

Alma Mater Studiorum – Università di Bologna

DOTTORATO DI RICERCA IN

Scienze Biotecnologiche e Farmaceutiche

Ciclo XXX

Settore Concorsuale: 05/E1

Settore Scientifico Disciplinare: BIO/10

**Targeting RAD51-BRCA2 interaction  
in the search for novel anticancer and  
chemosensitizer drug candidates**

**Presentata da:** Fabrizio Schipani

**Coordinatore Dottorato**

Prof. Santi Mario Spampinato

**Supervisore**

Prof. Andrea Cavalli

Co-Supervisore

Dott.ssa Stefania Girotto

**Esame finale anno 2018**

# TABLE OF CONTENTS

<b>Abstract</b>	<b>I</b>
<b>Chapter 1: Introduction</b>	
<b>1.1 The DNA damage response</b>	<b>1</b>
1.1.1 The signal transduction: sensors, mediators and effectors	3
1.1.2 Targeting DDR in cancer for drug development	4
<b>1.2 Double strand breaks</b>	<b>8</b>
1.2.1 Classical non homologous end joining (C-NHEJ)	9
1.2.2 Alternative non homologous end joining (A-NHEJ)	10
1.2.3 Homologous Recombination (HR)	10
<b>1.3 RAD51-BRCA2: a key interaction for a successful DSB pathway repair</b>	<b>12</b>
<b>1.4 Double strand break repair pathway: a target for cancer therapy</b>	<b>21</b>
1.4.1 Impairment of the double strand break repair pathway and synthetic lethality	25
<b>Goal of the research project</b>	<b>28</b>
<b>Chapter 2: Materials and Methods</b>	
<b>2.1 Materials</b>	<b>33</b>
2.1.1 Growth media	33
2.1.2 Solutions and reagents	36
2.1.3 Peptide	36
2.1.4 Antibodies	36
<b>2.2 Molecular biology</b>	<b>37</b>
2.2.1 Plasmids for recombinant proteins expression	37
2.2.2 Competent cells	40
2.2.3 XL1-Blue, BL21(DE3)pLysS and Rosetta2(DE3)pLysS cells transformation	41
2.2.4 Plasmid DNA purification	41
2.2.5 hRAD51 subcloning into pET15b vector	41
2.2.6 Site-directed mutagenesis	42
<b>2.3 Biochemical methods</b>	<b>44</b>
2.3.1 Agarose gel electrophoresis	44
2.3.2 Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE)	44
2.3.3 In-gel digestion of proteins separated by polyacrylamide gel electrophoresis	44
2.3.4 Western blot	45
2.3.5 Recombinant GST-hRAD51: expression, purification and protease inhibition trials	46
2.3.6 Recombinant His-hRAD51: expression, purification and degradation trials	48
2.3.7 His-hRAD51-BRC4 complex: expression and purification	50
2.3.8 His-hRAD51 F86E-BRC4 complex: expression and purification	52

<b>2.4 Biophysical methods</b>	53
2.4.1 Protein and plasmid DNA quantification	53
2.4.2 Surface Plasmon Resonance (SPR)	54
2.4.3 Microscale Thermophoresis (MST)	54
2.4.4 Nuclear Magnetic Resonance (NMR)	55
2.4.5 Crystallization trials	56
<b>2.5 His-hRAD51 characterization</b>	58
2.5.1 Size exclusion chromatography	58
2.5.2 Negative staining transmission electron microscopy	58
2.5.3 Dynamic Light scattering (DLS)	58
<b>Chapter 3: Results and discussion</b>	
<b>Optimization of a protocol for the overexpression and purification of full-length hRAD51 in bacteria cells</b>	61
<b>3.1 Expression and purification of GST-hRAD51</b>	61
3.1.1 Degradation trials	64
<b>3.2 Expression and purification of His-hRAD51</b>	68
3.2.1 Degradation trials	71
3.2.2 His-hRAD51 characterization	77
<b>Biophysical characterization of hRAD51-hit molecules interaction through Surface Plasmon Resonance (SPR), Nuclear Magnetic Resonance (NMR) spectroscopy and Microscale Thermophoresis (MST)</b>	80
<b>3.3 Biophysical characterization of hRAD51-small molecules interaction</b>	80
3.3.1 Surface Plasmon Resonance (SPR)	81
3.3.2 NMR WaterLOGSY experiments	83
3.3.3 Microscale Thermophoresis (MST)	84
<b>3.4 Biophysical characterization of hRAD51-fragments interaction</b>	89
3.4.1 NMR Fragment Based screening	89
3.4.2 Microscale Thermophoresis analysis of the hit fragments	91
<b>X-ray structural characterization of hRAD51-BRCA2</b>	94
<b>3.5 Expression and purification of hRAD51-BRC4 complex</b>	94
<b>3.6 Expression and purification of hRAD51 F86E-BRC4 complex</b>	97
<b>3.7 Crystallization trials of hRAD51 F86E-BRC4 complex</b>	99
<b>Chapter 4: Conclusions and perspectives</b>	102

## **Appendix**

<b>A1 Fragment based drug discovery (FBDD)</b>	109
<b>A2 NMR-based screenings</b>	111
A2.1 WaterLOGSY	111
A2.2 <sup>19</sup> F NMR-based experiments	113
A2.3 n-Fluorine Atoms for Biochemical screenings (n-FABS)	114
<b>A3 Microscale Thermophoresis (MST)</b>	116
<b>A4 Surface Plasmon Resonance (SPR)</b>	118
<b>Glossary</b>	122
<b>Bibliography</b>	126

## **Abstract**

Genome integrity is constantly affected by DNA damage and replication errors due to exogenous and endogenous sources. Cells have developed several strategies to counteract these threats detecting DNA damages and mediating its repair through different mechanisms depending on the specificity of the DNA lesion.

Cancer cells are more sensitive to these insults compared to normal cells and most therapeutic strategy indeed rely on inducing specific lethal DNA damages. The urgent need for new therapeutic strategies in cancer treatment is due to the frequent development of resistance to therapeutic induced DNA damage.

Through a multidisciplinary project, comprising a computational, a synthetic, a biology as well as a biophysical contribution, we aim at applying a synthetic lethality approach in cancer cell therapy by preventing, with small molecules, the viable hRAD51-BRCA2 interaction. The latter interaction plays a key role in the repair of double strand breaks (DSBs) within the homologous recombination (HR) pathway.

The leading idea of our project is mimicking the synthetic lethality induced in BRCA2 defective oncology patients by olaparib, a poly ADP ribose polymerase (PARP) inhibitor. We aim at increasing cancer cells sensitivity to PARPi and at broadening PARPi applicability by administering, in combination with the latter inhibitors, hRAD51-BRCA2 small molecules disruptors also to individuals without BRCA2 mutations.

The goal of my PhD research project is the characterization of the hRAD51-BRCA2 interaction through a biophysical approach, in order to gain kinetic and thermodynamic insights on this critical junction of the HR pathway, and through a structural approach, to elucidate the atomic details of the interaction. My research work is concurrently a critical support for the overall project, which aims at the development of a new drug and at the validation of a broaden applicability of the synthetic lethality concept.

---

---

Chapter one

# Introduction

---

---

## **1.1 The DNA damage response (DDR)**

Genome integrity is constantly affected by DNA damage and replication errors due to exogenous and endogenous sources. Physical or chemical agents are the most common cause of environmental DNA damage. Ionizing radiations (IR) and ultraviolet (UV) light are examples of physical genotoxic agents. IR can induce oxidation of DNA bases and cause single strand or double strand DNA breaks (SSB or DSB), while UV can induce up to 105 DNA lesions per cell per day<sup>1</sup> generating pyrimidine dimers and pyrimidine (6-4) pyrimidone photoproducts. Chemical agents used in chemotherapy can also generate several DNA lesions depending on their mechanism of action: alkylation of DNA bases (i.e. temozolomide), intra or interstrand crosslinks, linking bases of the same or of different DNA strands respectively (i.e. cisplatin), SSB or DSB, locking the formation of topoisomerase-DNA covalent cleavage complexes (i.e. etoposide).

Cells have actually already developed DNA repair mechanisms to retrieve specific types of DNA errors and lesions. The DNA damage response (DDR) is a complex transduction pathway that senses replication stress and DNA damage in order to counteract these genome integrity threats<sup>2,3</sup>. Cells are provided with all the enzymatic tools (nuclease, helicases, polymerases, topoisomerases, recombinases, ligases, glycosylases, demethylases, kinases and phosphatases) required to remodel and repair DNA. Nonetheless, their activities must be regulated in a spatiotemporal way to prevent the formation of deleterious DNA structures during physiological cellular processes. Eukaryotic cells have developed specific strategies to recruit the right factors in the right place at the right time. Cells present different mechanisms of DNA repair depending on the type of DNA lesion involved (Figure 1.1). Mismatch repair (MMR) corrects mispaired DNA bases, while base excision repair (BER) acts through the excision of the damaged bases<sup>4,5</sup>. Nucleotide excision repair (NER) removes both pyrimidine dimers and intrastrand crosslinks

removing an oligonucleotide of 30 base pairs (bp) containing the damaged site or excising the interstrand crosslink with the help of proteins involved in the Fanconi anemia syndrome<sup>1,6</sup>, respectively. Single strand break repair (SSBR)<sup>7</sup> and double strand break repair (DSBR) repair SSB and DSB, respectively. Regarding DSBs, they can be restored by two different pathways: non homologous end joining (NHEJ) that promotes the potentially inaccurate rejoining of DSBs, and homologous recombination (HR) that provides an error free mechanism of repair<sup>8</sup>.

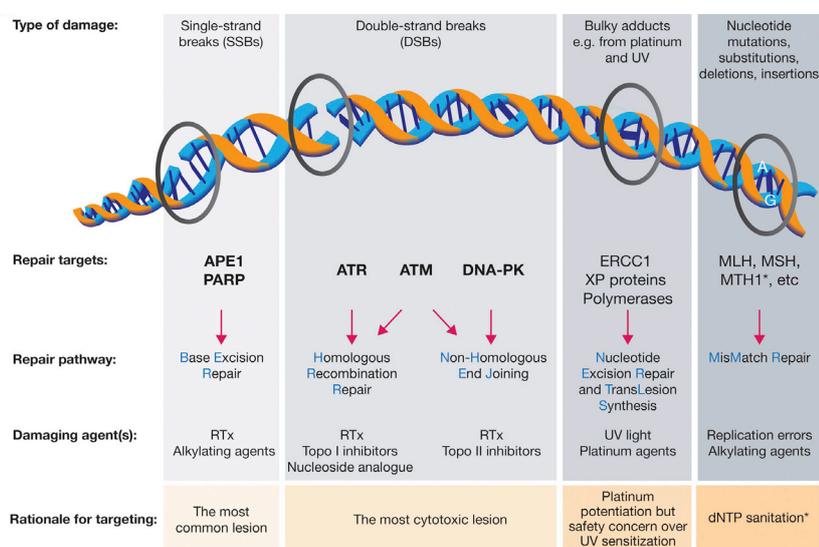


Figure 1.1: DDR pathways and cell cycle targets. (Taken from O'Connor et al.<sup>9</sup>).

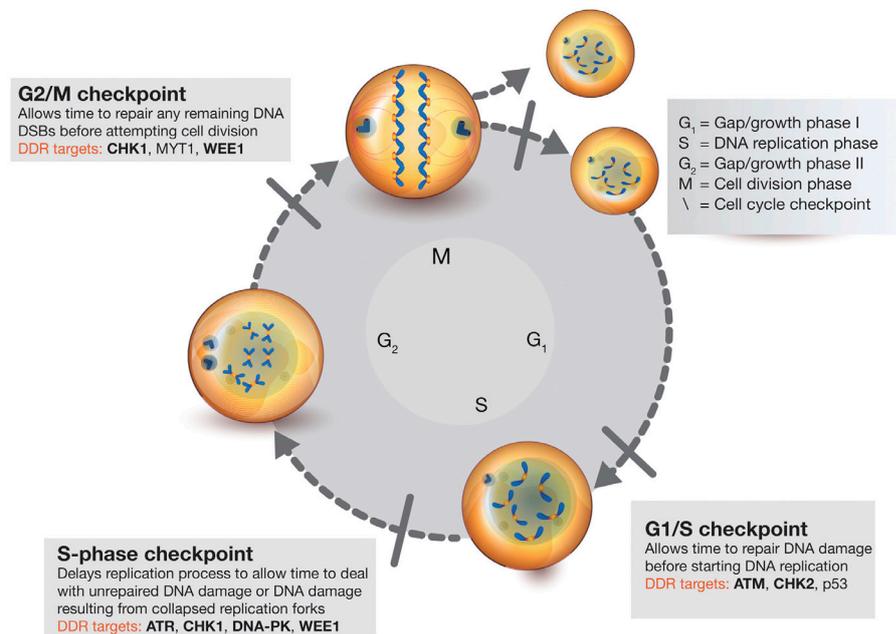
### **1.1.1 The signal transduction: sensors, mediators and effectors**

Upon a DNA lesion, the activation of the pathways involved in sending an SOS to initiate the repair of the damage by the cell is fundamental. The signal transduction mediated by the DDR occurs through the action of sensors, mediators, transducers and effectors. This regulatory pathway is well conserved in mammals; indeed, homologs exist for all the components. The most important DNA damage sensors belong to the phosphatidylinositol 3-kinase-like protein kinase (PIKKs) family (ataxia-telangiectasia mutated kinase (ATM), ataxia telangiectasia and Rad3-related protein (ATR) and DNA protein kinases, (DNA-PKc) and to the poly(ADP) ribose polymerase (PARP) family. ATM and DNA-PK are activated in presence of DSBs<sup>2,10</sup> and they mediate the HR and the NHEJ, respectively. ATR, in complex with its regulatory subunit ATRIP<sup>11</sup>, is recruited by the ssDNA-RPA complex in case of DNA polymerase stall therefore acting during SSB events. The PARP family has sixteen members but only PARP1 and PARP2 are involved in the DDR<sup>12</sup>. The role of the DNA damage sensors is the recruitment of the mediator proteins to activate a signaling cascade. Mediators can act as substrates, as recruiters of other substrates or as scaffolds to allow the formation of complexes for the repair at the damage site, where the phosphorylation of the histone variant H2AX on the serine 139 by ATM, ATR or DNA-PK<sup>13</sup> occurs. Phosphorylation of H2AX plays a key role in DDR and is required for the assembly of DNA repair proteins at the sites containing damaged chromatin and for activation of checkpoints proteins, which arrest the cell cycle progression<sup>14</sup>. Mediator proteins, such as MDC1, and BRCA1 for ATM, TopBP1 (topoisomerase-binding protein1) and Claspin for ATR can activate the effector kinases. Finally, the effectors execute the functions of the DDR allowing the direct control of DNA repair, genomic stability and cell cycle regulation.

The latter ones include several proteins such as substrates of the CHK and PIK kinases, protein involved in DNA replication (i.e. MCMs, ORCs, RFC), transcription regulation and cell cycle control (i.e. p53, Cdc25, Nbs1 and BRCA1).

### **1.1.2 Targeting DDR in cancer for drug development**

Treatment of cancer by radiotherapy or by chemotherapy causes significant collateral damages to normal tissues as well as side effects. The inhibition of the DDR in patients with cancer, lacking specific DDR functions could be a key strategy for target therapy. This concept is based on discoveries that show that the response to the damage is different depending on the cell cycle status (Figure 1.2). For example, cells in G1 (Gap 1 phase), lacking sister chromatid DNA, can repair DSB just through NHEJ pathway. Furthermore, there are also differences in the roles of the checkpoint at different stages of the cell cycle and in the DDR factors involved. For example, G1/S (synthesis phase) checkpoint allows time to repair DNA damage before starting DNA replication to avoid issues to DNA synthesis. The most important DDR factors involved are ATM, CHK2 and p53. The S phase checkpoint delays DNA replication to allow time to deal with unrepaired DNA damage or DNA damage resulting from collapsed replication forks. The DDR factors involved in this step are: ATR, CHK1, DNA PK and WEE1 (Wee1-like protein kinase). Finally, the G2 (Gap 2 phase)/M (mitosis) checkpoint allows to repair any remaining DNA DSBs before attempting cell division. In this step, DDR proteins such as CHK1, WEE1 and MYT (Membrane-associated tyrosine- and threonine-specific cdc2-inhibitory kinase) are recruited. The G2/M checkpoint represents the last opportunity for the cell to prevent DNA damage before starting the mitosis stage<sup>15</sup>.



**Figure 1.2:** DDR cell cycle targets. (Taken from O'Connor et al.<sup>9</sup>).

Three important key differences in cancer DDR provide a rationale for drug targeting (Figure 1.3).

