plasmid using the QuickChange site-directed mutagenesis kit (Stratagene), and primers that determined by the QuickChange Primer Design Program (Stratagene). For M61R: 5’-CCAAGAAGCTGATGACATGGCA\textcolor{red}{AGG}AAAAAGGGAAAT-ATG-3’ and 5’-CATATT\textcolor{red}{TCC}CTTTTTCCTTGCCATGTCACTGCTTCTTGG-3’. For F106E: 5’-ATTTCTTCTTTGAGAAAAACCTGAAGATGTCTCA\textcolor{red}{GAG}AGACTTGGTTCC-TCAAC-3’ and 5’-GTTGAAGGAACCAAGTCTC\textcolor{red}{CT}TGAGACATCTTTCCAGGTTTTTCTCA-AAGAAGAAAT-3’. The mutated residues are highlighted in which the methionine (ATG) was mutated to arginine (AGG) and the phenylalanine (TTC) was mutated to glutamic acid (GAG). Then, the double-mutant plasmid DNA was transformed into E. coli. Transformation of E. coli cells with the double-mutant DNA was performed as previously described. Colony selection and verification was also performed as previously described.

2.1.4 Insertion of Double-Mutants XRCC4\textsuperscript{1-157} (M61R and F106E) Gene into pET-32XT Expression Vector

The double-mutant XRCC4\textsuperscript{1-157} (M61R and F106E) was also transformed into pET-32XT expression vector that has thioredoxin tag at its N-terminus. pET-32XT is a modified vector
from pET-32a in which the original thrombin cleavage site encoded by the pET-32a vector was removed by site-directed mutagenesis (Arg → Gln) (Cai et al., 2003). pET-32XT has BamHI and XhoI restriction sites. First, both double-mutant pGEX-4T-1- XRCC4\textsuperscript{1-157} plasmid and pET-32XT-Smad6 MH2 domain plasmid were digested with BamHI and XhoI restriction sites for 2 hours at 37°C. Then, XRCC4\textsuperscript{1-157} was ligated to pET-32XT as described (see 2.1.2). Transformation of E. coli cells, colony selection and verification was also performed as previously described.

### 2.1.5 Production of Wild-Type and Double-Mutant XRCC4\textsuperscript{1-157} Proteins

Wild type and mutant XRCC4\textsuperscript{1-157} plasmids were expressed in E. coli BL21 (DE3) competent cells and induced using an autoinduction protocol for protein production (F.W. Studier, 2005). First, single colonies of the plasmids harboring wild-type or mutant XRCC4\textsuperscript{1-157} were grown overnight at 37°C with agitation in 5 mL LB medium containing 100 μg/ml ampicillin. Then, 1 ml of the overnight cultures was transferred into four 2 L flasks containing 500 mL ZYM-5052 medium (each 500 mL culture contains 5 gm tryptone, 2.5 gm yeast, 10 mL of 50XM solution, 10 mL of 50X5052 solution, 500 μL of 1000X trace metals and 500 μL of 100X trace metals and 500 μL of 100 μg/ml ampicillin). The four cultures were inoculated at 37°C with agitation for 3-4 hours until visible turbidity can be seen. Then, the cultures were left to grow at 20°C overnight. After that, time course of optical density (600 nm) measurements were taken of the four cultures until they reached saturation. Then, the cells centrifuged at 4000 rpm and 4°C for 30 minutes, and the dry pellets were stored at – 20°C.
Stock solutions:

**50XM**

<table>
<thead>
<tr>
<th>Solution</th>
<th>Concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Na$_2$HPO$_4$</td>
<td>25 mM</td>
</tr>
<tr>
<td>KH$_2$PO$_4$</td>
<td>25 mM</td>
</tr>
<tr>
<td>NH$_4$Cl</td>
<td>50 mM</td>
</tr>
<tr>
<td>Na$_2$SO$_4$</td>
<td>5 mM</td>
</tr>
<tr>
<td>50X5052</td>
<td></td>
</tr>
<tr>
<td>Glycerol (w/v)</td>
<td>0.5%</td>
</tr>
<tr>
<td>Glucose</td>
<td>0.05%</td>
</tr>
<tr>
<td>α-lactose</td>
<td>0.2%</td>
</tr>
</tbody>
</table>

**1000X Trace metals**

<table>
<thead>
<tr>
<th>Solution</th>
<th>Concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>FeCl$_3$</td>
<td>50 mM</td>
</tr>
<tr>
<td>CaCl$_2$</td>
<td>20 mM</td>
</tr>
<tr>
<td>MnCl$_2$</td>
<td>10 mM</td>
</tr>
<tr>
<td>ZnSO$_4$</td>
<td>10 mM</td>
</tr>
<tr>
<td>CoCl$_2$</td>
<td>2 mM</td>
</tr>
<tr>
<td>CuCl$_2$</td>
<td>2 mM</td>
</tr>
<tr>
<td>NiCl$_2$</td>
<td>2 mM</td>
</tr>
<tr>
<td>Na$_2$MoO$_4$</td>
<td>2 mM</td>
</tr>
<tr>
<td>Na$_2$SeO$_3$</td>
<td>2 mM</td>
</tr>
<tr>
<td>H$_3$BO$_3$</td>
<td>2 mM in 60 mM HCl</td>
</tr>
</tbody>
</table>

**500X MgSO$_4$**

2 mM MgSO$_4$

**2.1.6 Ni-NTA Purification**

Cell pellets were left to thaw in ice and the pellet in each bottle was resuspended in 5 ml of freshly prepared Basic Buffer N (50 nM phosphate pH 8, 10% glycerol, 5 mM fresh 2-mercaptoethanol (BME), 20 mM imidazole and 0.5 M NaCl) and 150 µL His compatible protease inhibitor (Merck, EMD). After that, the cells were transferred into a 50 mL centrifugal tube and sonicated 4X 45 seconds on 30% power (Sonicator from Ultrasonic Inc.). The lysed cells were diluted with 20 mL Basic Buffer N and centrifuged for 15 minutes at 15000 rpm at 4°C. Then, the supernatants were collected and mixed with 4 mL Ni-NTA resin (QIAGEN) and gently stirred for 1 hour at 4°C. After that, the flow through was collected and the resin washed with 20 mL of Basic Buffer N. The protein was eluted 6 times and each time was with 2 mL of Ni-NTA elution buffer (50 mM Na$_2$HPO$_4$, 300 mM NaCl and 250 mM imidazole; pH 8). All six fractions were analyzed by SDS-PAGE electrophoresis. Then, the elutions were combined and dialyzed in a 3-12 mL volume Slide-A-Lyzer dialysis cassette for 4-6 hours at 4°C against 1
2.1.7 Thrombin Cleavage of GST- XRCC4\textsuperscript{157} DM Protein

The cleavage of the GST moiety from the GST-XRCC4\textsuperscript{157} DM was carried out using thrombin (GE Healthcare). The concentration of the thrombin is 1 U/\textmu L and each unit can cleave > 90% of 100 \textmu g of fusion protein in 16 hours at 22°C. The protein was dialyzed overnight with 1X PBS buffer (140 mM NaCl, 2.7 mM KCl, 10 mM Na\textsubscript{2}HPO\textsubscript{4}, 1.8mM KH\textsubscript{2}PO\textsubscript{4}; pH 7.4). The dialyzed protein was collected and the concentration was measured. Then, one mg of protein was incubated with 1 or 2 units of thrombin. The reaction was mixed gently at room temperature and 10 \textmu L was taken after 2, 4, 16 and 18 hours. The protease activity was stopped with 1 \textmu L of His-compatible protease inhibitor cocktail (Calbiochem) and saved at -20°C until use. The results were checked on 12% SDS-PAGE.

2.2 mRNA Display and In Vitro Selection

2.2.1 DNA library Synthesis

Five different DNA libraries have been synthesized (by Dr. Hartman) and used in the selection process. Briefly, each library encodes a fixed cysteine at the N-terminus and another cysteine after 2, 4, 6, 8, or 10 random amino acids in order to produce different cyclic sizes of peptides. The DNA sequence of the random library is:
The detailed sequence composition of each library is as follows:

1- **TAATACGACTCCTAGGTTAACTTTACTAAAGGAGGACAGCTAAATGTGCNNSNN**
   **SNNSNNSNNSNNSNNSNNSNNSNNSNNSNNSNNSNNSNNSN**
   **SNNSGTCTAGCGGCTCCTTAGGCCCACCAT**
   **CACCATCACCACCGGTAGCTAGTTAGCTAG**

   The detailed sequence composition of each library is as follows:

   1- **TAATACGACTCCTATA**: T7 promoter, followed by GGG with the first “G” as the transcription start.

   2- **TTAAGGTGTTAG**: Epsilon enhancer.

   3- **TAAGGGAGG**: Shine-Dalgarno sequence, also known as the ribosome binding site.

   4- **ACAGCTAA**: the Spacer between ribosome binding site and the start codon, with “AA” at end.

   5- **ATGTGCNNS’sTGCNNS’s**: translation start codon for methionine (ATG) and the codon for the fixed cysteine (TGC), followed by a mix of random NNS codons and another fixed cysteine at the site indicated above. N denotes A, T, C, or G, and S denotes G, or C. The translated peptides have the sequence of MCX2CX10GSGSLGHis6, MCX4CX8GSGSLGHis6, MCX6CX6GSGSLGHis6, MCX8CX4GSGSLGHis6, or MCX10CX2GSGSLGHis6, where X can be one of the natural amino acids.

   6- **GGCTCCGCTAGCTAGGC**: codons for GlySerGlySer-LeuGly, the flexible linker with two out-of-frame stop codons.

   7- **CACCATCACCACCTAC**: codons for His5 tag. The sixth His of His6 tag is in the following sequence.

   8- **CACCGGGCTAT**: hybridization region for the Psoralen crosslinker (XL-PSO oligonucleotide, see below). The CAC encodes the sixth His of the intact His6 tag.

   9- **AGGTAGCTAG**: 3’-UTR to allow those non-crosslinked peptides to release at the in-frame
TAG stop codons.

2.2.2 PCR Amplification of Library DNA

The DNA libraries were synthesized in reverse orientation. For example, CTAGCTACCTATA-GCCGTTGGTATGGTGATGGTGCCCTAAGGAGCCGCTACCSNNSNNSNNSNNNSNNSN NSNNSNNSNNSNNSNNSNNSNNGCACATTTTAGCTGCTCCTCCTTAGTAAAGTTAACCCT ATAGTGAGTCGTATTA. Then, the DNA libraries were amplified by PCR using the following primers 5’-TAATACGACTCACTATAGGTTAACTTTACTAAGGAGGACAG-3’ and the 5’-CTAGCT-ACCTATAGCCGTTGGTGATGGTGATGGTGCTAAGGAGCCGCTACC-3’. Each PCR reaction contained 12 nM of DNA library, 0.5 µM of each primer, 1X ThermoPol buffer, 100 µM of dNTP’s and 5 U of Taq polymerase. The PCR reaction started by an initial heating of 94°C for 2 minutes followed by 18 cycles of 94°C for 15 seconds, 60°C for 30 seconds, and 72°C for 45 seconds followed by 72°C for 5 minutes. PCR product amplified was mixed with sample buffer and resolved on a 2% agarose gel containing ethidium bromide for visualization using UV. PCR product was cut from the gel and gel purified by QIAquick Gel Extraction Kit from QIAGEN.

2.2.3 StrataClone Cloning and Sequencing

The PCR product was ligated to StrataClone vector (Stratagene). The ligation product was transformed into StrataClone SoloPack competent cells following the StrataClone PCR Cloning Kit from Stratagene. After that, the transformation product was plated onto LB plates containing 20 mg/ml kanamycin and the plates were incubated overnight at 37°C. Then, 96 single colonies were selected and sent for sequencing to ensure that the libraries were
synthesized as designed.

2.2.4 Transcription and Purification of mRNA

Each library was transcribed by setting up a 1 ml transcription reaction that contains 0.1 µM DNA library, 0.6 µM T7 primer, 40 mM/0.1% Tris/Triton solution, 2.5 mM spermidine, 25 mM MgCl₂, 10 mM DTT, 5 mM CTP, 5 mM UTP, 5 mM ATP, 9 mM GTP, 0.2 U/µl RNase inhibitor, 1 U/µl inorganic pyrophosphatase and 1 U/µl T7 RNA polymerase (MA & Hartman, 2012). Then, the reaction was incubated overnight at 37°C in an incubator chamber. After that, the transcription reaction was removed from the incubator and 50 µL of Turbo DNase (Invitrogen) was added and the reaction was incubated for 15 minutes at 37°C. After incubation, 7.5 M urea was added and mixed well and the reaction was loaded on Urea PAGE by setting up a large 65 SequaGel (20 cm X 20 cm) on an Owl P10DS Dual Gel System. The gel was pre-run for 20–30 min by using constant power at 25 W supplied by EC Apparatus electrophoresis power supply EC 600. Then, the urea was flushed out from well and the sample was loaded. The gel was run at constant power of 25 W for 60-90 minutes until the bromophenol blue reaches the bottom of the gel. After that, the gel was transferred from the glass plates to Saran wrap on both sides and the mRNA band was visualized by UV shadowing and the band cut out with a fresh razor blade.

After that, the mRNA was eluted by using the Whatman Elutrap electroelution System. First, the gel slice was put into the chamber, and mashed into small pieces and the chamber filled with 0.5X TBE buffer until the gel is covered. The electroelution was run for 2 h at 300 V on a Bio-Rad PowerPac basic power supply. At the end of running, the electrodes switched and run backwards for 1 min at 300 V. Then, the solution between membranes was collected and ethanol
precipitated with 0.1 volume of 3 M KOAc and 3 volumes of 100% ethanol. The sample was mixed well and frozen at −20°C for 30 minutes. After that, the sample was centrifuged at 16,000 Xg for 20 minutes at 4°C. The supernatant was discarded and the pellet washed with 500 µL of 70% ethanol and centrifuged again at 16,000 xg for 1 min. The supernatant was discarded and the pellet air-dried for 5–10 min at room temperature. The mRNA was dissolved in ddH₂O and an absorbance at 260 nm was measured on a spectrophotometer. The concentration was calculated by using the online software Oligonucleotide Properties Calculator (http://www.unc.edu/~cail/biotool/oligo/index.html). Then, the mRNA was stored at -20°C.

2.2.5 Psoralen Photo-Crosslinking

The five mRNA libraries were mixed together with a final concentration of 2 µM in order to be photo-crosslinked with an XL-PSO oligonucleotide. The sequence of the XL-PSO oligonucleotide is 5’-PsoC6-(UAGCCGGUG)2’-OMe-15xA-2xSpacer9-ACC-Puro-3’. Spacer9 is triethyleneglycol (TEG) phosphoramidites and it is used to tether 5’-dCdC-puromycin to the 3’-end. The Psoralen photo-crosslinking reaction contains 20 mM HEPES-KOH, pH 7.6, 7.5 µM XL oligo, 100 mM KCl, 1 mM spermidine, 1 mM EDTA and the mixture of the mRNA libraries. The reaction was mixed well and transferred into PCR tubes with 100 µl in each tube. The PCR tubes were placed in a PCR machine, heated to 70°C for 5 min, then cooled to 25°C over 5 min (0.1°C/s). After this, each sample was transferred to a crosslink plate, 100 µL per well. A 365 nm UV lamp was used to irradiate the plate for 20 minutes at 4°C. The samples were collected in 1.5 ml eppendorf tube and ethanol precipitated as previously described. The pellet was resuspended in the required volume.
2.2.6 *In vitro* translation