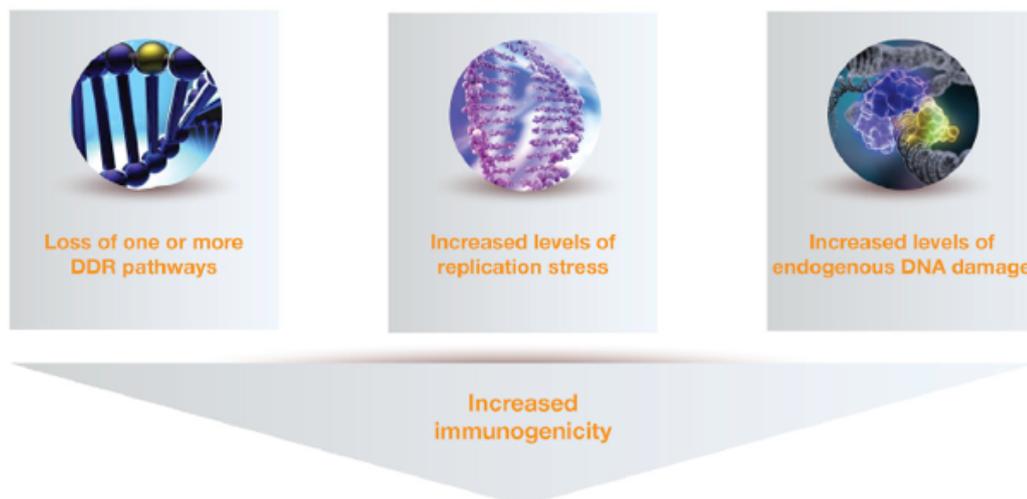
One relevant difference between normal and cancer cells is that most cancers have lost one or more DDR pathways. DDR deficiency results in a dependency on a particular DDR pathway or target for its survival. Thus, identifying a specific inhibitor for that target or relevant pathway is possible to specifically target the tumor cells. This approach is well known as synthetic lethality<sup>16,17</sup>. The proof of this theory is given by the regulatory approval of olaparib (Lynparza), the first PARP inhibitor on the market<sup>18</sup>. Olaparib was initially approved in cancer treatments (i.e. breast and ovarian cancer) in the presence of mutated BRCA1 or BRCA2 genes, which leads to increase in genomic instability as well as increased susceptibility to the treatment with PARP inhibitors.

Another important aspect, proper of cancer cells, is the replication stress. It occurs when the DNA polymerase is uncoupled from the replisome helicase activity<sup>19</sup>, generating extended

ssDNA at the replication fork. After the recognition of ssDNA by RPA (replication protein A) a DDR mediated by ATR<sup>20</sup> is induced. Insufficiency in nucleotide pools can induce a replicative stress too. In cancer cells the deficiency in pRB (retinoblastoma protein), the amplification of cyclin D1 or cyclin E<sup>21</sup> or the deletion of CDK2NA locus<sup>22</sup>, lead to the loss of the G1/S checkpoint. Thus, cells prematurely entry into S phase and DNA replication can start without all the necessary components, causing several issues, especially in the early S phase<sup>23</sup>. Cyclin E or MYC amplification and KRAS mutations increase the DNA replication initiation causing a conflict between the replication and transcription processes<sup>24-26</sup>. MYC overexpression leads also to the production of ROS species (reactive oxygen species) promoting the formation of damaged nucleotide as the 8 oxoguanine<sup>27</sup>. Replication stress, generated by chemotherapy, can therefore act more in cancer than in normal cells<sup>9</sup>. Currently, inhibitors of ATR (AZD6738 and VX970) and of its effector kinase CHK1 (GDC0575, LY2606368 and MK8776) are investigated in clinical trials<sup>9</sup>. However, it has been discovered that there is also a backup pathway involving DNA PK and CHK1 that can stabilize and repair the replication fork, therefore cancer cells can survive to the treatment with ATR inhibitors.

The third approach to target cancer cells is increasing the level of endogenous DNA damage, which is achieved through radiotherapy and systemic chemotherapy. Even though there have been improvements to reduce the exposure of normal tissue to radiation, acute and chronic normal tissue toxicities represent a limiting factor for radiation dose delivery. A strategy to improve the specificity is combining the radiotherapy with the administration of DDR inhibitors. Several examples of DDR targeting agents combined with radiotherapy in preclinical and clinical trials have already obtained satisfactory results<sup>28,29</sup>. However, even though this approach represents a good starting point, many aspects are unknown, for example the dose of DDR inhibitors that

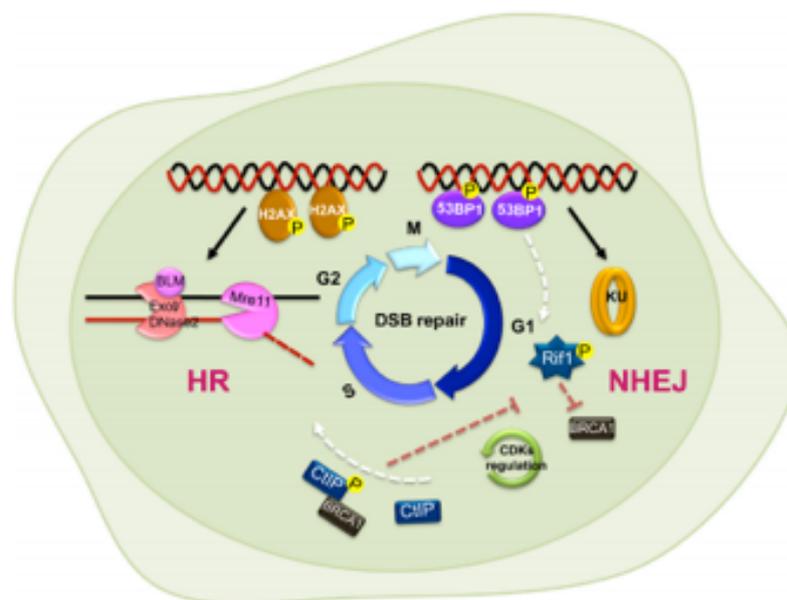
act as radiosensitizer and how much benefit is obtained from an extended exposure to the DDR agent following the irradiation. The main issues related to this strategy are the systemic delivery of chemotherapies and some overlapping side effects of DDR agents (i.e. gastrointestinal and bone marrow toxicity). In spite of these issues, preclinical and clinical studies have shown that chemo-resistant cancers can be sensitized through combination with DNA damaging agents<sup>17,30</sup>, therefore demonstrating the efficacy of combined therapies.



**Figure 1.3:** Major differences between normal and DDR cancer cells that can provide different strategies for drug targeting. (Taken from O'Connor et al.<sup>9</sup>).

## 1. 2 Double strand breaks

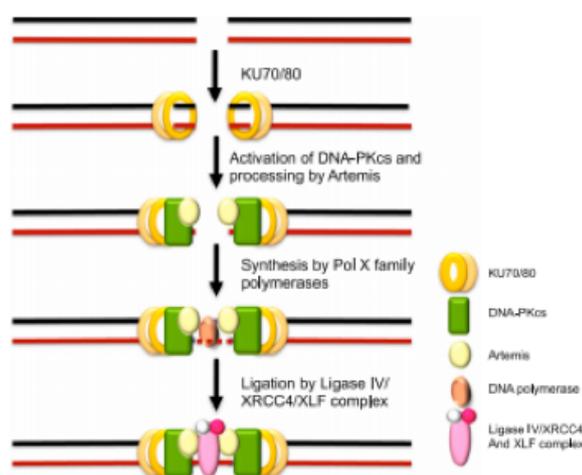
Among the several DNA lesions, DSBs are highly toxic since they can lead to aneuploidy, genetic aberration or cell death. As previously mentioned, DSBs can be generated by different sources, including treatment with genotoxic chemicals and ionizing radiation, collapsed replication forks and other endogenous DNA breaks. DSB repair uses three strategies: a homology-dependent error free HR, a potentially error prone KU dependent classical non homologous end-joining (C-NHEJ) and an error prone alternative non homologous end-joining pathway (A-NHEJ). The choice of one of these pathways depends on the phase of the cell cycle and on the DNA end resection<sup>31</sup> (Figure 1.4).



**Figure 1.4:** Choice between Homologous Recombination and Non Homologous end Joining pathway for DSB repair: HR predominantly occurs in S and G2 phases; NHEJ is active throughout the cell cycle, playing a major role during G1 and M phases (Taken from Srivastava et al.<sup>32</sup>).

### 1.2.1 Classical non homologous end joining (C-NHEJ)

C-NHEJ occurs throughout the cell cycle but especially during G1 and M<sup>33</sup> (Figure 1.5). The KU protein (heterodimer KU70-KU80 in eukaryotes) initially recognizes and binds the DSB in a sequence independent manner. The KU-DNA complex interacts with DNA PK forming a functional basket shaped structure, which encircles a molecule of dsDNA<sup>34</sup>. The asymmetrical ring allows the KU70 subunit to bind the major groove proximal to the DSB and the KU80 subunit the minor groove distal to the break site<sup>35</sup> respectively, preventing the end access. The activation of the DNA-PK complex protein with its auto-phosphorylation induces both the opening of the central DNA-binding cavity and its release from the DNA ends. The following step is the DNA end processing, which depends on the nature of the damage and one or more enzyme activities can be involved. Indeed, in case of abasic or 5'-dRP/AP sites, the end processing occurs through the KU protein lyase activity; the tyrosyl DNA phosphodiesterase activity of TDP2 hydrolyzes 5' phosphotyrosyl-DNA bonds and the polynucleotide kinase/phosphatase (PNKP) generates 5' phosphates and 3' hydroxyl ligable ends. The recruitment of these enzymes depends on their interaction with the KU protein<sup>36</sup>. Finally, DNA ligase IV, together with XRCC4 and XLF (XRCC4 like factor), ties the processed ends.



**Figure 1.5:** Schematic representation of NHEJ repair pathway (Taken from Srivastava et al.<sup>32</sup>).

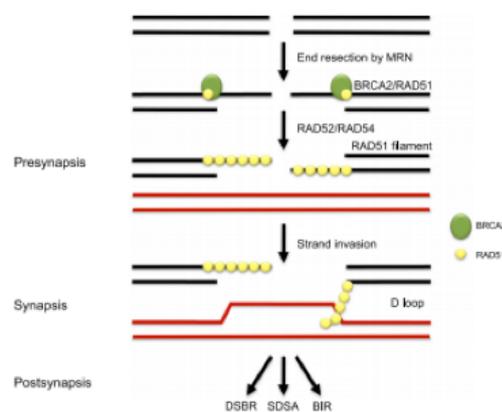
### **1.2.2 Alternative non homologous end joining (A-NHEJ)**

A-NHEJ occurs especially when cells enter the S phase<sup>37</sup>. Many A-NHEJ events, also known as microhomology-mediated end-joining (MMEJ), require end resection and join the ends by base pairing at microhomology sequence, usually 5-25 nucleotides long, generating deletions at the junction. It has been shown that other A-NHEJ pathways do not require microhomology sequences. A-NHEJ shares similar step with C-NHEJ and the two pathways are competitive; in fact A-NHEJ occurs in the absence of the key C-NHEJ factors. Since KU limits poly ADP ribose synthesis, PARP1 tether DSBs under KU defective conditions<sup>38</sup>. PARP1 enzymatic activation leads to autoPARylation, resulting in decrease of PARP1 affinity for nucleosomes and subsequent relaxation of the chromatin. Importantly, PARP1 allows the recruitment of MRN complex, interacting with Mre11 and NBS1 proteins. In addition, it can promote the DNA end processing during the A-NHEJ beside its interaction with PNK, XRCC1 and ligase III $\alpha$ <sup>39</sup>.

### **1.2.3 Homologous Recombination (HR)**

HR occurs during the S/G2 phases when sister chromatids are available as homologous templates<sup>40</sup> (Figure 1.6). HR is fundamental in meiosis in which it allows the proper segregation of homologous chromosomes and it also favors the generation of genetic diversity. As a key early step, the trimeric complex MRN and CtBP (C-terminal binding protein) interacting protein (CtIP) process the DNA ends. Further resection occurs by the 5'-3' exonuclease activity of EXO1 or by the combined nuclease helicase activities of BLM/DNA2 (Bloom Syndrome RecQ Like Helicase/DNA replication helicase nuclease 2). The generated 3' ssDNA tails are bound by RPA, which is subsequently replaced by RAD51 thanks to the help of several mediators, such as BRCA2 (breast cancer-associated 2) and RAD54. Once it is recruited RAD51 multimerizes forming nucleofilaments on ssDNA and several RAD51 paralogs, such as XRCC2, XRCC3,

RAD51B, RAD51C and RAD51D act promoting or stabilizing RAD51 nucleofilaments. Then, RAD51 nucleoprotein filament searches for a homologous DNA sequence invading the intact dsDNA to form a heteroduplex DNA. Thus, a D-loop structure is generated in which RAD51 also interacts with the dsDNA. The invading strand is then extended by DNA polymerase  $\delta$ , PCNA and RFC<sup>41</sup>. The next step is the processing of the D loop, which can take place following different routes: the double strand break repair model (DSBR) and the synthesis-dependent strand annealing (SDSA). In the DSBR the second end of DSB is engaged to stabilize the D loop structure generating a double Holliday Junction (dHJ), which can be solved to produce crossover or non-crossover products or dissolved to exclusively generate non-crossover products. In the SDSA, the invading strand can be displaced from the D loop by DNA helicases followed by annealing with its complementary strand and DNA synthesis. SDSA mechanism preferentially occurs over DSBR during mitosis, instead DSBR is dominant throughout the meiosis in order to produce crossover products. Beyond HR, there are two more pathways to repair DSBs in a RAD51 independent fashion: NHEJ described above and the SSA (single strand annealing) pathway in which ssDNA sequences, formed during DSB processing, contain regions of homology at both sides of DSB that can be annealed and ligated<sup>42</sup>.



**Figure 1.6:** Schematic representation of homologous recombination pathway. Depending on the second end capture, three independent pathways can operate: DSBR, SDSA and BIR. (Taken from Srivastava et al.<sup>32</sup>).

### **1.3 RAD51-BRCA2: a key interaction for a successful DSB repair**

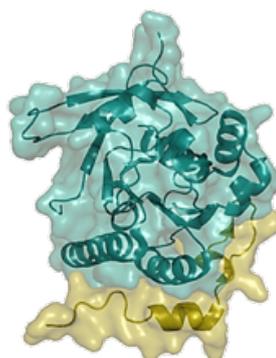
RAD51, via interaction with BRCA2, plays a key role in HR and thus in the DSBs repair. The overexpression of RAD51 has been reported in several cancers, including breast, bladder, prostate, glioblastoma, lung and leukemia. RAD51 overexpression causes high levels of HR events, leading to genomic instability<sup>43</sup>. It has also been known that germline mutations of BRCA2 gene lead to an increased susceptibility to breast, ovarian and other cancer types. The inactivation of the HR pathway affects the survival of cells leading to a 300 fold increased sensitivity to anticancer drugs. For example, primary ovarian cancers are often responsive to cisplatin treatment, which has been correlated to decreased level of HR proteins, such as BRCA1 and FANCF, or mutations in BRCA1 and BRCA2. Furthermore, platinum-based therapy resistance in ovarian cancer cells is related to the re-expression of FANCF or genetic reversion of BRCA1 or BRCA2 mutations. These correlations emphasize the role of the HR pathway not only in response to the therapy, but also as a resistance mechanism. In this context, many lines of evidence indicate that the RAD51-BRAC2 complex, which is critical for cellular survival, constitutes a crucial element in tumorigenesis and its disruption provides an important target in anticancer therapies and as a chemosensitizer.

RAD51 is a highly conserved 339 amino acids recombinase enzyme involved in the essential steps of HR. RAD51 gene expression depends on several transcriptional activators and repressors<sup>44</sup>, but it is not affected by DNA damage<sup>45,46</sup>. RAD51 protein, as well as its homologous RadA and RecA proteins, is mainly present in homo-oligomeric form in the cytosol in normal cycling cells. This suggests that the level of nuclear RAD51 is due to regulated changes in its subcellular distribution and the nuclear availability of RAD51 must be regulated to enable appropriate levels of recombination. The N terminal domain structure of RAD51 is highly conserved from prokaryotic to eukaryotic organisms: it comprises five  $\alpha$ -helices, a  $\beta$  strand

useful for protomer-protomer interactions and a typical  $\alpha/\beta$  ATPase domain with nucleotide binding Walker motifs. Since RAD51 misses a nuclear localization signal (NLS) sequence, it has to be transported into the nucleus by the interaction with other proteins. This role is mainly played by the BRCA2 protein<sup>47</sup>, even though there are some evidences which show that the nuclear transport can also occur in a BRCA2 independent way, in particular with the help of RAD51 paralog, RAD51C<sup>48</sup>.

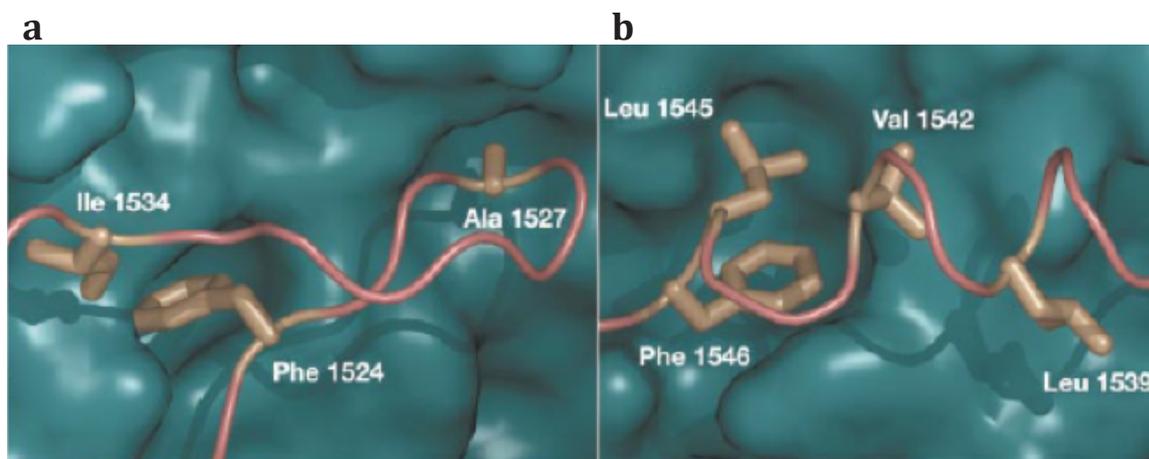
BRCA2, a 3418 amino acids protein, interacts with RAD51 through its eight repeats constituted by 35-40 amino acids (BRC repeats). BRC repeats are well conserved in different species and mutations in some of them confer sensitivity to DNA damage, suggesting their critical role for BRCA2 function. Among the eight BRCA2 motifs, BRC3 and BRC4 show the highest affinity for RAD51, instead BRC5 and BRC6 are inefficient RAD51 binders<sup>49,50</sup>.

The only published X-ray structure of a RAD51-BRC4 complex (Figure 1.7; PDB: 1N0W) shows the most prominent points of contact characterizing the interaction between the two proteins. The crystal structure was obtained by removing the first 97 amino acidic residues of RAD51 and the two proteins, RAD51 and BRC4, were purified as a fusion protein linked by a flexible peptide<sup>51</sup> ((Thr-Gly-Ser)<sub>4</sub>-Met-Gly).



**Figure 1.7:** Structure of RAD51 (C terminal region; blue) in complex with BRC4 peptide (yellow). (Taken from Pellegrini et al.<sup>51</sup>).

The binding of BRC4 to RAD51 covers the core parallel  $\beta$ -sheet structure of RAD51 with two short antiparallel strands and an amphipathic  $\alpha$ -helix (amino acids 1536-1542), which lies against the RAD51 exterior. This continuous contact between BRC4 and RAD51 is brought about by a 28 amino acid (1521-1548) segment of BRC4, which includes hydrophobic contacts as well as a number of charge-charge interactions along the interface. Two BRC4 sequences are essential for the interaction with RAD51 protein: 1524-FX<sub>1</sub>X<sub>2</sub>A-1527 and the sequence at the C terminal 1545-LFDE-1548. Phe1524 is located on the  $\beta$  strand hairpin and its aromatic ring is located in a RAD51 hydrophobic pocket formed by the side chains of residues Met158, Ile160, Ala190, Ala192, Leu203, Ala202 and Met210 (Figure 1.8a). Ala1527, in the second position of the hairpin loop, inserts its  $\beta$  carbon into a small pocket formed by the side chains of RAD51 residues Phe166, Pro168, Leu171, Leu186 and Val189 (Figure 1.8b).



**Figure 1.8:** Hydrophobic interaction at RAD51-BRC4 interface (a) apolar interactions between the  $\alpha$  helix of BRC4 (brown) and RAD51 surface (green); (b) hydrophobic interaction between the  $\beta$  hairpin of BRC4 and RAD51 surface. (Taken from Pellegrini et al.<sup>51</sup>).

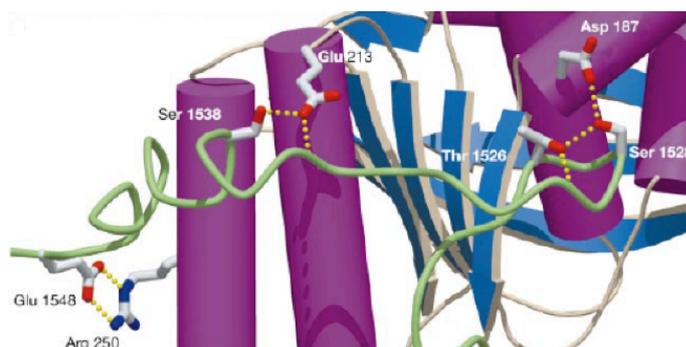
The FX<sub>1</sub>X<sub>2</sub>A sequence acts mimicking the structure of an adjacent RAD51 monomer, thus it inhibits RAD51 assembly blocking oligomerization. The mimicked RAD51 sequence 85-GGFTTATE-91 is highly conserved and it is located in the linker between the N terminal region and the catalytic core of the protein (Figure 1.9). The mutant forms RAD51 F86E and RAD51 A89E *in vitro* lose the ability to produce RAD51 filament, although they are still able to bind BRC4.

RAD51	H. sapiens	85 - G F <sup>*</sup> T T A <sup>*</sup> T E - 91
RAD51	C. griseus	85 - G F T T A T E - 91
RAD51	X. laevis	82 - G F T T A T E - 88
RAD51	D. melanogaster	82 - G F L S A R T - 88
RAD51	S. cerevisiae	143 - G F V T A A D - 149
DMC1	H. sapiens	85 - G F L T A F E - 91
RADA	P. furiosus	95 - T F M R A D E - 102
RecA	E. coli	25 - S I M R L G E - 31
BRC4	H. sapiens	1523 - G F H T A S G - 1529

**Figure 1.9:** Alignment of evolutionary conserved RAD51 sequences. Asterisks indicate two critical hydrophobic amino acid residues. (Taken from Pellegrini et al.<sup>51</sup>).

The amino acidic sequence 1545-LFDE-1548 of BRC4 is highly conserved in diverse organisms. In particular, the motif LF is mainly present in all eight human repeats as two hydrophobic residues. Phe1545 and Phe1546 provide hydrophobic interactions with a pocket between RAD51 helices A4 and A5, instead the glutamic acid forms a salt bridge with RAD51 Arg250. The interaction is enhanced by other hydrophobic contacts, such as Ile1534 in the linker region and Leu1539 and Val1542 in the  $\alpha$  helix (Figure 1.8b). Beyond the hydrophobic interaction, polar and charged bindings occur between RAD51 and BRC4 (Figure 1.10). Asp187 of RAD51 accepts a hydrogen bond from Ser1528 and electrostatically interacts with Lys1530. Finally,

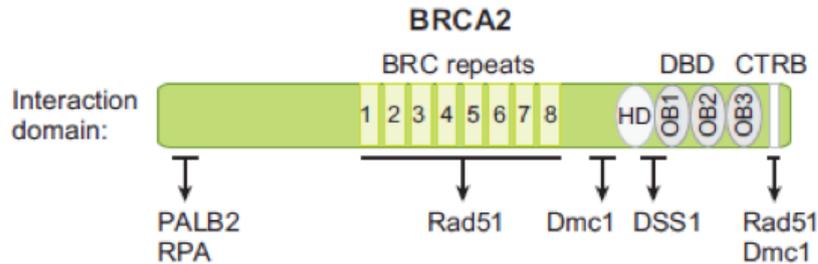
Glu213 of RAD51 receives a hydrogen bond from Ser1538 of BRC4. It has been reported that mutations of relevant contact points of the RAD51-BRC4 interface affect the formation of the complex.



**Figure 1.10:** Polar interaction between RAD51 (ribbon) and BRC4 (purple tubes). (Taken from Pellegrini et al.<sup>51</sup>).

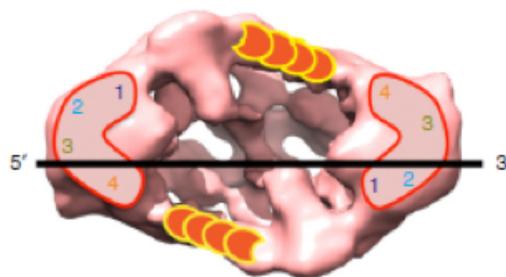
It has also been predicted that physiological modifications can decrease RAD51 affinity destabilizing the structure of BRC4. For instance, phosphorylation of Tyr1526 or Ser1538 of BRC4 abolish the polar contact with Asp187 and Glu213 of RAD51, respectively<sup>52</sup>. The determination of the atomic structure of the BRC4 repeat in complex with the catalytic core of RAD51 allowed expanding the knowledge of the protein-protein binding and identifying the key residues of the interaction. However, the molecular description of the system is limited to static snapshots of the crystal structure and it is not sufficient to coherently describe the molecular processes occurring upon binding.

A mechanism explaining how BRCA2 modulates the DNA binding selectivity of RAD51 and drives RAD51 filament formation has been recently proposed<sup>53,54</sup>. Beyond the BRC repeats, BRCA2 possesses at the C terminal, a second essential region for HR pathway modulation (Figure 1.11).



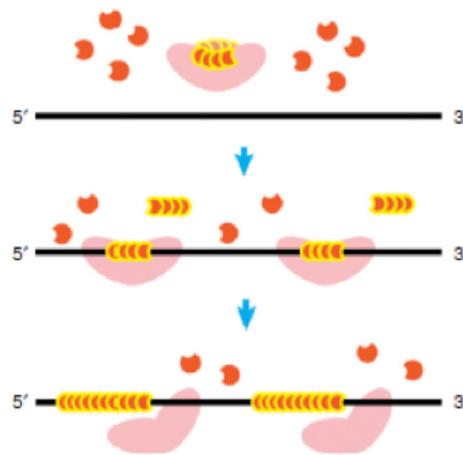
**Figure 1.11:** Functional domains in human BRCA2. (Taken from San Filippo et al.<sup>55</sup>).

BRCA2 contains a DNA Binding Domain (DBD), characterized by a  $\alpha$  helix followed by three OB folds, able to bind both ssDNA and dsDNA. BRCA2 is predominantly present in a dimeric form able to bind both ssDNA and RAD51 without dissociation. BRCA2 dimers offer mechanistic advantages because they impose the correct orientation to ssDNA and they can recruit two sets of RAD51 molecules arranged in opposite directions, one of which will be productive in ssDNA binding. The DBD binds weakly 12-20 nucleotides of ssDNA and tightly a sequence of 36 nucleotides of ssDNA, resembling the RPA DBD<sup>56</sup>. In this context, irrespective of its relative polarity, ssDNA can bind two of four DNA binding domains in one BRCA2 monomer and the other two different domains in the other BRCA2 monomer, allowing the correct orientation of ssDNA in both (Figure 1.12).



**Figure 1.12:** Structure of dimeric BRCA2 (pink) impose the correct polarity to ssDNA (black) by its DNA binding domains (1-4). ssDNA could simultaneously bind to domains 3 and 4 (OB2 and OB3) at the 5' end (left side) of one BRCA2 monomer and domains 1 and 2 ( $\alpha$ -helical domain and OB1) at the 3' end (right side) of the second monomer. Dimeric BRCA2 recruits two sets of RAD51 molecules (orange) in opposing direction. (Taken from Shahid et al.<sup>54</sup>).

Recombination is initiated at the 3' ssDNA tails generated by DSB resection. These tails enable multiple BRCA2 mediated nucleation events to originate a RAD51 nucleofilament in the 3'-5' direction. The unidirectional growth of the nucleofilament is most likely due to the presence of BRCA2 at 3' end blocking dissociation, while allowing RAD51 self-polymerization at the other end (Figure 1.13).



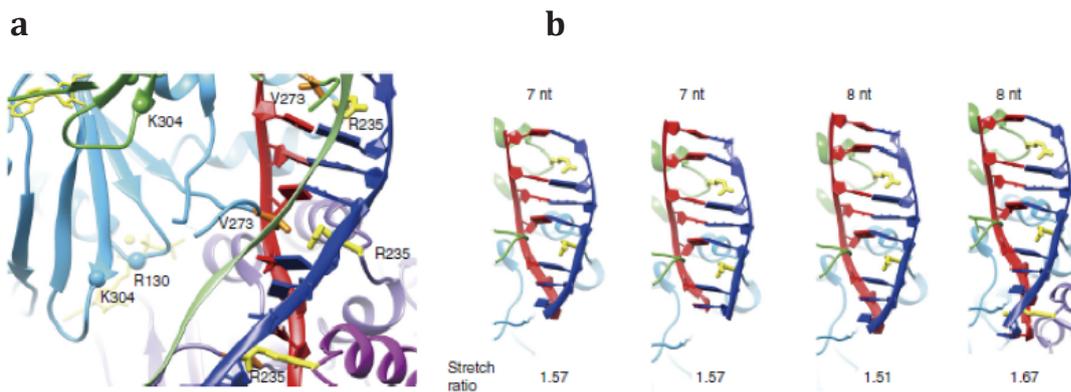
**Figure 1.13:** Predicted model for nucleoprotein filament formation and elongation mediated by BRCA2. (Taken from Shahid et al.<sup>54</sup>).

Besides the formation of RAD51 nucleofilament assembly, BRCA2 regulates the DNA binding selectivity of RAD51. It has been shown that BRC domains can apparently have two opposing effects: they can block assembly of RAD51 nucleofilament on dsDNA and stabilize the nucleofilament<sup>47,57</sup> itself, reflecting a complexity in the mechanism by which BRCA2 regulates RAD51. It has been observed that the stabilization of the nucleofilament on ssDNA by the single BRC4 repeat requires ATP hydrolysis. BRC4 acts preventing the dissociation of the ADP-RAD51 from the ssDNA, enabling the exchange of ADP with ATP within the filament and blocking ATP hydrolysis. The result is the formation and preservation of active ATP-RAD51-ssDNA nucleofilament, which is the form required for DNA pairing function. Thus,

BRC4 acts in two ways: it enhances the pairing activity of RAD51 by slowing ATP hydrolysis meanwhile preventing the assembly of dsDNA-RAD51 filaments through impairment of RAD51 nucleofilament formation on dsDNA.

Recently, the different conformations of human RAD51-DNA complex, in the presynaptic and postsynaptic complexes, have been structurally described. The structures show as RAD51 protomers form a helical assembly via three interfaces<sup>58</sup>. The first one is mediated by an ATP molecule located between two adjacent protomers. In particular, the Walker motif A stabilizes the AMP-PNP (Adenylyl-imidodiphosphate) molecules with conserved residues K133 and T134. The second interface involves the packing of a short  $\beta$  strand of the linker region against the central  $\beta$  sheet of the ATPase domain in the adjacent protomer. Finally, the third interface corresponds to an aromatic packing between Tyr54 in the N terminal region of one protomer with F195 in the ATPase domain of the adjacent protomer. The ATP mediated interface represents the responsible for the correct presynaptic and postsynaptic complexes assembly, while the other two interfaces lead to the assembly stabilization. In the presynaptic complex three consecutive protomers hold a nucleotide triplet by the interaction of RAD51 protomers residues with the phosphate backbone of the DNA (Figure 1.14a). These interactions promote the stretching of ssDNA to 1.5 times the length of B-DNA. In the postsynaptic complex, RAD51 forms mainly identical interactions with the invading strand. In particular, the presence of different RAD51-DNA configurations, which depend on the length of the micro-homology region, has been predicted. Indeed, a micro-homology sequence of eight nucleotides corresponds to the threshold for a successful pairing event between two homologous DNA regions. In this context, Arg235 plays a critical role in the stabilization of the complementary strands in the stretched conformation: under a range of eight nucleotides the DNA joint is labile

because no more than two Arg235 are engaged, instead, micro-homology regions over eight nucleotides allow to reach the extension ratio up to 1.67 involving four RAD51 protomers in which three Arg235 residues stabilize the nascent DNA joint (Figure 1.14b).



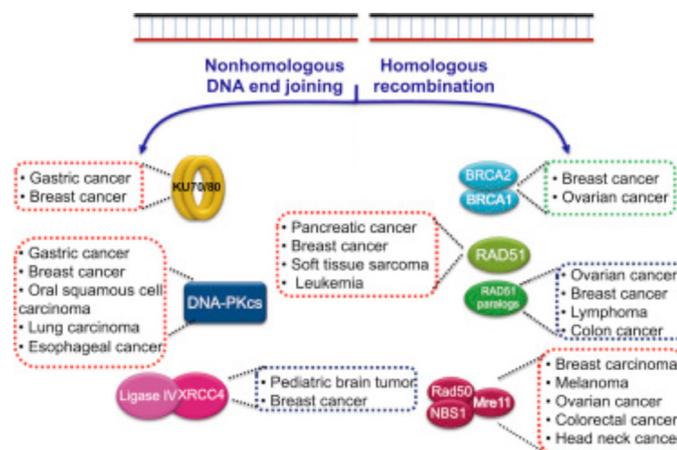
**Figure 1.14:** Model of RAD51-mediated DNA-strand exchange (a) RAD51 interacts with three DNA strands in the intermediate state. (b) Different RAD51-DNA configurations can exist, resulting in different lengths of DNA compared with B form DNA. (Taken from Xu et al.<sup>58</sup>).