A 5 ml translation reaction was prepared, which contained 8 mM putrescine, 1mM spermidine, 5 mM potassium phosphate, 95 mM potassium chloride, 5 mM ammonium chloride, 5 mM magnesium acetate, 0.5 mM calcium chloride, 1mM dithiothreitol, 1 µg/ml inorganic pyrophosphatase, 4 µg/ml creatine kinase, 1.1 µg/mL nucleotide diphosphate kinase, 30 µM (6R,S)-5,10-formyl-5,6,7,8-tetrahydrofolic acid, 93 µg/mL myokinase, 10 mM creatine phosphate, 2 mM ATP, 2mM GTP, 2.4 mg/ml *E. coli* total tRNA, 0.2 µM MTF, 1.0 µM IF1, 0.3 µM IF2, 0.7 µM IF3, 3.2 µM EF-Tu, 0.6 µM EF-Ts, 0.5 µM EF-G, 0.3 µM RF1, 0.4 µM RF3, 0.1 µM RRF and 0.5 µM ribosomes. In addition, the reaction contained 18 natural amino acids (200 µM each), 10 µM cysteine, 10 µM methionine and 20 aminoacyl tRNA synthetases (0.1 µM MetRS, 0.3 µM LeuRS, 0.6 µM GluRS, 0.2 µM ProRS, 1.0 µM GlnRS, 1.0 µM HisRS, 0.25 µM PheRS A294G, 1.5 µM TrpRS, 0.2 µM SerRS, 0.2 µM IleRS, 0.4 µM ThrRS, 0.6 µM AsnRS, 0.6 µM AspRS, 0.5 µM TyrRS, 0.5 µM LysRS, 0.4 µM ArgRS, 0.2 µM ValRS, 0.2 µM AlaRS, 0.5 µM CysRS, and 0.06 µM GlyRS). 0.3 µM of 35S-Met is added to isotopically label the peptides. The translation mix was incubated for 15 minutes at 37°C. Then, the 2 µM photocrosslinked mRNA was added to the translation mix and incubated for 1 hour at 37°C. After that, the translation reaction was quenched with 550 mM KCl and 50 mM MgCl₂ followed by an incubation for 90 minutes at 37°C. Then, the translation reaction was chilled overnight at -80°C.

2.2.7 Oligo(dT) Purification and Cyclization

The goal of oligo(dT) purification and cyclization is to remove all non photo-crosslinked mRNA-peptide fusions and to produce cyclic peptides at the same time (Figure 2-1). The translation reaction was thawed and vortexed and 5 µL removed for scintillation counting to
determine the total radioactivity of $^{35}$S-Met added to the translation reaction. Then, five 100 mg Oligo(dT)-cellulose Type 7 (GE Healthcare) were added to five 20 mL Bio-Rad Econo-Pac columns (BioRad) and rinsed once with 5 ml ddH$_2$O to swell cellulose and twice with oligo(dT) binding buffer (20 mM Tris–HCl, pH 7.8, 10 mM EDTA, 1 M NaCl, 0.2% Triton X-100. Add 0.5 mM fresh TCEP). Then, each one ml of the translation reaction was added to one column and the reaction was brought up to 4 mL with the binding buffer. The columns were placed on a rocking platform shaker in 4°C cold room and shook for 30 min. After that, the flow through was removed and the reaction was rinsed twice with the oligo(dT) binding buffer. Then, 4 mL of cyclization buffer (20 mM Tris–HCl, pH 7.8, 0.66 M NaCl, 3 mM α,α’-dibromo- m -xylene (Sigma-Aldrich/Fluka), 33% acetonitrile (v/v), 0.5 mM fresh TCEP) was added to each reaction and rocked for 30 minutes at room temperature. Then, the flow through was removed and each column was washed twice with 4.5 mL oligo(dT) wash buffer (20 mM Tris–HCl, pH 7.8, 0.3 M NaCl, 0.1% Triton X-100); first wash was with 5 mM fresh BME and the other wash was with 0.5 mM fresh TCEP. Then, the columns were eluted 8 times with ddH$_2$O and filtered to remove out residual Oligo(dT)-cellulose in the fusion solution. One µL of each eluent was counted in a scintillation counter and the fractions with significant radioactivity were combined and ethanol precipitated using 0.3 volume of 3 M KOAc, pH 5.2, 0.002 volume of glycogen (5mg/mL) (Applied Biosystems), and 3 volume of 100% ethanol. Then, the pmole amount of the mRNA-peptide fusion was calculated based on the scintillation counts and the pellet re-dissolved in ddH$_2$O so that the final concentration of the fusion is 100 nM (0.1 pmole/µL).

2.2.8 Reverse Transcription

The goal of reverse transcription (RT) is to convert the single-stranded mRNA portions
of the fusions to heteroduplex of RNA/DNA in order to eliminate any unwanted RNA secondary
structures and render the nucleic acid portion of the fusion more stable. The RT reaction was
carried out following the Superscript III First Strand Synthesis Kit (Invitrogen). In a
microcentrifuge tube, the 100 nM mRNA-peptide fusions was added to 0.5 mM dNTPs and 0.5
µM RT-primer TTTTTTTTTTTTTGTGATGGTGATGGTGGCCTAAGC. The tube was
incubated for 5 minutes at 65 °C in a heating block. After five minutes, the tube was immediately
placed on ice and incubated for at least 1 minute. Then, the following reagents were added in a
final concentration of 5 mM MgCl₂, 1 mM DTT, 2 U/µL RNaseOUT (Invitrogen), 5 U/µL
Superscript III (Invitrogen). The RT reaction was incubated for 15 minutes at 55°C for
elongation, then 15 minutes at 70°C to inactivate the Superscript III. At the end of the RT
reaction, 0.5 µL of each RT reaction was counted in a scintillation counter in order to calculate
how much of mRNA-peptide fusions was recovered.

2.2.9 Ni-NTA Purification

Ni-NTA purification was performed in order to remove any truncated peptide fusions
(Figure 2-2). The following buffers were prepared for Ni-NTA purification under native
condition:

<table>
<thead>
<tr>
<th>Denaturing Binding buffer, pH 8</th>
<th>Wash buffer, pH 8</th>
<th>Elution buffer, pH 8</th>
</tr>
</thead>
<tbody>
<tr>
<td>100 mM NaH₂PO₄</td>
<td>100 mM NaH₂PO₄</td>
<td>50 mM NaH₂PO₄</td>
</tr>
<tr>
<td>10 mM Tris HCl</td>
<td>300mM NaCl</td>
<td>300mM NaCl</td>
</tr>
<tr>
<td>6 M guanidinium HCL</td>
<td>0.2% Triton X-100</td>
<td>250mM imidazole</td>
</tr>
<tr>
<td>0.2% Triton X-100</td>
<td>5 mM BME (fresh)</td>
<td>0.2% Triton X-100</td>
</tr>
<tr>
<td>5 mM BME (fresh)</td>
<td></td>
<td>5 mM BME (fresh)</td>
</tr>
</tbody>
</table>

First, 100 µl of Ni-NTA agarose was added into spin filter tube. Then, the RT reaction
was diluted 5 fold with the Ni-NTA denaturing binding buffer and transferred into filter tube.
The peptide fusion was bound to the Ni-NTA for one hour at 4°C in a tumbler in a cold room. After that, the flow through was removed by centrifuging at 5900 rpm for 1 minute. The resin was washed 3 times with Ni-NTA wash buffer and centrifuged at each time. Then, the peptide fusion was eluted 6 times with portions of 50 µL Ni-NTA elution buffer and 0.5 µL of each eluent was counted in a scintillation counter to measure how much of mRNA-peptide fusions recovered. All eluents with high counts were combined and dialyzed against 1 L pre-cooled selection buffer (50 mM Tris–HCl, pH 8.0, 150 mM NaCl, 4 mM MgCl₂, 0.25% Triton X-100, 0.1 mg/ml BSA). The overnight-dialyzed peptide fusions were collected and 1 µL was counted in a scintillation counter.