Even though these recent studies have allowed a partial understanding of the DNA recombination process mediated by RAD51, the molecular mechanism that takes place starting with the recruitment and transportation of RAD51 into the nucleus until its release at the end of the process, has not been completely described yet.

### 1.4 Double strand break repair pathway: a target for cancer therapy

Impairing the DSB repair mechanism administrating drugs that either interfere with protein-protein interactions or inhibit enzyme activities both represent an approach for sensitizing cancer cells to chemo- and radio-therapy. The HR and NHEJ pathways are not mutually exclusive, so one route cannot substitute the other one. As already mentioned, germline mutations in BRCA1 and BRCA2 are responsible for the hereditary susceptibility to breast and ovarian cancers. Homologous deletion of the ATR gene is lethal for embryos as well as knockouts of the MRN complex<sup>59,60</sup>. Conversely, ATM knockout mice are viable although they show growth defects and a higher incidence of cancer<sup>61,62</sup>.

Knockout of RAD51 leads to embryonic lethality, whereas RAD51 paralogs knockout exhibit nearly normal phenotype<sup>63</sup>. Furthermore, inhibition of either Ligase IV or XRCC4 determines embryonic lethality, instead loss of Artemis or DNA-PKcs results viable, although they are associated with severe combined immunodeficiency (SCID)<sup>64</sup>. Defects in NHEJ and HR proteins can predispose the cells to cancer. In particular, mutations in DSB repair pathways can determine genomic instability and higher mutation rates, which may provide an advantage to cancer cells. Polymorphisms, upregulation or hyperactivation of DSB repair also support tumorigenesis (Figure 1.15).



**Figure 1.15:** HR and NHEJ proteins defects in various cancers. (Taken from Srivastava et al.<sup>32</sup>).

Inhibitors of HR are the most investigated DNA repair inhibitors. Mirin was the first identified inhibitor of the nuclease activity of the MRN complex<sup>65</sup>. It acts by killing BRCA2 deficient cells, but since it has to be administered in micromolar concentration, it can enhance the risk of undesired off-target effects acting upstream of NHEJ. Another strategy to impair HR is targeting the ATR-CHK1 and ATM-CHK2 pathways. KU-55933 represents the first selective inhibitor of ATM. It competes with the ATP binding site of ATM, inhibiting its catalytic activity<sup>66</sup>. VE-821 and NU6027 have been identified as ATR inhibitors and they are especially toxic in cell deficient in p53. VE-821 determines radio- and chemo-sensitization in pancreatic cells, whereas NU6027 is a pyrimidine analogue of adenosine triphosphate (ATP) molecule and it can act as a CDK inhibitor or it can confer cisplatin cytotoxicity independently of CDK inhibition. Small molecule inhibitors, ML216 for RecQ helicase BLM and NSC19630 for WRN have been identified. ML216 inhibits the DNA-dependent ATPase activity of BLM, but acts in a non-competitive manner with respect to ATP. NSC 19630 affects DNA unwinding by interfering with the activity of WRN helicase and in combination with topotecan it acts synergistically to inhibit cell proliferation<sup>67</sup>.

Among NHEJ proteins, many attempts have been made to target DNA PKcs. Wortmannin and LY294002 inhibit DNA PKcs and possess similar properties, indeed both sensitize cancer cells to IR treatment. Unfortunately, they show *in vivo* toxicity and off-target effects. NU7026 and NU7441 are analogs of LY294002. Both have been reported to be very selective and potent DNA-PK inhibitors. However, their poor solubility and low bioavailability affect their administration *in vivo*. As far as KU proteins are concerned, no inhibitors have been published yet.

Another strategy to impair the NHEJ pathway is targeting the final step of the process mediated

by Ligase IV. In this context, SCR7 inhibits NHEJ *in vitro* and *in vivo* and its mechanism of action is based on the accumulation of unrepaired DSBs. Importantly, it can potentiate the effects of chemo- and radio-therapy when coadministrated<sup>68</sup>.

Finally, many efforts have been conducted to validate inhibition of DNA end processing as a therapeutic target. Encouraging data support the possibility of impairing the activity of human polynucleotide kinase/phosphatase (hPNKP), the enzyme responsible for phosphorylation of 5' hydroxyl termini and dephosphorylation of 3' phosphate termini, which are often altered after IR treatment. Interestingly, the radiosensitization of breast and lung cells<sup>69</sup> has been observed after treatment with this inhibitor.

*State of the art strategies for targeting RAD51 using small molecules*

In addition to the role in DSB repair, RAD51 acts stabilizing the stalled replication DNA fork<sup>70,71</sup> to prevent nucleolytic degradation of nascent DNA. RAD51 overexpression in cancer cells enhances HR efficiency and leads to resistance after chemo- and radio-therapies treatments. RAD51 overexpression and the increase in foci formation could be due to the rapidly dividing cancer cells resulting in increased DNA replication stress. Interestingly, a major RAD51 expression is not a consequence of gene amplification but it depends on increased gene transcription levels. *In vivo* studies have shown that RAD51 causes the formation of DNA-RNA hybrid structures, which represent a danger for genome stability<sup>72</sup>. Furthermore, RAD51 inhibition is correlated with a reduction of breast cancer migration suggesting that it can also contribute to metastases<sup>73</sup>. Taken together, these data show that beyond its canonical role as DNA repair protein, RAD51 can determine genotoxic effects and can be implicated in DNA damage and promotion of genomic alterations.

Recently, many efforts have been made to target RAD51 with small molecules as an anticancer therapy. Reduction of RAD51 expression levels helps to increase the sensitivity of cancer cells to the treatments with relatively smaller effects in noncancerous human cell lines<sup>74</sup>. Based on the knowledge of the processes in which RAD51 is involved, small molecules have been developed for targeting RAD51. It is possible to distinguish three classes of inhibitor compounds depending on their mechanism of action: compounds that inhibit RAD51-ssDNA filament formation, compounds that stimulate RAD51-ssDNA binding and compounds that inhibit RAD51 D-loop formation<sup>75</sup>.

Compounds that inhibit RAD51-ssDNA nucleoprotein filament formation have been investigated with the purpose of affecting the capacity of RAD51 to bind ssDNA and consequently impairing both DSB repair and stalled replication forks. 4,4'-Diisothiocyanostilbene-2,2'-disulfonic acid (DIDS) was originally used as an inhibitor of membrane transporters and ionic channels, but it functions also as a RAD51-mediated D loop formation inhibitor with an  $IC_{50}$  of 1-10  $\mu$ M. It acts by binding at or near the RAD51's ssDNA binding site. However, further effects of DIDS were underestimated, in fact, it induces elevated cell toxicity. Another small compound, B02 interacts directly with RAD51 ( $K_d = 5.6 \mu$ M) preventing the formation of nucleoprotein filaments. B02 inhibits HR and increases cancer sensitivity to several DNA damaging agents, such as cisplatin, mytomicin C, doxorubicin and etoposide<sup>76</sup>. A similar effect has been observed using the compound Chicago Sky Blue (CSB,  $IC_{50} = 0.4 \mu$ M) and its analog compound RI-1. Unfortunately, the chloromaleimide group of this class of compounds covalently binds the thiol group on the C319 residue of RAD51 leading to off-target effects. A SAR study has allowed discovering another analog, RI-2, which reversibly binds RAD51 but displays a six fold decrease in potency.

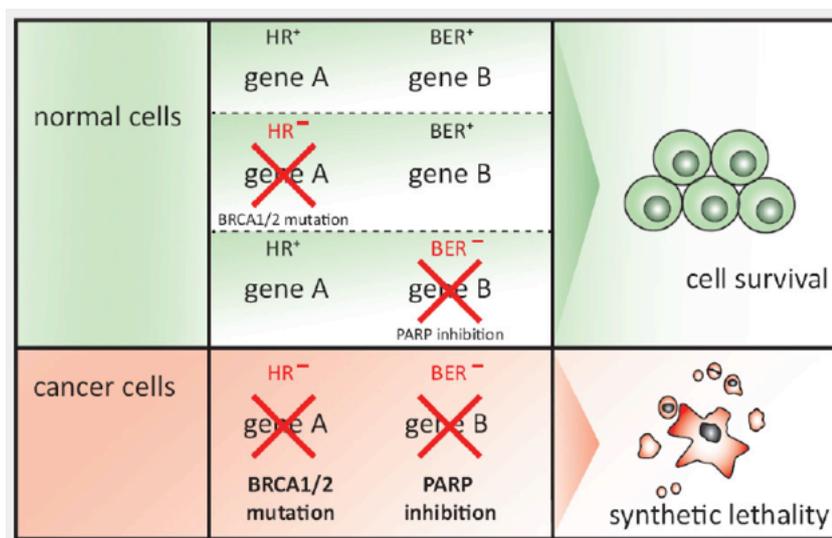
A different strategy is the stabilization of RAD51-ssDNA nucleofilaments by stimulating RAD51-ssDNA binding. RS-1 allows RAD51 to bind ssDNA tightly by locking RAD51 in the active conformation, therefore without affecting active site ATP hydrolysis. It has been shown that RS-1 acts mainly in malignant cells, which display increased RAD51 expression and its activity is dependent on RAD51 and RAD54B/RAD54L translocase<sup>77</sup>.

Another class of RAD51 inhibitors acts by impairing the strand exchange activity of RAD51 while keeping its ability to bind ssDNA. For instance, halenaquinone decreases the IR induced RAD51 foci and might act by blocking the formation of ternary complexes and rendering the presynaptic complex catalytically inactive<sup>78</sup>. Budke *et al.* successfully developed other compounds with similar effect, such as RI(dI)-1 and its analogs<sup>75</sup>. Among these class of compounds, RI(dI)-2 represents the best optimized RAD51 D-loop inhibitors. It sensitizes cancer cells to IR and can also inhibit SSA.

#### **1.4.1 Impairment of the double strand break repair pathway and synthetic lethality**

Impairment of the DSB repair mechanism can also lead to sensitization of cancer cells to therapeutic DNA damaging agents. Actually, PARP inhibitors (PARPi) are currently administrated to women with germline mutations in the BRCA2 gene for treatment of breast and ovarian cancers. Indeed, these mutations make cancer cells more sensitive to PARPi. Most PARPi act trapping PARP1 at the site of damage, preventing both autoPARylation and PARP1 release from the DNA. The PARP1 catalytic cycle is therefore blocked with consequent DNA replication fork stalling. Thus, in BRCA-mutant tumors, in which the HR pathway is compromised, the administration of PARPi leads to cell death. Interestingly, the identification of PARPi for treatment of BRCA-mutant related cancers represents the first proof of principle study validating synthetic lethality as a cancer therapy (Figure 1.16). In 2014, olaparib (Lynparza™),

developed by AstraZeneca, became the first PARPi to be approved by FDA to treat advanced ovarian cancer associated with defective BRCA genes. Clinical trials are currently studying this drug to treat breast and pancreatic BRCA-mutant tumors. The synthetic lethality approach is based on the assumption that two genes are synthetically lethal, if loss or silencing of one of them is compatible with cell viability, while simultaneous loss of both genes leads to cell death. Synthetic lethality is here obtained by administering PARPi and simultaneously impairing the HR for DSB repair (BRCA2 mutation), therefore inhibiting cancer cells attempts to restore the therapeutic DNA damage induced.



**Figure 1.16:** Synthetic lethality. If genes essential for a repair pathway (i.e. gene A) are inactivated in normal cells, alternative pathways (i.e. gene B) can function to respond to the DNA damage. Conversely cancer cells which already present mutations in genes essential for a DNA repair pathway cannot respond to the DNA damage if mutation or silencing occurs also in the alternative DDR pathway. (Taken from Huhn et al.<sup>79</sup>).

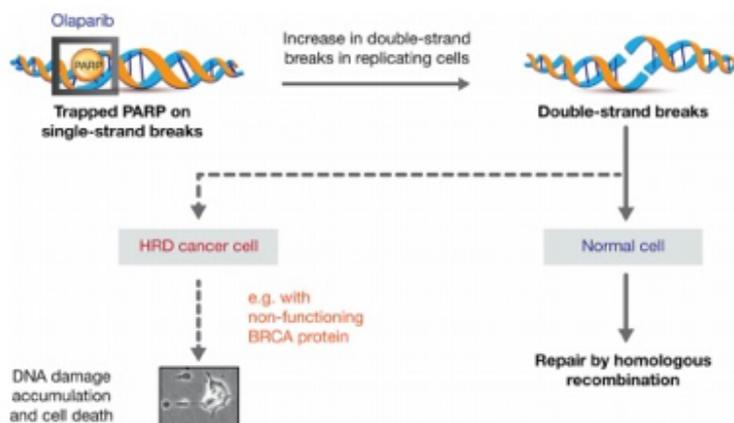
The identification of a small molecule, able to disrupt RAD51-BRCA2 interaction, could in principle pharmacologically induce synthetic lethality by inhibiting the HR pathway in analogy with BRCA2 mutations. Inhibition of RAD51-BRCA2 interaction can be effective in treating patients without BRCA2 mutations and concurrently broaden the potential applications of PARPi in anticancer therapy.

Few molecules have so far been described that are able to disrupt the RAD51-BRCA2 interaction. Among these small molecules, IBR2 has been reported to bind RAD51 and to inhibit *in vitro* and *in vivo* the interaction with BRCA2 ( $IC_{50} = 10 - 20 \mu\text{M}$ ). Unfortunately, IBR2 suffers of significant stability and low bioactivity issues<sup>80</sup>. New studies for the identification of molecules able to prevent the RAD51-BRCA2 interaction are therefore necessary to overcome these IBR2 issues and exploit the inhibition of the RAD51-BRCA2 interaction as an effective tool to validate and boost up the synthetic lethality approach.

## **Goal of the research project**

Human RAD51 (hRAD51), via interaction with BRCA2, plays a key role in homologous recombination (HR) and thus in the repair of DNA double-strand breaks (DSBs). The overexpression of hRAD51 has been reported in a wide range of cancers including breast, bladder, glioblastoma, prostate, pancreas, lung and leukaemia<sup>81-84</sup>. hRAD51 has been also correlated with resistance to DNA damaging agents, such as temozolomide, and to radiotherapy<sup>43</sup>. In addition, hRAD51 down-regulation has been shown to increase cell chemo- or radio-sensitivity both *in vitro* and *in vivo*<sup>85</sup>.

Oncology patients, carrying BRCA2 mutations, show enhanced sensitivity to some chemotherapeutic agents and radiotherapy<sup>86</sup>. Importantly, BRCA2-defective oncology patients show an increased susceptibility to poly ADP ribose polymerase inhibitors (PARPi), such as olaparib, which has been approved by FDA for treatment of BRCA-mutant tumors<sup>87</sup>. Olaparib acts by trapping PARP onto single-strand breaks, leading to the formation of DSB during the replication process (Figure 1.17). In tumors with HR deficiency due to BRCA1/2 mutations, this mechanism of action determines genomic instability and consequently cell death. This approach exploits the theory of synthetic lethality, which is a type of genetic interaction where the simultaneous occurrence of two genetic inactivations results in cellular death, but deletion of just one of them does not affect cell viability. In fact, targeting the synthetic lethal partner of a mutated gene can selectively kill cancer cells sparing normal cells.



**Figure 1.17:** Mechanism of action of olaparib in homologous recombination deficiency cancer cells (Taken from O’Connor et al.<sup>9</sup>).

Efficacy of PARPi is lost in tumors with active HR and resistance to the treatment with PARPi is often associated to an increased HR activity, due to either hRAD51 overexpression<sup>88</sup> or restoration of non-mutated BRCA2<sup>89</sup>. Based on the above assumptions, our hypothesis is that disrupting hRAD51-BRCA2 interaction could affect the survival of cancer cells, making them sensitive to the treatment with PARPi. The long term aim of this research project is the identification of a small molecule, acting as an inhibitor of protein-protein hRAD51-BRCA2 interaction, able to trigger a full synthetic lethality process in patients without BRCA2 mutations. This compound will be eventually administered to these individuals in combination with olaparib, mimicking the synthetic lethality induced by olaparib in BRCA2 defective oncology patients. The rationale of this approach lies on the evidence that spontaneous DNA alterations occurs more frequently in tumor cells rather than in normal cells because of their much higher genetic instability; therefore, cancer cells heavily depend on DNA repair mechanisms. Normal cells, which experience much less DNA damage, may be affected by DNA repair inhibitors at much lesser extent relative to cancer cells. Simultaneous administration of olaparib with RAD51-BRCA2 inhibitors should preferentially affect cancer cells hence supporting our

approach from the selectivity standpoint.

In this context, we propose a multidisciplinary project (Figure 1.18) to validate the protein-protein interaction hRAD51-BRCA2 as a target for anticancer drug discovery in combination therapies with PARPi. To achieve this aim, the first step is the identification of small molecules able to disrupt the hRAD51-BRCA2 interaction through a virtual screening campaign on the only available X-ray crystal structure of the hRAD51-BRC4 complex (PDB: 1N0W). The identified compounds are synthesized and tested through biochemical and biophysical assays for their ability to interfere with hRAD51-BRC4 interaction. Their affinity ( $K_d$ ) and kinetic ( $k_{on}$ ,  $k_{off}$ ) parameters for the binding to hRAD51 are also determined. In parallel, cell-based experiments are carried out to investigate the anticancer activity of the novel compounds alone or in combination with PARPi. Pancreatic adenocarcinoma cells are used as the model for the cell-culture experiments because of its clinical relevance. For this tumor form, there is a well-studied cell line lacking functional BRCA2 (the Capan-1 cell line). The BRCA2-proficient cell lines for the study are BxPC3 and Panc-1. BxPC3 cells offer an appropriate comparison to Capan-1 because they match the Capan-1 line in terms of tumor type and differentiation state. The effects of the most promising inhibitors on BRCA2 deficient cell lines derived from human breast and ovarian cancers are also studied. The obtained results will contribute to the identification of hit and lead compounds, which will be further improved through the information gained from the X-ray crystal structure of hRAD51-inhibitors complex. Virtual screenings and the synthesis of the compounds are carried out in Italian Institute of Technology (IIT), while cell-based experiments are performed by our collaborators at the University of Bologna. Biochemical assays are carried out by our collaborators in Perugia. Support for structural sample preparation and data collection is provided by the National Center for Protein

Science Shanghai (NCPSS).

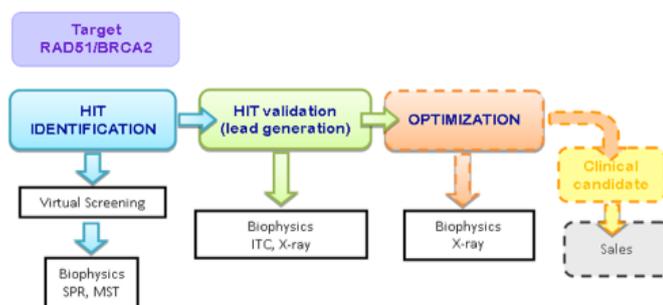


Figure 1.18 : Overview of the research project.

Within this project, the goal of my PhD research activity is the characterization, through biophysical techniques, of the hRAD51-BRCA2 interaction and of the hRAD51-new inhibitor molecules interactions. My research work is critical to confirm the ability of the newly identified compounds to specifically target the protein-protein hRAD51-BRCA2 interaction, which is viable for cellular survival. Moreover, the biophysical characterizations provide strong support to the identification of new molecules able to interfere with the hRAD51-BRCA2 interaction.

The aims of my PhD project are:

- **Optimization of a protocol for the overexpression and purification of full-length hRAD51 in bacteria cells;**
- **Biophysical characterization of hRAD51-hit molecules interaction through surface plasmon resonance (SPR), Nuclear Magnetic Resonance (NMR) spectroscopy and Microscale Thermophoresis (MST)**
- **X-ray structural characterization of hRAD51-BRCA2**

---

---

Chapter two

# Materials and Methods

---

---

## 2.1 Materials

### 2.1.1 Growth media

Bacterial cultures were grown in the following media: Luria-Bertani (LB) medium, M9 medium, SOC medium, ZYM-5052 and Terrific Broth (TB)-5052 auto-induction media.

#### *Luria-Bertani medium*

LB medium is the most widely used broth for bacterial cells growth. Its composition is reported in the following Table 2.1:

Reagents	LB medium	LB agar
	Weight (g)	Weight (g)
NaCl	10	10
Yeast extract	5	5
Tryptone	10	10
Agar		15
<b>Final Volume</b>	1 L	1L

**Table 2.1:** Composition of LB broth and LB agar.

LB medium and LB agar were sterilized in autoclave at 121 °C for 20 minutes.

#### *M9 medium*

M9 medium is defined as a minimal broth. In fact, it contains only salts and nitrogen, so it is generally supplemented with a carbon source (e.g. glucose), amino acids and vitamins as needed. The composition is reported in Table 2.2:

Reagents	M9 medium
	Weight (g)
Na <sub>2</sub> HPO <sub>4</sub>	33.9
KH <sub>2</sub> PO <sub>4</sub>	15
NH <sub>4</sub> Cl	5
NaCl	2.5
<b>Final Volume</b>	1 L

**Table 2.2:** Composition of M9 medium.

M9 medium sterilized in autoclave at 121 °C for 20 minutes was supplemented with sterilized solutions of glucose 0.4% (w/v), MgSO<sub>4</sub> 1 mM and CaCl<sub>2</sub> 3 mM.

*Super Optimal Catabolite repression (SOC) medium*

SOC medium (Table 2.3) is defined as a rich bacterial growth medium that is used to maximize the transformation efficiency of *E. coli* in the final step of bacterial cell transformation.

<b>SOC medium</b>	
<b>Reagents</b>	<b>Weight (g)</b>
Tryptone	20
Yeast extract	5
NaCl	0.58
KCl	0.19
MgCl <sub>2</sub>	0.95
MgSO <sub>4</sub>	1.20
Glucose	3.6
<b>Final Volume</b>	<b>1L</b>

**Table 2.3:** Composition of SOC-medium.

SOC medium was sterilized in autoclave at 121 °C for 20 minutes.

*Auto-induction media*

Auto-induction media, like ZYM-5052 and TB-5052 are very useful for protein production and usually result in significantly higher protein yields compared to conventional media. In strains such as BL21, the T7 RNA polymerase is under control of the lacUV5 promoter, inducible with lactose. The addition of glucose, glycerol and  $\alpha$ -lactose (5052 solution) into the medium allows to obtain an auto-induction medium. In fact, bacteria use glucose as primarily source of carbon which, at the same time, prevents the activation of *lac* operon by lactose, promoting the catabolite repression. In this condition, the production of the heterologous protein is not allowed. Hence, exhaustion of the glycerol causes two effects: bacteria will use glycerol as carbohydrate source and lactose will be internalized into the cells to promote the activation of *lac* operon with consequent expression of the recombinant protein.

*ZYM-5052 medium*

In ZYM-5052 auto-induction medium potassium and sodium phosphate avoid the pH drop of the medium (6.75 at 25 °C). In Table 2.4 the medium composition is reported:

Reagents	ZY	5052 (50X) solution
	Weight (g)	Weight (g)
Tryptone	9.4	
Yeast extract	4.7	
Glycerol	4.7	25
Glucose		2.5
$\alpha$ -lactose		10
<b>Final Volume</b>	1L	100 mL

**Table 2.4:** Composition of ZYM-5052 medium.

ZY and 5052 (50X) solutions, sterilized in autoclave at 121 °C for 20 minutes, were supplemented with sterilized M solution ( $\text{Na}_2\text{HPO}_4$  25 mM,  $\text{KH}_2\text{PO}_4$  25 mM,  $\text{NH}_4\text{Cl}$  50 mM,  $\text{Na}_2\text{SO}_4$  5 mM) and  $\text{MgSO}_4$  2 mM to prepare the ZYM-5052 medium.

*Terrific Broth-5052*

Terrific Broth (TB) is an enriched medium, which helps the recombinant *E. coli* strains of maintaining an extended growth phase. The large amount of yeast and tryptone provides additional nutrients and growth factors required for enhanced growth of cells. The potassium phosphates are present to prevent a drop in the pH ( $7.2 \pm 0.2$  at 25 °C) of the medium during bacterial growth. The medium composition is reported in the following Table 2.5:

Reagents	Terrific broth	5052 (50X)
	Weight (g)	Weight (g)
Tryptone	12	
Yeast extract	24	
Glycerol	4 mL	25
$\text{KH}_2\text{PO}_4$	2.2	
$\text{K}_2\text{HPO}_4$	9.4	
Glucose		2.5
$\alpha$ -lactose		10
<b>Final Volume</b>	1L	100 mL

**Table 2.5:** Composition of TB-5052 medium.

TB and 5052 (50X) solutions were sterilized in autoclave at 121 °C for 20 minutes. TB-5052, obtained adding 20 mL of 5052 (50X) solution to 980 mL of TB, was supplemented with a sterilized solution of  $\text{MgSO}_4$  (2 mM final concentration).

### **2.1.2 Solutions and reagents**

**SDS-PAGE sample loading buffer:** 4X: 0.25 M Tris-HCl pH 6.8, 8% SDS, 0.3M DTT, 30% Glycerol, 0.02% Bromphenol blue;

**MOPS SDS running buffer:** 50 mM MOPS, 50 mM Tris Base, 0.1% SDS, 1 mM EDTA, pH 7.7; (Invitrogen);

**Staining solution:** (40% EtOH, 10% acetic acid, 0.05% w/v comassie blue G-250);

**Destaining solution:** (25% ethanol, 8% acetic acid);

**Transfer buffer:** NuPAGE transfer buffer (Invitrogen);

**TBS-T buffer:** 137 mM NaCl, 2.7 mM KCl, 19 mM Tris base, 0.1% (v/v) Tween20;

**Blocking solution:** TBS-T buffer supplemented with 5% (w/v) milk;

**DNA sample loading buffer:** Gel Loading Dye, Blue (6X) (New England Biolabs);

**Tris-Borate-EDTA (TBE) buffer:** 89 mM TrisBase, 89 mM Boric acid, 2 mM EDTA;

**Agarose gel:** Tris-Borate-EDTA (TBE) buffer: 89 mM TrisBase, 89 mM Boric acid, 2 mM EDTA supplemented with (w/v) % agarose depending on the desired final gel concentration.

### **2.1.3 Peptide**

**BRC4:** BRC4 peptide was purchased from Thermo Scientific, Invitrogen.

### **2.1.4 Antibodies**

**Anti-RAD51:** polyclonal antibody produced in mouse, purified immunoglobulin (SIGMA-ALDRICH);

**Anti-Mouse + HRP:** Rabbit Anti-Mouse IgG H&L (HRP) (Abcam).

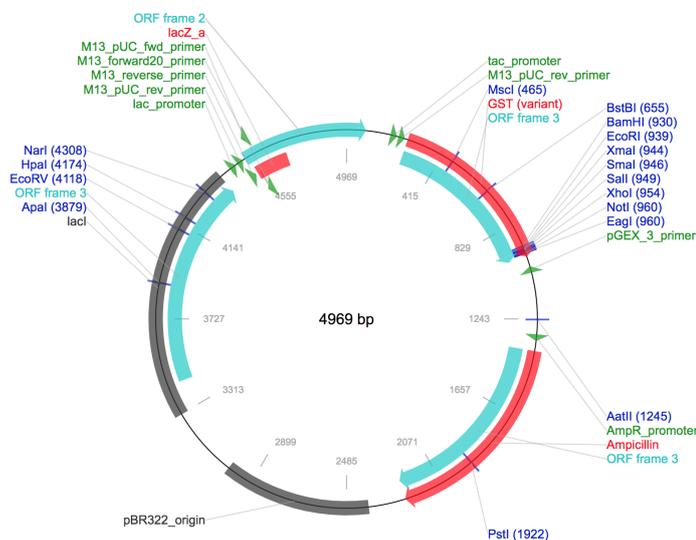
## 2.2 Molecular Biology

### 2.2.1 Plasmids for recombinant protein expression

#### *pGEX-4T-1 plasmid*

pGEX-4T-1 plasmid (GE Healthcare) is designed for high-level intracellular expression of genes or gene fragments. Expression is inducible in *E. coli* by Isopropyl  $\beta$ -D-1-thiogalactopyranoside (IPTG) addition. It allows to produce tagged proteins with a GST moiety at the amino terminus or at the carboxyl terminus (Figure 2.1). The gene coding for  $\beta$ -lactamase is present in the plasmid's DNA sequence conferring ampicillin resistance to cells transformed with this plasmid.

hRAD51 gene sequence was cloned between BamHI and SalI restrictions sites to produce a recombinant protein with a GST tag at the N terminal region.

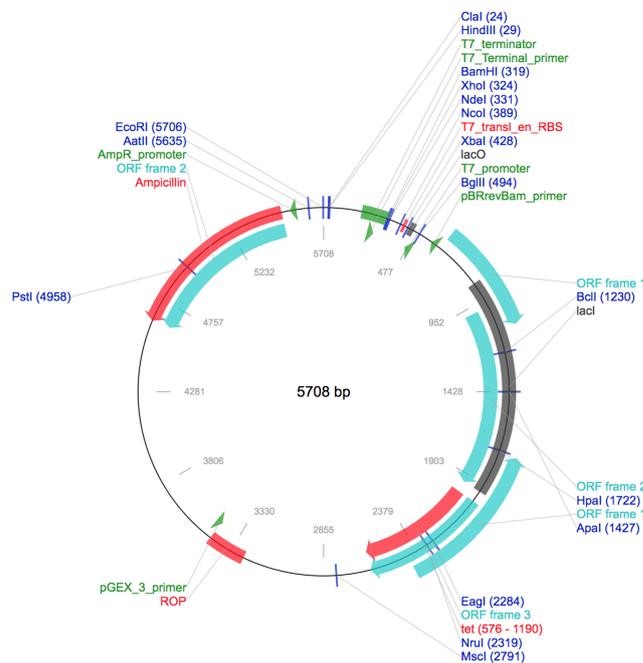


**Figure 2.1:** *pGEX-4T-1* plasmid map. (Taken from <https://www.addgene.org/vector-database/2876/>).

*pET15b plasmid*

pET15b plasmid (Novagen) is a medium copy number plasmid characterized by the presence of a T7 promoter for the expression of heterologous protein by induction with IPTG. pET15b carries a N-terminal His-tag sequence followed by a thrombin site and three cloning sites (Figure 2.2). The gene coding for  $\beta$ -lactamase present in the plasmid's DNA sequence confers ampicillin resistance to cells transformed with this plasmid.

hRAD51 and hRAD51 F86E gene sequences were subcloned between BamHI and Bpu1102I restriction sites in order to overexpress the protein of interest with a hexahistidine moiety at the N terminal region.

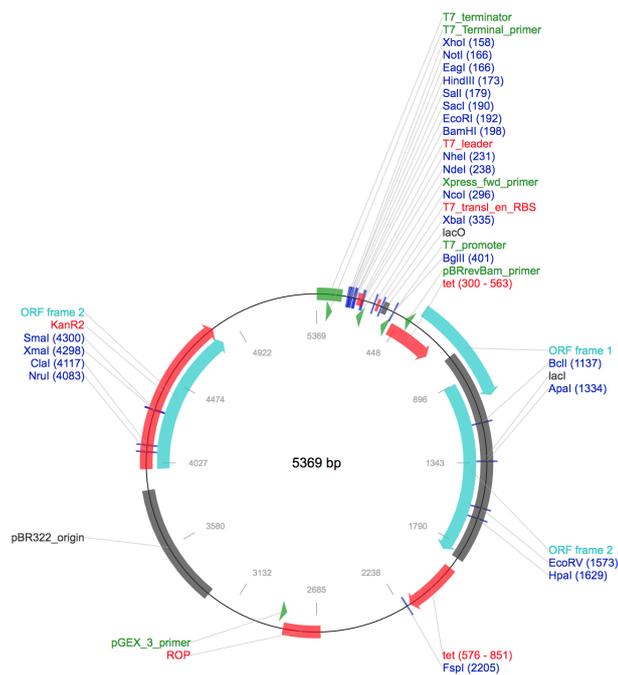


**Figure 2.2:** *pET15b plasmid map.* (Taken from: <https://www.addgene.org/vector-database/2543/>).

*pET28a plasmid*

pET28a plasmid (Novagen) is a medium copy number plasmid. The expression of recombinant protein, by IPTG induction, is regulated by the T7 promoter. pET28a carries an N terminal His-Tag-thrombin-T7-Tag configuration plus a C terminal His-Tag sequence (Figure 2.3). Bacterial cells transformed with this plasmid can be selected in media containing kanamycin.

Gene sequence coding for BRC4 peptide was subcloned between NcoI and Sall restriction sites in order to express the heterologous protein with a hexahistidine tag at the N terminal region.



**Figure 2.3:** *pET28a plasmid map.* (Taken from: <https://www.addgene.org/vector-database/2565/>).

### 2.2.2 Competent cells

#### *BL21(DE3)pLysS cells*

*E. coli* BL21(DE3)pLysS (Promega) strains are lysogenic for  $\lambda$ -DE3 that contains the T7 bacteriophage gene I, encoding T7 RNA polymerase under the control of the lac UV5 promoter. BL21(DE3)pLysS also contains a plasmid, pLysS, which carries the gene encoding T7 lysozyme. T7 lysozyme lowers the background expression level of the target genes under the control of the T7 promoter but does not interfere with the level of expression achievable after IPTG induction. The genotype is: F<sup>-</sup>, ompT, hsdSB (rB<sup>-</sup>, mB<sup>-</sup>), dcm, gal,  $\lambda$ (DE3), pLysS, Cmr.