2.2.10 In Vitro Selection

The selection process started by an overnight incubation of two proteins with Pierce Magnetic Glutathione beads (Thermo Fisher Scientific). The first sample contains GST protein only or GST-XRCC4157 DM for negative selection (pre-clearing) (Figure 2-3 and Figure 2-4) and the second sample contains GST-XRCC4157 fusion for selection (Figure 2-5). 50 µL of Pierce Magnetic Glutathione beads per 1 mL translation reaction was transferred into two 1.5 mL eppendorf tubes and washed three times with 400 µL GST beads wash buffer (125 mM Tris–HCl, pH 8.0, 150 mM NaCl); each time the magnetic beads were mixed gently and the supernatant was removed by using Magnetic separation stand (Invitrogen). 130 µg of each protein was mixed with 400 µL GST beads wash buffer and transferred to the tubes with magnetic beads and incubated at 4°C for overnight with rotation on a tube rotator. In the next day, the magnetic beads with GST protein were pelleted briefly and the supernatant was removed using the magnetic stand. The beads were washed twice with 400 µL GST beads wash buffer.
and once with 400 µL selection buffer. Then, the dialyzed mRNA-peptide fusions from the previous step were added to the tube at a final concentration of 0.1 mg/mL BSA was added. The tube was incubated at 4°C for 2 hours with rotation. After that, the supernatant was collected and the beads washed three times with 200 µL selection buffer. The beads from this step and 10 µL from the supernatant and washes were counted in a scintillation counter. Then, the remaining of the supernatant and the three washes were transferred to the other tube that contains the GST-XRCC4157 fusion. The tube was incubated at 4°C for 2 hours with rotation. Then, the supernatant was removed and the beads were washed three times with 400 µL of the selection buffer. Then, the beads were suspended in 100 µL of the selection buffer and 5 µL was counted in a scintillation counter. After that, the mRNA-peptide fusions that attached to the GST-XRCC4157 were amplified by PCR directly from the beads as described bellow.

2.2.11 PCR Amplification of Selected Fusions

The PCR reaction was carried out directly from the beads. First, a test of 100 µL PCR reaction was done with 5 µL of the suspended beads in order to find the optimum conditions for PCR. The PCR reaction was performed as previously described with the following primes 5ʹ- TAATACGACTCACTATAGGGTTAACTTTACTAAGGAGGACAGCTAAATG-3ʹ and the 5ʹ- CTAGCTACCTATAGCCGGTGGTGATGGTGGTGATGGTT-GGCCTAAGGAGCCGCTACC-3ʹ.

Starting in cycle 14, 5 µL was taken out of the PCR tube after the elongation step (step 4, at 72°C). The PCR machine was paused at 44 seconds of this step and 5 µL was taken from the PCR reaction. This was done for even numbered PCR cycles 14-32. The PCR product was checked on 2% agarose gel and the intensity of DNA bands was analyzed to find out the optimum conditions of the PCR. Then, large-scale PCR was set up using multiple PCR tubes
and all suspended beads were used to amplify the selected fusions. After PCR, the reaction mixtures were combined and an equal volume phenol:chloroform:isoamyl alcohol (25:24:1) (Sigma-Aldrich) was added. The sample was vortexed and centrifuged at 4000 rpm for 2 minutes in a cold room. The upper layer was transferred to a new tube and an equal volume of CHCl₃ was added to the tube and span down again. The aqueous layer was transferred to new tube and ethanol precipitated as previously described. The resulting precipitate was dissolved into a final volume of 100 µL ddH₂O, and passed over a NAP-5 column (GE Healthcare) to de-salt. The final volume from NAP-5 was 500 µL. This was the template for the next round of selection.
Figure 2-1: *In vitro* transcription, photo cross-linking and *in vitro* translation
Figure 2-2: mRNA-Peptide Fusions Purification Steps.
Figure 2-3: Pre-Clearing (Negative Selection) with GST Protein Only. We will do pre-clearing with GST protein only in the first three rounds.
Figure 2-4: Pre-Clearing (Negative Selection) with GST-XRCC4^{157} Protein. We will use GST-XRCC4^{157} DM for pre-clearing in advanced rounds. This will increase the stringency.
Figure 2-5: *In vitro Selection*. The mRNA-peptide fusions that bound to the XLF interface of XRCC4 will be amplified directly from the beads in the first four rounds. However, in advanced rounds, the selective inhibitors will be eluted with adding excess of Trx-XRCC4^{157} DM.
III. RESULTS

3.1 Cloning of Wild Type XRCC4\textsuperscript{157} into pGEX-4T-1 Vector and Production of WT GST-XRCC4\textsuperscript{157} Protein

As described in the introduction, the main aim of this study is to find peptide inhibitors of the XRCC4-XLF interaction in order to increase radiosensitivity of breast cancer cells. Therefore, we first cloned the WT and DM XRCC4 into expression vectors and produced GST and Thioredoxin tagged XRCC4 to be used in the selection of such peptides from a library.

Previous studies showed that the interaction between XRCC4 and XLF occurs through the head domains (Ropars et al., 2011; Malivert et al., 2010). Therefore, a truncated XRCC4\textsuperscript{157} gene was produced from a full length XRCC4\textsuperscript{336} (Figure 3-1). The process started by PCR amplification of a XRCC4\textsuperscript{157} fragment from the pFLAG-CMV-2 - XRCC4\textsuperscript{336} plasmid. The XRCC4\textsuperscript{157} fragment was separated on a 1% agarose gel (Figure 3-2). The band corresponding to our XRCC4\textsuperscript{157} fragment was on 504 bp (471 bp for the 157 residues of XRCC4 in addition to 33 bp of BamHI and XhoI restriction sites, stop codon site and 6xHis that we added on our fragment). Therefore, the XRCC4\textsuperscript{157} fragment (504 bp) was purified from the gel using QIAquick Gel Extraction Kit and cut with BamHI and XhoI restriction enzymes to prepare for ligation into the pGEX-4T-1 vector. The pGEX-4T-1 is an expression vector that contains a GST tag at its N-terminus in addition to a multiple cloning site where XRCC4\textsuperscript{157} gene can be cloned. The purpose of cloning XRCC4\textsuperscript{157} into this vector is to be able to immobilize the protein on magnetic beads during the selection process. Therefore, pGEX-4T-1 vector was first cut with
BamHI and XhoI restriction enzymes following by a cut with antarctic phosphatase that catalyzes the removal of 5’ phosphate of the vector in order to prevent self-ligation. The cut vector was resolved on 1% agarose gel in order to confirm that the cutting has been successful. Then, the linearized vector of length 4900 bp was cut out and purified using QIAquick Gel Extraction Kit.

After that, the purified XRCC4157-BamHI/XhoI fragment was ligated into the pGEX-4T-1-BamHI/XhoI vector. The ligation reaction was transformed into competent E. coli DH5α cells as described in Methods and plated onto medium with ampicillin. Single colonies were picked from the plates and were grown over night. Plasmid DNA was purified using a Miniprep Kit. To test for plasmid DNA with the correct insert, purified plasmid DNA was digested with BamHI and XhoI. The resulting DNA fragments were separated on a 1% agarose gel (Figure 3-3) to confirm whether the cloning of XRCC4157 into pGEX-4T-1 vector had been successful. It was expected to see two bands, one band on about 500 bp and one band on about 5000 bp. The results show that ten out of the twelve minipreps (lanes 1-5, 7, and 9-12) have bands on the right size, 504 bp and 4900 bp. Lane 6 and 8 probably contain a religated vector with no insert. Plasmid DNAs (from minipreps 2, 4, 9 and 11) were sent for sequencing and confirmed as correct (Figure 3-4).