#### *Rosetta2(DE3)pLysS cells*

*E. coli* Rosetta™ 2 host strains (Novagen) are BL21 derivatives designed to enhance the expression of eukaryotic proteins that contain codons rarely used in *E. coli*. The main difference with BL21(DE)pLysS cells is due to the presence of tRNAs for 7 rare codones (AGA, AGG, AUA, CUA, GGA, CCC, and CGG) on a compatible chloramphenicol-resistant plasmid. The tRNA genes are driven by their native promoters. The genotype is: F<sup>-</sup> ompT hsdSB(rB<sup>-</sup> mB<sup>-</sup>) gal dcm (DE3) pLysSRARE2 (CamR).

#### *XL1-Blue cells*

*E. coli* XL1-Blue cells (Agilent Technologies) are endonuclease (endA) deficient cells, to improve the quality of mini-prep DNA, and are recombination (recA) deficient to promote insert stability. A hsdR mutation prevents the cleavage of cloned DNA by the EcoK endonuclease system. The lacIq Z $\Delta$ M15 gene on the F' episome allows blue-white color screening. The genotype is: recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac [F' proAB lacIq Z $\Delta$ M15 Tn10 (Tetr)].

### **2.2.3 XL1-Blue, BL21(DE3)pLysS and Rosetta2(DE3)pLysS cells transformation**

An aliquot of 50  $\mu$ L of competent cells (XL1-Blue, BL21(DE3)pLysS and Rosetta2(DE3)pLysS cells) were thawed on ice and transformed with about 25 ng of DNA plasmid. After plasmid DNA incubation for 30 minutes on ice, cells were subjected to heat shock at 42 °C for 45 seconds and successively 2 minutes on ice. Afterwards, cells were grown in SOC medium at 37 °C for 45 minutes under shaking (200 rpm) and then plated on agar plates supplemented with the appropriate antibiotic (ampicillin 100  $\mu$ g/mL and/or kanamycin 50  $\mu$ g/mL) for the selection of transformed cells.

### **2.2.4 Plasmid DNA purification**

A single colony of XL-1 Blue cells transformed with the plasmid of interest was grown at 37 °C in an appropriate volume (10 mL for mini-prep or 150 mL for maxi-prep) of LB medium supplemented with the appropriate antibiotic. The saturated bacterial cultures were centrifuged at 6000 rpm for 10 minutes. Afterwards, plasmid DNA purification was carried out using QIAprep Spin Kit (Qiagen) and Plasmid Maxi Kit (Qiagen) for mini-prep and maxi-prep, respectively, following the manufacturer's instructions.

### **2.2.5 hRAD51 subcloning into pET15b vector**

Two different primers, purchased from Metabion (Simmelweisstrasse, Germany) were designed in order to subclone the gene coding for hRAD51 into a pET15b vector:

- RAD51 BamHI Fw: 5' CGTGGATCCGATGGCAATGCAGA 3'

- RAD51 Bpu1102I Rev: 5' GCTCGGCTCAGCTCAGTCTTTG 3'

hRAD51 sequence was amplified by polymerase chain reaction (PCR). PCR reaction was carried out using the high fidelity *Pfu* polymerase for 33 cycles of 95 °C for 2 minutes → 95 °C for 30 seconds → 66 °C for 30 seconds → 68 °C for 30 seconds followed by 68 °C for 7 minutes

at 4 °C until electrophoresis.

The PCR product was purified through QIAquick PCR purification kit as described by manufacturer's instructions. Afterwards, the PCR product and the pET15b plasmid were double digested by BamHI and Bpu1102I (NewEngland BioLabs) upon incubation at 37 °C for 1.5 h. The digested products were run in 1% agarose gel. Bands corresponding to double digested hRAD51 and pET15b plasmid were extracted from the gel and purified using QIAquick Gel Extraction Kit following manufacturer's instructions. Finally, purified hRAD51 and pET15b DNA sequences were ligated by adding T4 DNA ligase (Invitrogen) and incubating for 25 minutes and room temperature. The new construct pET15b-His-hRAD51 was used to transform XL1-Blue cells through heat-shock method and then plated on agar plates containing ampicillin (100 µg/mL). After extraction of the plasmid using the QIAprep Spin Kit (Qiagen), the plasmid was sequenced. Plasmid with the correct sequence was transformed in Rosetta2(DE3)pLysS cells to express and produce the recombinant protein hRAD51 with a hexahistidine tail at the N-terminus.

### **2.2.6 Site-directed mutagenesis**

Primers to introduce the F86E mutation in the vector pET15b-His-hRAD51 were purchased from Metabion (Simmelweisstrasse, Germany) The QuikChange II Site-Directed Mutagenesis kit (Agilent Technologies) using the polymerase chain reaction (PCR) was used to introduce the mutation.

- Primer RAD51 F86E Fw 5' CAGCTAAATTAGTTCCAATGGGTGAAACCACTGCAACTGAATTCC 3'

- Primer RAD51 F86E Rev 5' GGAATTCAGTTGCAGTGGTTTCACCCATTGGAACATAATTTAGCTG 3'

Mutated plasmid was transformed in *E. coli* XL1-Blue competent cells through heat-shock method and then plated on agar plates containing ampicillin (100 µg/mL). After extraction of

the plasmid using the QIAprep Spin Kit (Qiagen), the plasmid was sequenced. Plasmid with the correct mutation were transformed in Rosetta2(DE3)pLysS cells to express and produce the recombinant protein hRAD51 F86E with a hexahistidine tail at the N-terminus.

## **2.3 Biochemical methods**

### **2.3.1 Agarose gel electrophoresis**

Agarose gel electrophoresis promotes the separation of biomolecules, such as nucleic acids or protein, in an agarose matrix applying an electric field. Agarose gel electrophoresis was performed using 1% (w/v) agarose gel prepared as described in section 2.1.2.

Before loading DNA samples into the wells, an appropriate volume of gel loading dye was added to the samples. The running of DNA samples was carried out applying a constant voltage of 120 V for 50 minutes using the horizontal electrophoresis apparatus (Mini-Sub® Cell GT Systems; Biorad) in TBE buffer supplemented with SYBR Safe DNA Gel Stain (Invitrogen).

### **2.3.2 Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE)**

SDS-PAGE promotes the separation on polyacrylamide gels in the presence of SDS of proteins according to their molecular weights applying an electric field. SDS-PAGE was performed using precast polyacrylamide gels (NuPAGE 4-12% BisTris-Gel, Invitrogen) run in XCell SureLock™ Mini-Cell Electrophoresis System (Invitrogen).

Cell lysates or protein samples, after addition of an appropriate volume of loading sample buffer, were denatured by boiling them for 5 minutes at 95 °C. Samples were loaded into the wells; the running was performed with a constant voltage of 120 V for about 90 minutes in MOPS SDS running buffer.

The gels were then incubated in a staining solution for 1 h at room temperature. Finally, gels were destained using a destaining solution.

### **2.3.3 In-gel digestion of proteins separated by polyacrylamide gel electrophoresis**

Protein bands were excised from the polyacrylamide gel and cut in small cubes (1 mm). Gel particles were covered with acetonitrile and then rehydrated in 0.1 M  $\text{NH}_4\text{HCO}_3$ . After 5 minutes

of incubation, an equal volume of acetonitrile was added to wash gel pieces. The liquid was removed after 15 minutes and the gel particles were dried in a SpeedVac System. Gel pieces were swelled in 10 mM DTT/0.1 M  $\text{NH}_4\text{HCO}_3$  and incubated for 45 minutes at 56 °C to induce proteins reduction. The excess liquid was removed and the gel particles were shrunk with acetonitrile that was quickly replaced with 55 mM iodoacetamide in 0.1 M  $\text{NH}_4\text{HCO}_3$ . Gel pieces were incubated in this solution for 30 minutes in the dark. The liquid was removed and the gel particles were washed with 0.1 M  $\text{NH}_4\text{HCO}_3$  and acetonitrile as described above. Gel pieces were completely dried in a SpeedVac system and then rehydrated with a digestion buffer (50 mM  $\text{NH}_4\text{HCO}_3$ , 5 mM  $\text{CaCl}_2$ , trypsin 12.5 ng/ $\mu\text{L}$  (Boehring Mannheim sequencing grade) at 4 °C. After 120 minutes 40  $\mu\text{L}$  of digestion buffer without trypsin was added to the mixture to keep gel pieces hydrated during enzymatic cleavage overnight at 37 °C. The peptides were recovered from the gel by addition of a volume of 25 mM  $\text{NH}_4\text{HCO}_3$  sufficient to cover the gel pieces. After incubation for 15 minutes, the same volume of acetonitrile was added. After other 15 minutes incubation, the supernatant was recovered. The extraction was repeated two times with 5% formic acid. All the extracts were pooled and completely dried in the SpeedVac System before HPLC-MS analysis.

#### **2.3.4 Western Blot**

SDS-PAGE gels were run as described in section 2.3.2. Protein separated by SDS-PAGE were blotted onto nitrocellulose membranes (Protran BA83 Nitrocellulose Blotting Membrane, 0.2  $\mu\text{m}$ , 20 × 20 cm; Ge Healthcare) applying a constant voltage of 25 V for 2 hours in transfer buffer.

After blotting the proteins on membranes, they were blocked for at least 2 h using the blocking

solution. Subsequently, the membranes were incubated overnight at 4 °C with the primary polyclonal antibody against hRAD51 (1:1000 dilution in blocking solution). The day after, the membranes were subjected to three washes using TBS-T buffer and then incubated with the secondary antibody for 1 h at room temperature (1:10000 dilution in blocking solution). Finally, the membranes were washed three times in TBS-T solution. The membranes were exposed to the LiteAblot Extend Chemiluminescent Substrate (EuroClone) for a period sufficient to detect the bands and an image was recorded using the LAS4000 (Fujifilm) instrument.

### **2.3.5 Recombinant GST-hRAD51: expression, purification and protease inhibition trials**

#### *Protease inhibition trials*

In order to inhibit the GST-hRAD51 proteolytic degradation by ClpX protease and HslV-HslU protease complex different parameters were modified: addition to the bacteria growth medium of 1 mM ZnSO<sub>4</sub> or 300 mM KCl, alone or in combination, the bacteria growth media (M9, LB, ZYM-5052), the temperature of bacteria growth and induction (20 °C or 37 °C) and the time of induction (2 h, 3 h, 4 h or overnight for LB medium; 2 h and overnight for M9 medium; 16 h, 20 h and 24 h for ZYM-5052 medium).

Protease inhibition trials were performed starting from a pre-culture of BL21(DE3)pLysS/pGEX-4T-1-GST-hRAD51 cells grown overnight in LB medium at 37 °C. The saturated overnight culture was added to fresh LB supplemented with or without ZnSO<sub>4</sub> or KCl alone or in combination, or to the media cultures mentioned above starting from an initial OD<sub>600</sub> of 0.1 - 0.2. For growth and induction in LB or M9 media, flasks were incubated at 200 rpm and 37 °C until cells reached OD<sub>600</sub> of 0.6 - 0.7. Then, IPTG was added at 0.5 mM final concentration and induction was performed overnight at 37 °C or 20 °C.

For auto-induction medium ZYM-5052, flasks were incubated shaking at 200 rpm and 37 °C or 20 °C.

Proteolytic degradation was tested through western blot analysis on volumes of bacterial culture determined by Equation 2.1. Bacteria were centrifuged and all the pellets were resuspended in the same amount of sample loading buffer to obtain equal cells concentrations.

**Equation 2.1**

$$\text{Bacterial culture volume} = \left( \frac{1}{OD_{600}} \right) * 0.5$$

*Protocol for expression and purification of GST-hRAD51*

hRAD51 was expressed in *E. coli* BL21(DE3)pLysS cells. A saturated overnight culture of BL21(2DE3)pLysS/pGEX-4T-1-GST-hRAD51 was added into a fresh LB medium containing ampicillin to reach an initial OD<sub>600</sub> of 0.1 - 0.2. Flasks were shaken at 200 rpm and 37 °C until the cells reached the mid-log phase (OD<sub>600</sub> 0.6 - 0.7). Then IPTG was added to a final concentration of 0.5 mM. The bacterial culture was induced at 37 °C for 4 h. Subsequently, cells were centrifuged for 10 minutes at 6000 rpm and 4 °C. The cell pellet was frozen at -80 °C.

The bacteria cells pellet was subsequently resuspended in an appropriate volume of lysis buffer (100 mM Tris-HCl (pH 8.00), 7% (v/v) glycerol, 0.5 mM EDTA, 10% (w/v) Sucrose, 300 mM NaCl, 5 mM DTT, Urea 0.5 M) supplemented with protease inhibitor cocktail (SIGMAFAST™ Protease Inhibitor Cocktail Tablets, EDTA-Free). To the resuspended pellet PMSF 1 mM and lysozyme 100 µg/mL were added. The cell suspension was lysed on ice by sonication (12 rounds of 30"; Amplitude 85%; Tip KE76; Bandelin Sonoplus HD2070 sonicator). The disrupted cell suspension was centrifuged for 1 h at 13000 rpm. The supernatant fraction was

filtered with a 0.45  $\mu$ M membrane (MiniSart Syringe filter 0.45  $\mu$ M) to remove residual particles before chromatography. The supernatant was applied onto a GST-Trap column (GSTrap™ FF 1 mL; GE Healthcare) equilibrated with buffer A (50 mM Tris-HCl pH 8.00, 150 mM NaCl, 1 mM DTT, 0.5 mM EDTA, 1 mM Deoxy-cholate) using an Akta purifier™ 10 system (GE Healthcare). The column wash washed with buffer A until a stable baseline was reached and then the bound fusion protein GST-hRAD51 was eluted with a linear gradient over 10 column volumes using buffer B (buffer A supplemented with 10 mM reduced glutathione). The peak fractions containing the fusion protein were pooled and subsequently concentrated using an Amicon ultra centrifugal filter (cutoff 30 kDa; EMD Millipore) up to 1 mL final volume.

The pooled fractions were injected onto a gel filtration column (Superdex200 Increase 10/300 GL, GE Healthcare) equilibrated with buffer C (20 mM Tris-HCl (pH 8.00), 5% (v/v) Glycerol, 150 mM NaCl, 1 mM DTT, 1 mM Deoxy-cholate, 0.5 mM EDTA).

GST-hRAD51 purity was checked by SDS-PAGE and stored at -80 °C for further analysis.

### **2.3.6 Recombinant His-hRAD51: expression, purification and degradation trials**

#### *His-hRAD51 expression trials*

*E. coli* Rosetta2(DE3)pLysS cells were transformed with pET15b-His-hRAD51. Different parameters were modified for the optimization of the His-hRAD51 expression protocol: the bacteria growth medium (LB medium, ZYM-5052 and TB-5052 auto-induction media), temperature of *E. coli* growth (20 °C or 37 °C) and induction (20 °C or 37 °C), IPTG concentration (0.05 mM, 0.1 mM, 0.25 mM and 0.5 mM for growth in LB medium;) and time of induction (2 h, 3 h, 4, overnight for growth and induction in LB medium; 24 h, 32 h, 48 h, 56 h, 72 h for

growth in auto-induction media). His-hRAD51 expression trials were performed starting from a pre-culture in Rosetta2(DE3)pLysS/pET15b cells grown overnight in LB medium at 37 °C. The saturated overnight culture was added to fresh LB or to the auto-induction media to reach a starting OD<sub>600</sub> of 0.1 - 0.2. For growth and induction in LB medium, flasks were incubated at 200 rpm and 37 °C until cells reached OD<sub>600</sub> 0.6 - 0.7. Then, IPTG was added at 0.5 mM final concentration and induction was performed overnight at 37 °C or 20 °C.

For growth in auto-induction media ZYM-5052 and TB-5052, flasks were incubated shaking at 200 rpm and 20 °C. The optical density (OD<sub>600</sub>) of the bacterial cultures were measured at the end of the induction to characterize the bacteria growth under each condition.

#### *Degradation trials of His-hRAD51*

Proteolytic degradation of His-hRAD51 in Rosetta2(DE3)pLysS cells was evaluated by western blot loading the soluble fraction of proteins extracted from *E. coli* bacteria cells grown under different conditions as described above for GST-hRAD51.

#### *Protocol for the expression and purification of His-hRAD51*

hRAD51 was expressed in *E. coli* Rosetta2(DE3)pLysS cells. A saturated overnight culture of Rosetta2(DE3)pLysS/pET15b-His-hRAD51 was diluted (1:1000) into a fresh TB-5052 auto-induction medium containing ampicillin (100 µg/mL). The flasks were shaken at 200 rpm at 20 °C for 72 h. Subsequently, cells were centrifuged for 10 minutes at 6000 rpm and 4 °C. Cell pellet was frozen at -80°C.