After that, pGEX-4T-1 - XRCC4157 plasmid was transformed in E. coli BL21 for protein expression to produce GST tagged XRCC4157. The protein was produced and purified as described in Methods. Then, fractions of the lysate, flow through, washes, and eluted protein were run on a 10% SDS-PAGE stained with coomassie blue. A scanned image of the coomassie stained SDS-PAGE gel is shown in Figure 3-5. The result from the gel analysis shows a major band corresponding to the molecular size of the GST- XRCC4157 protein, which is around 44 kD. The concentration of the protein was measured by a spectrophotometer (A280) and calculated to
be 2.58 mg/mL (57.3 µM). The total amount of the purified fusion protein obtained from 1 L culture was 7.74 mg.
Figure 3-1: The strategy of cloning the WT XRCC4157 gene into the pGEX-4T-1 vector and production of GST tagged XRCC4157 protein.
Figure 3-2: The XRCC4\textsuperscript{157} fragment and pGEX-4T-1 Vector Cut. Lane 1: pGEX-4T-1 cut with \textit{BamHI} and \textit{XhoI} to form single band around 4900 bp. Lane 2: XRCC4\textsuperscript{157} fragment amplified by PCR from pFLAG-CMV2-XRCC4\textsuperscript{336} plasmid. The lower band on about 504 bp indicates the XRCC4\textsuperscript{157} constructed fragment (XRCC4 = 471 bp, \textit{BamHI} = 6 bp, \textit{XhoI} = 6 bp, stop codon = 3 bp and 6xHis = 18 bp). The pGEX-4T-1 vector and XRCC4\textsuperscript{157} fragment were cut out from the gel and purified.
Figure 3-3: Cloning of Wild Type XRCC4\textsuperscript{157} into pGEX-4T-1 Vector. The WT XRCC4\textsuperscript{157} fragment was cloned into pGEX-4T-1 vector. Then, the XRCC4\textsuperscript{157} was cut out from the pGEX-4T-1 vector with \textit{BamHI} and \textit{XhoI} restriction enzymes. Two bands were formed; the upper band is around 4900 bp indicating the pGEX-4T-1 vector and the lower band is around 504 bp indicating the XRCC4\textsuperscript{157} insert. The results show that the samples in lanes 1-5, 7, and 9-12 have an insert on the expected size.
**Figure 3-4: XRCC4**\textsuperscript{157} **sequence.** Nucleotide 1-6 is \textit{Bam}H1 sequence. 7-477 is the XRCC\textsuperscript{157} sequence. 479-495 is the 6xHis sequence. 496-498 is the stop codon sequence. 499-504 is the \textit{Xho}l sequence.
Figure 3-5: Protein Composition of WT GST-XRCC4\textsuperscript{157} as determined by SDS-PAGE. 10\% polyacrylamide gradient gel stained with coomassie blue. Lane 1: 30 μL of lysate was loaded (out of 5 mL). Lane 2: 30 μL of flow through was loaded (out of 5 mL). Lane 3: 30 μL of washes (out of 20 mL). Lane 4-9: 30 μL of different protein elutions (each elution was 1 mL). The results show that the samples in lanes 4-9 have the WT GST-XRCC4\textsuperscript{157} protein on the expected size (44 kD).
3.2 Cloning of Double-mutant XRCC4\textsuperscript{157} (M61R and F106E) into pGEX-4T-1 Vector and Production of DM GST-XRCC4\textsuperscript{157} Protein

The purpose of producing a double-mutant XRCC4\textsuperscript{157} comprising M61R and F106E is to use it in the pre-clearing process during the selection. This is to eliminate any non-specific peptide inhibitors that might bind to regions other than that of the XLF interface of XRCC4. The two mutations at M61R and F106E were introduced directly into the WT pGEX-4T-1-XRCC4\textsuperscript{157} plasmid by constructing mutagenic primers and following the QuickChange site-directed mutagenesis procedure (Figure 3-6). Then, plasmid DNA was transformed into competent \textit{E. coli} DH5α cells as described in Methods and plated onto medium with ampicilllin. Single colonies were picked from the plates and were grown overnight. Plasmid DNA was purified using QIAGEN Miniprep Kit. Plasmid DNA’s (from minipreps 1-8) were sent for sequencing and confirmed positive (Figure 3-7).

After that, the DM pGEX-4T-1 - XRCC4\textsuperscript{157} plasmid was transformed into \textit{E. coli} BL21 for protein expression to produce GST tagged DM XRCC4\textsuperscript{157}. The protein was produced and purified as described in Methods. Then, fractions of the lysate, flow through, washes, and eluted protein were run on a 10% SDS-PAGE stained with coomassie blue. Ascanned image of the coomassie stained SDS-PAGE gel is shown in Figure 3-8. The result from the gel analysis shows a band corresponding to the molecular size of the GST-XRCC4\textsuperscript{157} DM protein, which is around 44 kD. The concentration of the protein was measured by a spectrophotometer (A280) and calculated to be 5.96 mg/mL (132 μM). The total amount of the purified fusion protein obtained from 1 L culture was 17.88 mg.
Figure 3-6: The strategy of cloning the DM XRCC4\textsuperscript{157} gene into the pGEX-4T-1 vector and production of GST tagged XRCC4\textsuperscript{157} protein.
Figure 3-7: Double-mutant XRCC4^{157} sequence. Red circles indicate the sites of the two mutations inserted in which the Methionine (ATG) was mutated to Arginine (AGG) and the Phenylalanine (TTC) was mutated to Glutamic acid (GAG).
Figure 3-8: Protein Composition of DM GST-XRCC4^{157} as determined by SDS-PAGE. 10% polyacrylamide gradient gel stained with coomassie blue. Lanes 1-5: 30 µL of different protein elutions (each elution was 1.5 mL). Lane 6: 30 µL of washes (out of 20 mL). Lane 7: 30 µL of flow through was loaded (out of 5 mL). Lane 8: 30 µL of lysate was loaded (out of 5 mL). The results show that the samples in lanes 1-5 have the DM GST-XRCC4^{157} protein on the expected size (44 kD).
3.3 Thrombin Cleavage of GST- XRCC4$^{157}$ DM Protein

We have tried to remove the GST moiety from the double-mutant GST tagged XRCC4$^{157}$ protein. The idea is to use a non-GST tagged XRCC4$^{157}$ DM during the elution step in the selection process. Once the mRNA-peptide fusions are mixed with the WT GST-XRCC4$^{157}$ and captured on magnetic glutathione beads, the reaction will be washed with the DM XRCC4$^{157}$ in order to remove all non-specific mRNA-peptide inhibitors. This is another way to increase the stringency of the selection system.

We used thrombin (GE Healthcare) to cleave the GST from double mutant XRCC4$^{157}$ protein. Thrombin cleaves Arg-Gly bonds in specific sequences (LeuVal-Pro-Arg$^{1}$Gly-Ser). After incubating the thrombin with DM GST-XRCC4$^{157}$ protein, different samples were removed at various time points and analyzed by 12% SDS-PAGE to estimate the yield, purity and extent of thrombin digestion. The results show that the GST moiety was not completely removed (Figure 3-9). We have tried many times to cleave the GST with different conditions but still large amount of protein could not be cleaved. Therefore, we cloned the DM XRCC4$^{157}$ into pET-32XT expression vector (see below).
Figure 3-9: Thrombin cleavage of GST-XRCC4<sup>157</sup> DM as determined by SDS-PAGE. 12% polyacrylamide gradient gel stained with coomassie blue. Lane 1: 25 µL (10 µg) of uncleaved GST-XRCC4<sup>157</sup> before dialysis. Lane 2: 25 µL (10 µg) of uncleaved GST-XRCC4<sup>157</sup> after an overnight dialysis with 1 L of 1X PBS buffer. Lane 3-6: 10 µL (out of 230 µL total reaction) of the cleaved protein (1 mg) with 1 unit thrombin after 2, 4, 16 and 18 hours. Lane 7-9: 10 µL (out of 230 µL total reaction) of the cleaved protein with 2 units thrombin after 2, 4 and 18 hours. The GST molecular weight is about 26 kD and the XRCC4<sup>157</sup> molecular weight is about 18 kD.
3.4 Cloning of Double-mutant XRCC4\textsuperscript{157} (M61R and F106E) into pET-32XT Vector and Production of DM Thioredoxin (Trx) - XRCC4\textsuperscript{157} Protein

We were not able to remove the GST tag from the DM GST-XRCC4\textsuperscript{157}. Therefore, we cloned the DM XRCC4\textsuperscript{157} into pET-32XT expression vector to produce DM XRCC4\textsuperscript{157} protein with Thioredoxin (Trx) tag. The production of DM Trx-tagged XRCC4\textsuperscript{157} started by cutting both DM pGEX-4T-1 – XRCC4\textsuperscript{157} and pET-32XT - Smad6 MH2 domain with BamHI and XhoI (Figure 3-10). Then, the cut plasmid DNA’s were resolved on 1% agarose gel in order to confirm that the cutting has been successful (Figure 3-11). Then, the DM XRCC4\textsuperscript{157} fragment on 504 bp and the pET-32XT vector on 5.9 Kp were cut out and purified using QIAquick Gel Extraction Kit.

After that, the purified DM XRCC4\textsuperscript{157}-BamHI/XhoI fragment was ligated into the pET-32XT-BamHI/XhoI vector. The ligation reaction was transformed into competent *E. coli* DH5α cells as described in Methods and plated onto medium with ampicillin. Single colonies were picked from the plates and were grown overnight. Plasmid DNA was purified using a Miniprep Kit. To test for plasmid DNA with correct insert, purified DNA was digested with BamHI and XhoI. The resulting DNA fragments were separated on a 1% agarose gel to confirm whether the cloning of XRCC4\textsuperscript{157} into pET-32XT vector had been successful (Figure 3-12). The results show that eight out of the twelve minipreps (lanes 2-6, 8, 9 and 12) have bands on the right size, 504 bp and 5.9 Kp. Lane 1 probably contains supercoid uncut plasmid. Lane 7 and 10 seems to have pET-32XT harboring the original insert (Smad6-MH2 domain). Lane 11 contains a religated vector with no insert. Minipreps 4 and 8 were sent for sequencing and confirmed positive.
After that, the DM XRCC4\textsuperscript{157} – pET-32XT plasmid was transformed in \textit{E. coli} BL21 for protein expression to produce Thioredoxin-tagged XRCC4\textsuperscript{157} DM. The protein was produced and purified as described in Methods. Then, fractions of the lysate, flow through, washes, and eluted protein were run on a 10\% SDS-PAGE stained with coomassie blue. Scanned image of the coomassie stained SDS-PAGE gel is shown in Figure 3-13. The result from the gel analysis shows a band corresponds to the molecular size of the Trx- XRCC4\textsuperscript{157} protein, which is around 37 Kd. The concentration of the protein was measured by a spectrophotometer (A280) and calculated to be 5.49 mg/ml (149.5 \textmu\text{M}). The total amount of the purified fusion protein obtained from 1 L culture was 21.96 mg.
Figure 3-10: The strategy of cloning the DM XRCC4\textsuperscript{157} gene into the pET-32XT vector and production of Thioredoxin tagged XRCC4\textsuperscript{157} protein.
Figure 3-11: DNA segment encoding DM XRCC4$^{157}$ (M61R and F106E) was cut out from the pGEX-4T-1 vector and the pET-32XT was linearized. Lane 1: The DM XRCC4$^{157}$ is cut out from the pGEX-4T-1 vector with *Bam*HI and *Xho*I. Two bands were formed; the lower band on about 504 bp indicates the DM XRCC4$^{157}$ insert while the upper on 4900 bp indicates the rest of pGEX-4T-1 vector. Lane 2: to form single band around 4900 bp. Lane 2: Removal of Smad6 MH2 domain insert from pET-32XT vector by cutting the DNA plasmid with *Bam*HI and *Xho*I. Two bands were formed; the upper band on 5900 bp indicates the pET-32XT vector while the lower band on 550 bp indicates the Smad6 MH2 domain insert. The DM XRCC4$^{157}$ fragment and the pET-32XT vector were cut out from the gel and purified.
Figure 3-12: Cloning of Double-Mutant XRCC4<sup>157</sup> (M61R and F106E) into pET-32XT Vector. The DM XRCC4<sup>157</sup> fragment was cloned into pET-32XT vector. Then, the XRCC4<sup>157</sup> was cut out from the pET-32XT vector with BamHI and XhoI restriction enzymes. Two bands were formed; the upper band is around 5900 bp indicating the pET-32XT vector and the lower band is around 504 bp indicating the DM XRCC4<sup>157</sup> insert. The results show that the samples lanes 2-6, 8, 9 and 12 have an insert on the expected size.
Figure 3-13: Purification of DM Trx-XRCC4\textsuperscript{157} protein as determined by SDS-PAGE. 10% polyacrylamide gradient gel stained with coomassie blue. Lane 1: 25 µL of lysate was loaded (out of 10 mL). Lane 2: 25 µL of flow through was loaded (out of 10 mL). Lane 3: 30 µL of washes (out of 20 mL). Lane 4-9: 25 µL of different protein elutions (each elution was 2 mL). The results show that the samples in lanes 4-9 have the DM Trx-XRCC4\textsuperscript{157} protein on the expected size (37 Kd).
3.5 mRNA Display and In Vitro Selection

3.5.1 Preparation of the DNA libraries

The five DNA libraries have been synthesized in an innovative way such that they can produce peptide libraries that have both building block and large scaffold diversity. The structure of each DNA library was described in Methods. In order to ensure that the libraries were synthesized as designed, the composition and nucleotides distribution of the DNA libraries were verified. Each library was first PCR amplified and resolved on 2% agarose gel (Figure 3-14). As expected a band appears around 150 bp that corresponds to the size of the DNA template. Then, each library was gel purified by QIAquick Gel Extraction Kit and ligated into the StrataClone vector. The ligated reaction was transformed into StrataClone SoloPack competent cells and plated onto medium with kanamycin. Ninety-six single colonies were picked from the plates and sent for sequencing.

The sequences of the libraries were analyzed by using Bio-Edit software. As shown in Figure 3-15, there are two fixed cysteines; one at the N-terminus and another one after 2, 4, 6, 8 or 10 random amino acids, as expected. In the random region of our libraries we used NNS codons in which N denotes A, T, C, or G, and S denotes G, or C. The benefit of using NNS codons is to increase the diversity of the libraries and to reduce stop codon usage, which will improve the percentage of library members without stop codons to 62% compare to NNN codons. Therefore, once would expect for an ideal library that each library will have 25% distribution of each nucleotide in the NN part and 50% distribution of either G or C in the S part of the NNS codon. As it is illustrated in Figure 3-16, the A, T, C, and G in the NN part are almost equally distributed in each library. Also, there is almost 50% distribution of G and C in the S part. The presence of A or T in the S part may be due to mutation or frameshift.
Figure 3-14: DNA library on 2% agarose gel. Lane 1: CX2CX4 DNA library. Lane 2: CX4CX8 DNA library. Lane 3: CX6CX6 DNA library. Lane 4: CX8CX4 DNA library. Lane 5: CX10CX2 DNA library. Lane 6: no template. The upper bands were cut and gel purified.
Figure 3-15: DNA libraries sequence. Each library was PCR-amplified and sample clones were sequenced. The predicted amino acid sequence from each clone is shown.
Figure 3-16: Distribution of the nucleotides in the NNS codon. Panel A shows the distribution of A, T, C, and G in the NN part. Panel B shows the distribution of the G and C in the S part.
3.5.2 Small Scale of *In Vitro* Transcription and Translation

The transcription of the five DNA templates was carried out to produce five mRNA libraries. The transcription reaction of each template was loaded on urea PAGE in order to separate the mRNA from the DNA (Figure 3-17). The mRNA band was visualized by UV shadowing and the band cut out of the gel. The mRNA was eluted and ethanol precipitated and the concentration of each library was measured by spectrophotometer (A260).