Pellet was subsequently resuspended in an appropriate volume of buffer A (20 mM Tris-HCl (pH 8.00), 500 mM NaCl, 10 mM Imidazole, 5 mM DTT, 10% (v/v) Glycerol) supplemented with protease inhibitor cocktail (SIGMAFAST™ Protease Inhibitor Cocktail Tablets, EDTA-

Free). The cell suspension was lysed on ice trough sonication (24 rounds of 30"; Amplitude 85%; Tip MS72; Bandelin Sonoplus HD2070 sonicator). The disrupted cell suspension was centrifuged for 1 h at 13000 rpm. The supernatant fraction was filtered with a 0.45  $\mu$ M (MiniSart Syringe filter 0.45  $\mu$ M) membrane to remove residual particulates before chromatography. The supernatant was applied onto a His-Trap column (His-Trap<sup>TM</sup> FF 5mL, GE Healthcare) equilibrated with buffer A. A wash step was performed using 10% of buffer B (20 mM Tris-HCl (pH8.00), 500 mM NaCl, 500 mM Imidazole, 10% (v/v) Glycerol). The protein was then eluted with a linear gradient from 10% to 100% of buffer B over 10 column volumes. Fractions (0.5 mL) were collected and analyzed by SDS-PAGE. Collected fractions corresponding to the recombinant protein were dialyzed overnight at 4 °C against buffer C (50 mM Tris-HCl (pH 8.00), 200 mM KCl, 0.25 mM EDTA, 2 mM DTT, 10% (v/v) glycerol). Dialyzed protein was loaded onto an anion exchange column (ResQ, GE Healthcare) equilibrated in buffer C. The elution was performed with a linear gradient of buffer B (50 mM Tris-HCl (pH 8.00), 1 M KCl, 0.25 mM EDTA, 2 mM DTT, 10% (v/v) glycerol). Fractions (0.5 mL) were collected and analyzed by SDS-PAGE. Fractions containing His-hRAD51 were pooled and dialyzed against the storage buffer (20 mM Hepes (pH 8.00), 250 mM KCl, 0.1 mM EDTA, 2 mM DTT, 10% (v/v) glycerol). The protein yield was determined from the optical absorption at 280 nm (extinction coefficient 14900 M<sup>-1</sup>cm<sup>-1</sup>) of the final sample.

### **2.3.7 His-hRAD51-BRC4 complex: expression and purification**

#### *Protocol for the expression and purification of His-RAD51-BRC4 complex*

hRAD51-BRC4 complex was expressed in *E. coli* Rosetta2(DE3)pLysS cells. A saturated overnight culture in LB-medium of Rosetta2(DE3)pLysS co-transformed with pET15b-His-

hRAD51 and pET28a-BRC4 was diluted (1:1000) into a fresh TB-5052 auto-induction medium containing ampicillin (100 µg/mL) and kanamycin (50 µg/mL). The flasks were shaken at 200 rpm at 20 °C for 72 h. Subsequently, cells were centrifuged for 10 minutes at 6000 rpm and 4 °C. Cell pellet was frozen at -80 °C.

Pellet was subsequently resuspended in an appropriate volume of buffer A (20 mM Tris-HCl (pH 8.00), 500 mM NaCl, 10 mM Imidazole, 5 mM DTT, 10% (v/v) Glycerol) supplemented with protease inhibitor cocktail (SIGMAFAST™ Protease Inhibitor Cocktail Tablets, EDTA-Free) and PMSF 1 mM. The cell suspension was lysed on ice through sonication (24 rounds of 30"; Amplitude 85%; Tip MS72; Bandelin Sonoplus HD2070 sonicator). The disrupted cell suspension was centrifuged for 1 h at 13000 rpm. The supernatant fraction was filtered with a 0.45 µm (MiniSart Syringe filter 0.45 µm) membrane to remove residual particles before chromatography. The supernatant was applied onto a His-Trap column (His-Trap™ FF 5 mL, GE Healthcare) equilibrated with buffer A. A wash step was performed using 5% of buffer B (20 mM Tris-HCl (pH 8.00), 500 mM NaCl, 500 mM Imidazole, 10% (v/v) Glycerol). Then, the protein was eluted with a linear gradient from 5% to 100% of buffer B over 10 column volumes. Fractions (0.5 mL) were collected and analyzed by SDS-PAGE. Fractions containing His-hRAD51-BRC4 complex were pooled and then concentrated up to 0.5 mL using a concentrator (Amicon Ultra-15 Centrifugal Filter Units; cut-off 10 kDa; Merck-Millipore). The concentrated sample was injected onto a size exclusion chromatographic column equilibrated with buffer C (50 mM Tris-HCl (pH 8.00), 200 mM KCl, 0.1 mM EDTA, 2 mM DTT, 5% (v/v) glycerol). Fractions (0.5 mL) were collected and analyzed by SDS-PAGE.

### **2.3.8 His-hRAD51 F86E-BRC4 complex: expression and purification**

*Protocol for the expression and purification of His-RAD51 F86E-BRC4 complex*

hRAD51 F86E-BRC4 complex was expressed in *E. coli* Rosetta2(DE3)pLysS cells. A saturated overnight culture in LB medium of Rosetta2(DE3)pLysS co-transformed with pET15b-His-hRAD51 F86E and pET28a-BRC4 was diluted (1:1000) into a fresh TB-5052 auto-induction medium containing ampicillin (100 µg/mL) and kanamycin (50 µg/mL). The flasks were shaken at 200 rpm at 20 °C for 72 h. Subsequently, cells were centrifuged for 10 minutes at 6000 rpm and 4 °C. Cell pellet was frozen at -80 °C.

The His-hRAD51 F86E-BRC4 complex was purified as described in section 2.3.7.

## 2.4 Biophysical methods

### 2.4.1 Protein and plasmid DNA quantification

Protein absorbance was measured at 280 nm using a UV-visible Nanodrop2000 spectrophotometer (ThermoFisher) and protein concentration was calculated according to the Lambert-Beer law (Equation 2.2).

Equation 2.2

$$A = \epsilon * c * l$$

where  $\epsilon$  represents the extinction molar coefficient expressed in  $M^{-1} * cm^{-1}$ ;  $c$  is the protein concentration (mg/mL) and  $l$  is the light path in the sample (usually 1 cm).  $\epsilon$  was calculated for GST-hRAD51 and His-hRAD51 (Table 2.6) sequences reported in this thesis using ProtParamTool ([www.expasy.org/protparam](http://www.expasy.org/protparam)). Instead, it was not possible to measure the absorbance at 280 nm of His-BRC4 because there are no tryptophan, tyrosine or cysteine residues in the aminoacidic sequence of the peptide.

Protein	$\epsilon$ ( $M^{-1} cm^{-1}$ )	Molecular Weight (Da)
GST-hRAD51	57760	63258.74
His-hRAD51	14900	39715.30
His-BRC4		6410.26

**Table 2.6:** Extinction molar coefficients ( $\epsilon$ ) and molecular weight of GST-hRAD51, His-hRAD51 protein and His-BRC4 proteins.

Plasmid DNA concentration was obtained measuring the absorbance of the nucleic acids at 260 nm, where they present their maximum of absorbance.

### **2.4.2 Surface Plasmon Resonance (SPR)**

Two different sensor chips were used to immobilize the His-hRAD51 protein: a COOH5 and a His-Cap sensor chip (PALL Fortebio). The analyses were performed using a Pioneer FE SPR system (PALL Fortebio).

His-hRAD51 (isoelectric point, pI = 5.42) was diluted in acetate buffer (pH = 5.00) to confer a significant positive charge to the protein and to promote its binding to the COOH5 surface. The low ionic strength in the buffer determined aggregation and consequently precipitation of His-hRAD51 in acetate buffer. Thus, it was not possible to perform the following step of protein immobilization on the sensor chip.

A second attempt was performed using the His-Cap sensor chip. A solution of 500  $\mu\text{M}$   $\text{NiCl}_2$  was injected (10  $\mu\text{L}$ ; 10  $\mu\text{L}/\text{min}$ ) to promote the activation of the His-Cap surface. His-hRAD51 diluted in the SPR running buffer (20 mM Hepes (pH 8.00), 250 mM KCl, Glycerol 2%) was later injected over the sensor surface (50  $\mu\text{L}$ ; 10  $\mu\text{L}/\text{min}$ ) setting up a dissociation time of 1200 seconds. Regeneration of the surface was performed injecting a solution of 350 mM EDTA (75  $\mu\text{L}$ ; 30  $\mu\text{L}/\text{min}$ ) followed by the injection of 1 M imidazole (75  $\mu\text{L}$ ; 30  $\mu\text{L}/\text{min}$ ). This immobilization cycle reaction was repeated using four different His-hRAD51 concentrations (96.2 nM, 192 nM, 385 nM, 770 nM).

### **2.4.3 Microscale Thermophoresis (MST)**

His-hRAD51 was labelled with the NT-495 blue dye (NanoTemper Technologies) that binds the hRAD51 amine group of lysine residues. MST measurements simultaneously performed on 16 capillaries containing a constant concentration (80 nM) of NT-495-labelled-His-hRAD51 protein and 16 different concentrations of BRC4 peptide or ARN22064 allowed to determine a concentration dependent MST binding curve. The highest BRC4 concentration

tested was 16  $\mu$ M. The highest ARN22064 concentration used was 3 mM.

Measurements were carried out in MST buffer (20 mM Hepes (pH 8.00), 200 mM KCl, 0,1 mM EDTA, 5% (v/v) Glycerol). The ARN22064 compound required the addition of 5% DMSO in the MST buffer for solubility issues.

His-hRAD51 was also labelled with NT-647 red dye (NanoTemper Technologies) that binds the hexahistidine tail of the recombinant protein. MST measurements were performed as above described titrating a constant concentration (100 nM) of NT-647-HisRAD51 with 16 concentrations of BRC4 peptide or each of the following compounds: ARN22142, E5, A6 and F11. The highest concentrations used for the titrants were: 16  $\mu$ M for BRC4, 6 mM for ARN22142 and 4 mM for E5, A6 and F11.

The MST buffer was supplemented with 5% DMSO to increase the solubility of the molecule and the fragments.

A displacement assay was set up pre-incubating NT-647-labelled-His-hRAD51 (100 nM) with ARN22142 (6 mM) or with E5 (1 mM). A constant concentration of complexes was then titrated as described above with BRC4. The highest BRC4 concentration used was 1 mM. The MST analysis was performed using the NT.115 Monolith (Nanotemper Technologies). Binding curves were fitted using the Affinity Analysis software of Nanotemper Technologies to obtain binding affinity data.

#### **2.4.4 Nuclear Magnetic Resonance (NMR)**

All NMR experiments were recorded at 25 °C with a Bruker FT NMR Avance III 600-MHz spectrometer equipped with a 5-mm CryoProbe QCI  $^1\text{H}/^{19}\text{F}-^{13}\text{C}/^{15}\text{N}$ -D-Z quadruple resonance, a shielded z-gradient coil, and the automatic sample changer SampleJet NMR system with

temperature controls. SPAM filter approach was used to characterize the chemical triazole class and the compounds ARN22064 and ARN22142 in buffer solutions. 1D  $^1\text{H}$  water suppression in all proton experiments was achieved with the excitation sculpting sequence<sup>90</sup>; for a total number of 256 scans with a repetition time of 5 s. The two water selective  $180^\circ$  square pulses and the four PFGs were 2.5 and 1 ms in duration, respectively. WaterLOGSY experiments were acquired with 512 scans, 5 s repetition time, and 1.5 s mixing time.

For the NMR Fragments screening a monodimensional  $^{19}\text{F}$  NMR spectrum and a  $R_2$  filter experiment were recorded for a total of 64 scans with a repetition time of 5 s and proton decoupling during the acquisition period. For the  $R_2$  filter experiments a spin-echo scheme with total  $\tau = 0.2$  s was used. The data were multiplied by an exponential function corresponding to a line broadening of 1 Hz prior to Fourier transformation. The reference standard in the  $^{19}\text{F}$  spectra was  $\text{CFCl}_3$ .

#### **2.4.5 Crystallization trials**

##### *His-hRAD51 crystallization trials*

His-hRAD51 protein was concentrated up to  $\sim 5 - 10$  mg/mL. Crystals of His-hRAD51 were grown with the hanging drop method in 96 well plates. Drops were prepared mixing equal volume (0.1  $\mu\text{L}$ ) of protein and reservoir solutions using a Mosquito Crystal Nanolitre Protein Crystallization robot (TTP Labtech). Crystal Screen and Crystal Screen 2 reagent kits (Hampton research) were used for protein crystallization.

##### *His-hRAD51 F86E-BRC4 crystallization trials*

In a first set of experiments His-hRAD51 F86E-BRC4 complex was concentrated up to 26 mg/mL. Crystals of His-hRAD51 F86E-BRC4 were grown with the hanging drop method in

96 well plates. Drops were prepared mixing equal volume (0.2  $\mu$ L) of protein and reservoir solutions using a Mosquito Crystal Nanolitre Protein Crystallization robot (TTP Labtech). The following screening kits were used: ProPlex 1-96, JCSG-plus (Molecular Dimensions) and Crystal screen (Hampton research).

In a second set of experiments crystals of His-hRAD51 F86E-BRC4 complex were concentrated up to 13.5 mg/mL and grown as above described. The following crystallization screening kits were used: PEG/Ion, PEGRx, Crystal screen, Salt Rx (Hampton research), MCSG 1, MCSG 2, JCSG-plus, ProPlex 1-96 (Molecular dimensions).

## **2.5 His-hRAD51 characterization**

### **2.5.1 Size exclusion chromatography**

The recombinant pure His-hRAD51 was incubated at 37 °C for one hour in the absence or presence of the BRC4 peptide in a 1:4 molar ratio. After incubation, protein samples were loaded onto a size exclusion chromatographic column (Superdex200 Increase 10/300 GL, GE Healthcare) equilibrated with the following buffer: 20 mM Hepes (pH 8.00), 250 mM KCl, EDTA 0.1 mM, Glycerol 5%, DTT 2 mM.

### **2.5.2 Negative staining transmission electron microscopy**

Recombinant pure His-hRAD51 protein (1 mg/mL) was placed onto ultra-thin carbon film 300-mesh copper grids and negatively stained in 0.2% - 0.5% uranyl acetate at pH 4.5 - 5.5. The samples were observed with a FEI Tecnai G2 F20 field emission gun transmission electron microscope (FEI Company, The Netherlands), working at an acceleration voltage of 80 keV and equipped with a 2k X 2K Gatan Ultrascan CCD camera (Gatan company, USA).

### **2.5.3 Dynamic Light scattering (DLS)**

His-hRAD51 (27 µM) was analyzed in DLS buffer (20 mM Hepes (pH 8.00), 250 mM KCl, 0.1 mM EDTA, Glycerol 5%) and the particle size of His-hRAD51 was analyzed using a Zetasizer Nano Particle Analyzer (Malvern) at 25 °C using the standard operating procedure for size measurements. Data were collected at five different time points and expressed in terms of intensity distribution.

---

---

Chapter three

# Results and discussion

---

---

The literature reports, describing the efforts to characterize through biophysical and structural approaches the human hRAD51 (hRAD51) protein and its mechanism of action, evidence that the expression of the recombinant full-length form of the protein is actually quite challenging. The first attempt to obtain a crystal of the full-length RAD51 was done by Shin *et al.*<sup>91</sup> who tried to shed light on RAD51 interactions able to control and affect homologous recombination (PDB: 1SZP). However, to achieve this aim, a humanized RadA protein (from *Pyrococcus furiosus*), homologous of hRAD51 (50% identity), was expressed.

Until today, the only published structure describing the hRAD51-BRC4 complex, (PDB ID: 1N0W), was obtained by removing the first 97 amino acids of hRAD51<sup>51</sup>. The two proteins were purified as a single fusion protein linked by a flexible polypeptide sequence (Thr-Gly-Ser)<sub>4</sub>Met-Gly.

The N-terminal region of hRAD51 is highly conserved among eukaryotic RAD51 proteins, suggesting a specific role for this RAD51 region. This domain, highly disordered<sup>92</sup>, is relevant for protomer-protomer interaction through the polymerization motif 85-GGFTTATE-91.

In 1999, Aihara *et al.*<sup>92</sup> partially solved, by NMR, the structure of the N-terminal region of hRAD51 expressed as a truncated form of the first 114 amino acids. It consists of five short helices even though most of the peptide (residues 1 to 15 and 86 to 114) is disordered.

Many of the issues encountered by researchers, working with the full-length hRAD51, come from the significant mobility of the N terminal region of the protein and from the protein natural tendency to self-aggregate that can lead to considerable protein instability.

In this research work, we face the challenge of expressing, purifying and characterizing the recombinant full-length hRAD51 in order to fully characterize the interaction of hRAD51 with BRCA2 (BRC4) and clearly unveil the hot spots of the protein-protein interactions through biophysical characterization as well as X-ray crystallography.