Following transcription, a small scale of an *in vitro* translation was carried out for each library. The purpose of this is to ensure that our mRNA templates are translatable. We used the DYKM template as a standard mRNA template. DYKM was originally designed to test analogs of 4 of the natural amino acids (Hartman et al., 2007). It has a translatable region of 4 amino acids (D, Y, K, and M) and 6xHis tag in the c-terminus. We are using DYKM template here to compare the translation efficiency of our mRNA libraries to it. The translation reactions were carried out with $^{35}$S-methionine as the N-terminal residue. The average amount of the translated peptides was about 0.3 pmole (Figure 4-18). This is lower than what we were expecting; nevertheless, this indicates that the mRNA templates can be translated. Also, we will use large scale of *in vitro* translation reaction, which will increase the amount of translated peptides.
**Figure 3-17: Separation of the mRNA on a 6% PAGE/urea gel.** The 1 mL transcription reaction of CX4CX8 library was loaded on the gel without adding loading dye. The mRNA band (upper band) was visualized by UV shadowing. The lower band shows the separated DNA. 10 µL of bromophenol loading dye (LD) was loaded.
Figure 3-18: In vitro translation. A final translation reaction of 50 µL was carried out for each mRNA library showing that they are translatable.
3.5.3 In Vitro Selection

3.5.3.1 First Round

The formation of mRNA display natural cyclic peptide library started by the transcription of the five DNA templates that encode peptides with 15 random amino acids interspersed with two cysteine residues. Following transcription, the five mRNA libraries were combined together to a final concentration of 2 µM and photo cross-linked to a 3′-puromycin oligonucleotide. Then, the combined mRNA libraries were translated on a 5 mL translation scale in a completely reconstituted translation system with 20 natural amino acids. After that, mRNA peptide fusions were immobilized on an oligo-dT column and cyclized using dibromoxylene. These two processes were performed to remove peptides that are not conjugated to their mRNAs, to remove mRNAs that were not photo crosslinked by psoralen and to produce cyclic mRNA-peptide fusions. Then, the pmole amount of the mRNA-peptide fusions was calculated based on the scintillation counts of the $^{35}$S-Met to be 7 pmole. After that, the peptide-fusions were reverse transcribed to create double-stranded template and to eliminate any unwanted RNA secondary structures and render the nucleic acid portion of the fusion more stable. The percentage of the peptide fusions recovered after this step was 37% (2.6 pmole). Then, the mRNA-peptide fusions were purified by Ni-NTA purification. Ni-NTA binds to the C-terminal 6xHis tag incorporated into the peptide in order to remove the truncated mRNA peptide fusions. The percentage of the peptide fusions recovered after Ni-NTA was 45% (1.18 pmole). The eluents with high counts were collected and dialyzed overnight at 4°C in 1X selection buffer. The overnight dialyzed peptide mRNA-fusions was collected and the percentage calculated to be 122% (1.44 pmole), which is about a trillion peptide variants that are going to the selection procedure.
In the selection step of the first round, we first incubated our library with GST protein on magnetic beads (pre-clear). This is to remove all non-specific peptide inhibitors that bind to the GST or glutathione magnetic beads. The mRNA-peptide fusions in the flow through and washes were counted to find that 53% (0.76 pmole) of the peptide fusions was recovered. Then, the flow through and washes with high counts were collected and bound to WT GST-tagged XRCC4\textsuperscript{157} fragment. This fragment includes the complete head domain known to contain the XLF interface but lacks the C-terminal DNA ligase IV-binding region that could interact with the mRNA component of our mRNA-peptide fusions. Those peptides that bound to WT GST-XRCC4\textsuperscript{157} were captured on glutathione magnetic beads. Then, the beads were washed and suspended in selection buffer and the percentage of the \textsuperscript{35}S-labeled peptides bound to the resin was measured to be 2.1% (0.03 pmole). The DNA-mRNA duplex fused to the XRCC4-captured peptides was amplified by PCR.

3.5.3.2 Second round

The resulting DNA from the first round was \textit{in vitro} transcribed and then new mRNA-peptide fusions were formed. However, from the second round and subsequent rounds the translation scale is going to be 500 \textmu L. Also, for the second and third rounds, we used GST protein on glutathione magnetic beads for pre-clearing. The amount of the mRNA-peptide fusions after \textit{in vitro} translation, oligo dT purification and cyclizationin was 2.75 pmole. The percentage recovered after reverse transcription was 51% (1.4 pmole), after Ni-NTA was 52% (0.73 pmole), after dialysis was 47% (0.35 pmole) and after pre-clearing with GST protein was 95% (0.33 pmole). The percentage peptide fusions that bound WT GST-XRCC4\textsuperscript{157} and captured on glutathione magnetic beads was 1.14% (0.004 pmole).
3.5.3.3 Third round

The amount of the mRNA-peptide fusions after \textit{in vitro} translation, oligo dT purification and cyclization was 6.97 pmole. The percentage recovered after reverse transcription was 29% (2 pmole), after Ni-NTA was 56% (1.11 pmole), after dialysis was 49% (0.55 pmole) and after pre-clearing with GST protein was 134% (0.73 pmole). The percentage of the peptide fusions that bound to WT GST-XRCC4\textsuperscript{157} and captured on glutathione magnetic beads was 2% (0.011 pmole).
Figure 3-20: Recovery of the mRNA-peptide fusions during the first three rounds of *in vitro* selection process. Panel A shows the percentage of the mRNA-peptide fusions recovered after purifications, dialysis, pre-clearing and incubation with GST-XRCC4\textsuperscript{157} protein during the first three rounds. Panel B shows only the percentage of the mRNA-peptide fusions recovered after incubation with the GST-XRCC4\textsuperscript{157} protein at each round of selection.
Figure 3-21: Scheme of mRNA Display and *in vitro* selection for XRCC4\textsuperscript{157}. 
IV. DISCUSSION AND FUTURE DIRECTION

4.1 Therapeutic potential of suppressing DSB repair

Induction of DNA damage by a radio- or chemotherapeutic agent is still a key element in cancer therapy. However, the dose of DNA damaging agent that can be delivered to the cancer cells is limited due to the toxic side effects produced by such agents. Exposure to cytotoxic agents also results in the activation of accurate or inaccurate DNA repair mechanisms that allow the survival of cancer cells leading to inadequate tumor response to radio- or chemotherapy (Bolderson et al., 2009). Therefore, targeting the DNA repair mechanisms that are required for cancer cell survival will increase the efficiency of radio- or chemotherapy. Furthermore, since cancer cells often have disrupted DNA damage and repair responses, they may be more susceptible than normal cells to interference with normal repair mechanisms.

The majority of ionizing radiation-induced DSBs are repaired by NHEJ in mammalian cells, which has raised the possibility of directly targeting the NHEJ proteins and their interactions for the purpose of radiosensitzation of tumor cells. DNA-PK has been shown to be a good target for radiosensitzation of tumor cells, in that cells deficient in Ku70/80 or the DNA-PKcs are sensitive to DSBs induced by IR. The use of NU7441, a DNA-PK inhibitor, has shown an increase in the radiosensitivity of different cell lines and in xenograft models in preclinical trials (Zhao et al., 2006). However, inhibitors of DNA-PK will also interfere with its roles in telomere maintenance (Gilley et al., 2001; Goytisolo et al., 2001) and immune functions (W. Chu et al., 2000; Dragoi et al., 2005). Therefore, finding a more selective inhibitor is required.
Inhibition of XRCC4 is a particularly attractive target for overcoming tumor radioresistance because its only known function is NHEJ (Schaue & McBride, 2005). Also, it has been shown that breast tumor cells can be radiosensitized by adenovirus-mediated overexpression of a fragment of XRCC4 that binds ligase IV but does not support NHEJ (Jones et al., 2005). In addition, the interaction between XRCC4 and XLF should be a much more susceptible target for radiosensitization because, whereas XRCC4 and ligase IV form a tight complex in cells, the XRCC4-XLF interaction is transient. With any radiosensitizer, there will undoubtedly be some sensitization of normal tissues, and whether inhibition of NHEJ will ultimately improve therapeutic index is difficult to predict and will only be determined by future preclinical and clinical studies. However, as many breast tumor cells have partial deficiency in HRR and/or upregulated NHEJ (Gudmundsdottir & Ashworth, 2006; Nagaraju & Scully, 2007; H. Wang et al., 2001; Zhuang et al., 2006), raising the possibility that they may indeed be sensitized more than surrounding tissue when NHEJ is suppressed. Furthermore, when breast tumors treated with radiation metastasize, the subsequent tumors are often highly radioresistant due to upregulation of these downstream pathways (Johnston et al., 2006; Das et al., 2007) raising the possibility that the therapeutic index with metastasized tumors could be even greater.

Therefore, in order to inhibit the interaction between XRCC4 and XLF, a powerful technique such as mRNA display should be implemented. This technique has been used to find inhibitors of protein-protein and protein-ligand interactions (Schilppe et al., 2012; Getmanova et al., 2006). We chose mRNA display over the other selection techniques available because it possesses several advantages. For example, the presence of the covalent bond between the peptide and mRNA makes it highly stable compare to a non-covalent bond linkage in other techniques such as the ribosome display (Hui & Rihe, 2011). Therefore, in mRNA display the
selection process can be performed under stringent conditions in order to optimize the selection of specific binders from non-specific ones. Another advantage of mRNA display is the ability of producing large libraries of different unique sequences. mRNA display is totally an in vitro selection technique that can produce as many as $10^{12}$-$10^{14}$ unique sequences while other techniques that rely on an in vivo process have limitation such as phage display that produces diverse sequences of up to $10^{10}$ only. Other advantage of mRNA display includes the easy removal of abundant sequences from the starting library increasing the chance of selection non-abundant sequences.

Therefore, we produced WT GST-XRCC4\textsuperscript{157} protein in order to use it for the selection of the peptide inhibitors of the XRCC4/XLF interaction. We used truncated XRCC4 because the interaction between XRCC4 and XLF has been shown to be through the head domains (Ropars et al., 2011; Malivert et al., 2010). We also produced DM GST- XRCC4\textsuperscript{157} comprising M61R and F106E; M61 and F106 are two residues that are required for the interaction between XRCC4 and XLF. Production of DM GST-XRCC4\textsuperscript{157} protein is to use it in a pre-clearing step to eliminate any non-specific peptide inhibitors that might bind to regions other than the XLF interface of XRCC4. We have conducted circular dichroism (CD) study to see if the mutated XRCC4 protein is folded properly. As shown in Figure 4-1, the DM XRCC4\textsuperscript{157} protein has similar CD spectrum to that of the WT protein suggesting that the mutant protein is folded correctly and therefore will be effective in removing peptides that bind elsewhere on XRCC4. We also produced DM Trx-XRCC4\textsuperscript{157} protein to increase the stringency of the selection system. The DM Trx-XRCC4\textsuperscript{157} will be used during the elution step in the selection process.

Five DNA libraries have been synthesized and we have sequenced sample clones of each library to find that they have the sequences as they were designed. We have finished three
rounds of the *in vitro* selection of the XRCC4/XLF interaction inhibitors. About a trillion mRNA-peptide variants went to the selection in the first round. We were expecting about 10 trillions unique peptide fusions to start the selection; nevertheless, the amount we obtained should be sufficient to find selective inhibitors. The average number of unique molecules produced by an mRNA display technique is about $10^{12}$-$10^{14}$ unique sequences. The percentage of the mRNA-peptide fusions that bound to the XRCC4 after the first round was 2.1% of those peptide fusions collected after pre-clearing. This is a good result since abundant sequences of non-selective inhibitors were removed leaving the selective ones to go for the next round. The recovery after the second round was 1.14%. The decrease in the percentage of the peptide fusions recovered after the second round may be due the fact that we reduced the amount of the translation reaction from 5 mL in the first round to 500 µL in the second round. The third round had a recovery of 2%. Even though the recovery is not yet increasing significantly, we would expect to have a higher percentage of peptide fusions recovered after subsequent rounds. Moreover, the progression in recovery from the pre-clearing step in rounds 2-3 suggest that those peptides bind to the GST tag are in fact being removed by pre-clearing.

### 4.2 Future Direction

The selection process will continue for several rounds until the percentage of the mRNA-peptide fusions eluted begins to plateau signifying that enrichment has stopped. The DNA from that round will be cloned into vector and sequenced to determine the natural peptide motifs that bound to the XRCC4/XLF interface. Our lab experience suggests that this process will take 6-8 rounds and will lead to 6-8 families of function natural cyclic peptides. We will choose 1 or 2 peptides representative from each of the families for further analysis (Figure 4-2). To test these
peptides, we will use highly sensitive NHEJ assays (Akopiants et al., 2009). We expect that 50 pmoles of each peptide will be enough to do several replicate in vitro assays with that peptide.

Once the candidate peptides have been ranked with respect to potency in inhibiting NHEJ in extracts, 4-5 peptides will be chosen for evaluation of cytotoxicity. Typically, these will be the most potent inhibitors, but with some preference for diversity in primary structure as well. Each peptide will be synthesized on a mmole scale by microwave-enhanced Fmoc-based solid phase peptide synthesis and purified by HPLC. During the course of the synthesis we will append a C-terminal cell-penetrating peptide sequence (Arg₉), which will enable cell permeability of the peptides for the assays in cell culture (Goun et al., 2006).

Because the ultimate clinical use of the peptides would likely be for treatment of tumors resistant to radiotherapy, the peptides will be ranked according to their ability to radiosensitize MCF-7/HER2 cells. This line is a derivative of MCF-7 that has been engineered to overexpress HER2, and consequently is resistant to both radiation and chemotherapeutic drugs (Liang et al., 2003). Clonogenic survival assays will be performed and one or two peptides will be chosen for further studies in multiple cell lines. For example, they will be examined in MDA-MB-231-BR cells, a clone derived from MDA-MB-231 cells that specifically metastasized to brain in mice (Yoneda et al., 2001), as well as in 4T1 cells, a murine mammary tumor that metastasizes to lung. These lines are chosen, in part, in anticipation of future preclinical studies of metastatic disease in mice.

Also, we would like to evaluate the modes of radiation-induced cell death following irradiation in the presence or absence of peptide inhibitors, specifically, senescence (β-galactosidase staining, cell morphology), apoptosis (fluorescent DNA end-labeling (TUNEL) and prodipidium iodide flow cytometry), mitotic catastrophe (formation of micronuclei in
binucleated cells) and autophagy (acridine orange staining, LC3-1 cleavage by western blot, and electron microscopy) (Di et al., 2009; Klionsky et al., 2008).
Figure 4-1: CD spectra of wild type and double-mutant GST-XRCC4\textsuperscript{157}. The similarity in the UV spectra between the wild type (blue line) and double-mutant (pink line) GST-XRCC4\textsuperscript{157} indicates that the double mutant is folded correctly.
Figure 4-2: Overall scheme for synthesis, selection and evaluation of XRCC4/XLF interaction inhibitors.
REFERENCES


American Cancer Society. What are the key statistics about breast cancer?, 2012 (7/15/2012).


Mohammed Al Mohaini was born on June 06, 1982 in Alahsa, Saudi Arabia. He graduated with a degree of Bachelor in Pharmaceutical Sciences from King Saud University, Riyadh, Saudi Arabia in 2006. He worked in King Abdulaziz Hospital, Alahsa, Saudi Arabia as pharmacist in the Pharmaceutical Care department from 2006 to 2008. He worked at King Saud bin Abdulziz University for Health Sciences (KSAU-HS) from 2008-2009 as teaching assistant of pharmacology. He was awarded a scholarship from KSAU-HS to study Master and PhD abroad. He joined Virginia Commonwealth University in 2010.