## **Optimization of a protocol for the overexpression and purification of full-length hRAD51 in bacteria cells**

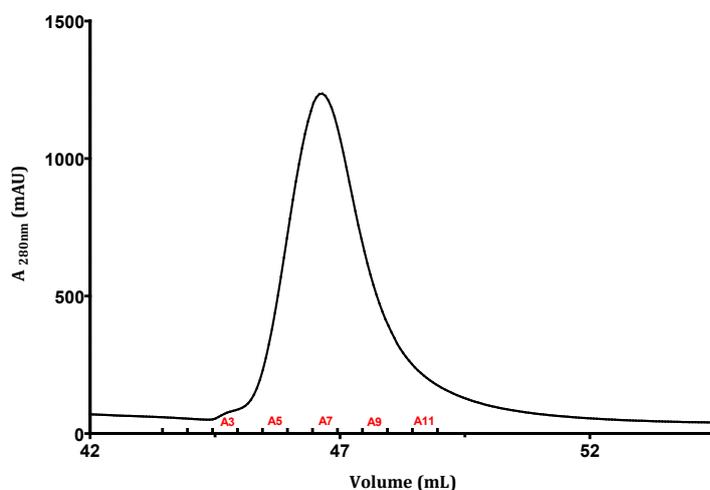
### **3.1 Expression and purification of GST-hRAD51**

The pGEX-4T-1-GST-hRAD51 vector for the expression in *Escherichia coli* (*E. coli*) of hRAD51 with a GST tag at the N-terminus was already available within the project. The GST moiety at the N-terminus was inserted to improve the solubility of the recombinant protein. BL21(DE3)pLysS *E. coli* cells were transformed with the above mentioned vector in order to express the heterologous protein. The expression protocol for hRAD51 in *E. coli* cells was pursued setting up a series of expression trials with variable parameters: temperature of bacteria growth before and after induction, concentration of IPTG and time of induction.

The protocol for GST-hRAD51 purification was optimized using two chromatographic steps: an affinity chromatography (GST-Trap column) and a size exclusion chromatography (Superdex200 column).

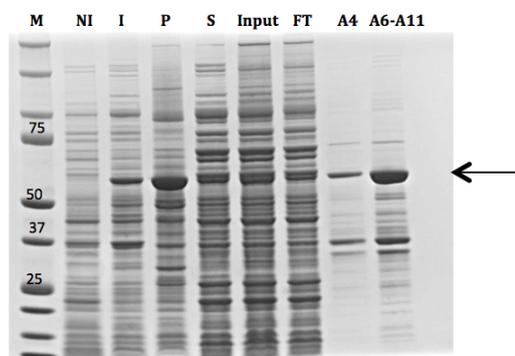
#### *Affinity chromatography*

Proteins extracted from the bacterial culture through cells sonication were loaded onto an affinity GST-Trap column (see Material and Methods). The elution chromatogram, resulted from the addition of 10 mM reduced glutathione to the elution buffer, is reported in Figure 3.1 and it shows the presence of a single peak (fractions A6-A11) with a small shoulder (fraction A4).



**Figure 3.1:** Chromatogram corresponding to the elution profile of the affinity chromatography (GST-Trap column).

Samples from the expression and first purification step of GST-hRAD51 were analyzed by SDS-PAGE (Figure 3.2).

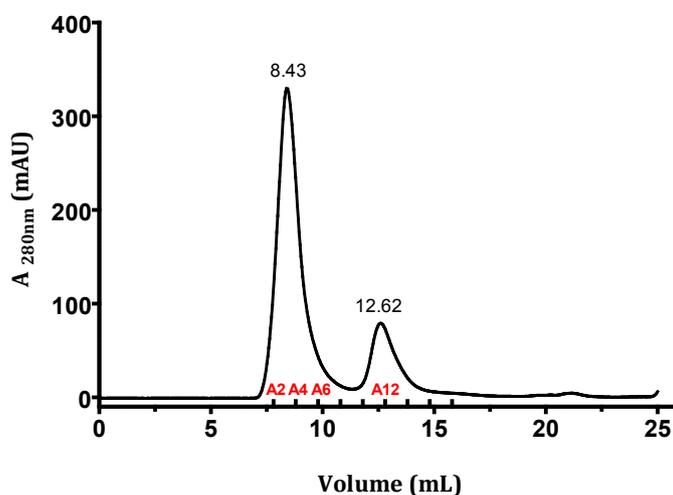


**Figure 3.2:** SDS-PAGE of samples from the expression and first step of GST-hRAD51 purification: *M*: marker; proteins contained in a not induced (*NI*) and induced (*I*) bacteria sample; insoluble (*P*) and supernatant (*S*) fractions after bacteria cells lysis; loading input of the chromatographic column (*Input*); flow through (*FT*) and fractions (*A4* and *A6-A11*) corresponding to the shoulder and main peak of the chromatographic elution (GST-Trap column) (see Figure 3.1). The black arrow indicates the full-length GST-hRAD51.

Even though a significant amount of insoluble fraction was detected after cell lysis, part of the protein was found in the soluble fraction. The recombinant protein binds the column resin as it can be observed by the presence, in the SDS-PAGE gel (pool fractions A6-A11), of a band at about 63 kDa, corresponding to the molecular weight of GST-hRAD51.

### Size exclusion chromatography

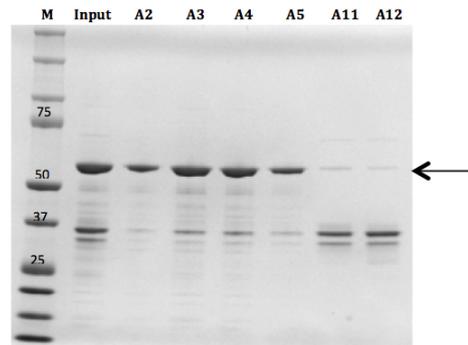
The A6-A11 fractions from the previous step of purification were pooled and concentrated. The obtained sample was then loaded onto a Superdex200 column (Material and Methods) in order to increase the purity of GST-hRAD51 and to evaluate the protein aggregation state. The size exclusion chromatographic profile (Figure 3.3) shows the presence of two peaks: the former is eluted in the column dead volume (8.43 mL), whereas the second one is eluted after 12.62 mL.



**Figure 3.3:** Chromatogram corresponding to the Superdex200 elution profile.

Fractions corresponding to the peaks eluted from the gel filtration were analyzed by SDS-PAGE (Figure 3.4). GST-hRAD51 (63 kDa) was the main component of the A2-A5 fractions, together with the presence of a weak band at 35 kDa. The latter band becomes, instead, the main constituent of the A11 and A12 fractions.

The size exclusion chromatographic step, besides not significantly improving the purity of the expressed GST-hRAD51 protein (compared to the loaded sample), shows that the protein was purified as a high molecular weight oligomer (elution in the column dead volume), most likely as an already self-assembled protein.

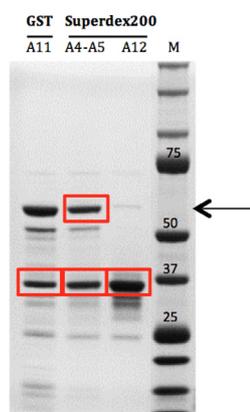


**Figure 3.4:** SDS-PAGE analysis of samples after gel filtration. *M*: marker; *Input*: loading input of the Superdex200 chromatographic column; *A2-A12*: fractions of the Superdex200 elution. The black arrow indicates the full-length GST-hRAD51.

The final protein yield, using the pGEX-4T-1-GST-hRAD51 construct, was 1.2 mg per liter of bacteria culture.

### 3.1.1 Degradation trials

In the attempt of improving purity and yield of the GST-hRAD51 protein we looked into the identity of specific proteins observed in the SDS-PAGE gel pattern of the peaks eluted from the two purification steps (Figure 3.5). Specific bands, as the one recurrently appearing at 35 kDa in the SDS-PAGE, were cut from the gel, extracted and digested (Material and Methods). The products were analyzed by HPLC mass spectrometry (HPLC-MS) analysis. Interestingly, the presence of two truncated forms of GST-hRAD51 were identified with a molecular weight of 35.60 kDa and 35.46 kDa (data not shown). Identification of all the proteins present in the elution samples of the first chromatographic step, through HPLC-MS analysis, allowed the identification of two specific *E. coli* proteases (the ClpX protease and the HslU subunit, which form an ATP dependent protease complex with HslV<sup>93</sup>). These two proteases, expressed in *E. coli* bacteria, are most likely responsible for the specific protein truncation observed. These data support our initial hypothesis i.e. purity and yield of the GST-hRAD51 protein are significantly affected by degradation issues.



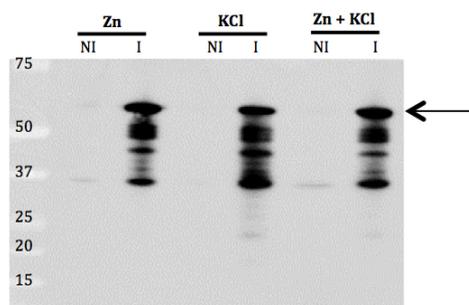
**Figure 3.5:** SDS-PAGE of samples (see Figure 3.3) eluted from affinity and size exclusion chromatographies and analyzed by HPLC-MS. Bands highlighted by red boxes were cut, extracted from the gel and analyzed by HPLC-MS to identify the amino acid sequence of the proteins corresponding to the bands. The black arrow indicates the full-length GST-hRAD51.

Degradation was confirmed storing the samples eluted from the affinity chromatography for one week at 4 °C. Comparison of SDS-PAGE gels of the same samples before and after storage showed that the band corresponding to the full-length GST-hRAD51 (~ 63 kDa) was converted, in time, into the band at about 35 kDa (not shown).

#### *Protease inhibition trials*

Previous studies have shown that ClpX protease can be inhibited by  $\text{ZnSO}_4$  or  $\text{KCl}$ <sup>94</sup>, while HslV-HslU protease complex can be inhibited by the peptide vinyl sulfone, carboxybenzyl-leucyl-leucyl-leucine vinyl sulfone<sup>95</sup>. Since the synthesis of the latter is quite prohibitive, several attempts were made to inhibit ClpX protease by the addition of  $\text{ZnSO}_4$  (1 mM) or  $\text{KCl}$  (300 mM) alone or in combination.

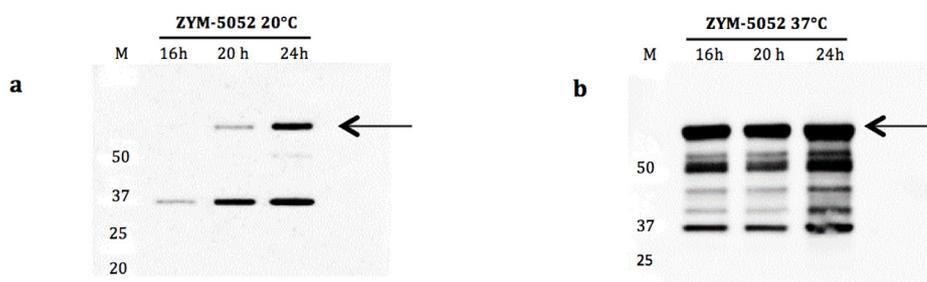
Addition of  $\text{ZnSO}_4$  1mM or  $\text{KCl}$  300 mM in the bacteria culture did not prevent protein degradation as shown in Figure 3.6 (western blot analysis). The recombinant protein is cleaved in different digestion products, independently from the addition of the two salts.



**Figure 3.6:** Western blot analysis using a polyclonal antibody against hRAD51 of not induced (NI) and induced BL21(2DE3)pLysS/pGEX-4T-1-GST-hRAD51 samples after addition in medium culture of  $ZnSO_4$  (1 mM) or KCl (300 mM) alone or in combination. The black arrow indicates the full-length GST-hRAD51.

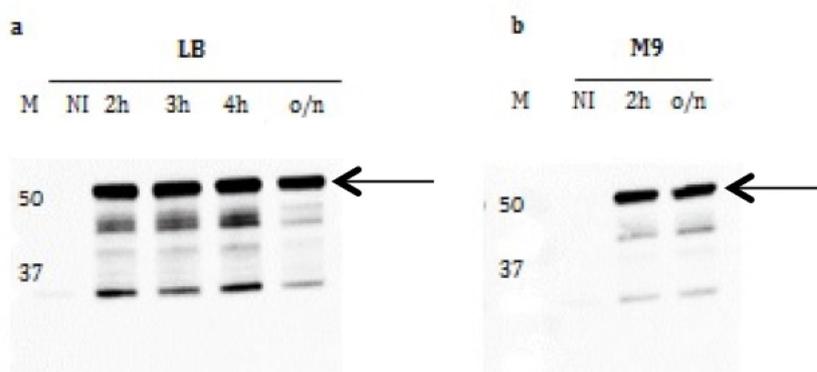
Bacteria growth and induction conditions were modified in an attempt to decrease the rate of degradation. Different parameters were modified: medium composition (ZYM-5052, LB, M9), temperature of growth and induction (20 °C or 37 °C).

Western blot analyses of the bacteria samples, grown in different culture media and induced at different temperatures, were performed. The analysis of bacterial cells grown in ZYM-5052 medium showed a consistent degradation of GST-hRAD51 at 37 °C (digestion products with molecular weights between 37-50 kDa), whereas degradation was less evident when cells were grown at 20 °C (Figure 3.7a and b).



**Figure 3.7:** Western blot analysis using a polyclonal antibody against hRAD51 to evaluate the proteolytic degradation of GST-hRAD51 in bacterial cells grown in ZYM-5052 medium at 20 °C (a) and 37 °C (b) after different incubation times (16 h; 20 h or 24 h). The black arrows indicate the full-length GST-hRAD51.

The proteolytic degradation in cells grown in LB medium, irrespective of the induction time, was significant. In this case, as shown by western blot analysis (Figure 3.8), several digestion products of the full-length protein were detected. Lower degradation levels were observed in cells induced at 20 °C. Low degradation was also observed in bacterial cells grown at 37 °C in M9 medium and induced at 20 °C or 37 °C (Figure 3.8).



**Figure 3.8:** Western blot analysis using a polyclonal antibody against hRAD51 to evaluate the proteolytic degradation of GST-hRAD51 in bacterial cells grown in LB medium (a) at different induction temperatures (20 °C or 37 °C) and times (2 h, 3 h, 4 h or overnight) or in M9 medium (b) at different induction temperatures (20 °C or 37 °C) and times (2 h or overnight). The black arrows indicate the full-length GST-hRAD51.

Even though the temperature of bacteria growth and induction affects the degradation of the overexpressed recombinant protein, it was not possible to inhibit GST-hRAD51 protein degradation significantly improving protein purity and yield.

The optimization of the protocol for the expression and purification of GST-hRAD51 from BL21(DE3)pLysS cells presented several challenges. A significant amount of recombinant protein was lost in the insoluble fraction after cell lysis. GST-hRAD51 was expressed in low yield and the final protein preparation was not pure enough for further biophysical analysis.

Additional analyses showed that proteolytic degradation, operated by two specific endogenous *E. coli* proteases, was significantly affecting the purification. Several attempts were

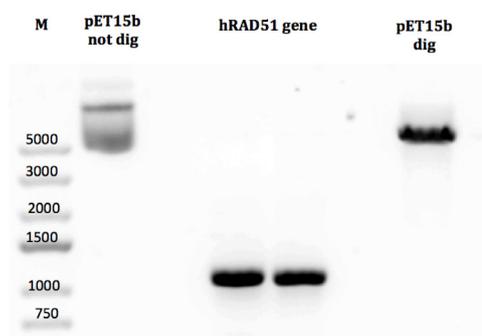
made to inhibit the degradation process. The use of potential proteases inhibitors (KCl or Zn<sup>2+</sup>) or the modification of *E. coli* cells growth and induction conditions did not allow preventing the proteolytic events.

Taken together these experiments showed the need to change expression and purification strategy in order to produce a full-length pure hRAD51 protein in high yield.

### 3.2 Expression and purification of His-hRAD51

#### *hRAD51 subcloning into pET15b vector*

Previous studies showed that proteolytic degradation, operated by endogenous proteases, can sometimes be avoided by changing both bacterial cell strain and vector<sup>96</sup>. Thus, a new attempt to express and purify the full-length hRAD51 was made by subcloning the gene sequence, coding for the hRAD51 protein, into a pET15b vector in order to overexpress hRAD51 with a His-tag at the N-terminal region (Materials and Methods). hRAD51 gene sequence (1107 bp), amplified by PCR reaction, and the pET15b vector (5708 bp) were double digested using BamHI and Bpu1102I restriction enzymes. hRAD51 gene sequence and pET15b vector, extracted and purified from agarose gel (Figure 3.9), were then ligated in order to obtain the pET15b-His-hRAD51 construct.



**Figure 3.9:** Agarose gel of hRAD51 gene and pET15b vector (pET15b dig) after double digestion with BamHI and Bpu1102I enzymes. Not digested pET15b vector (pET15b not dig) was also loaded as negative control.

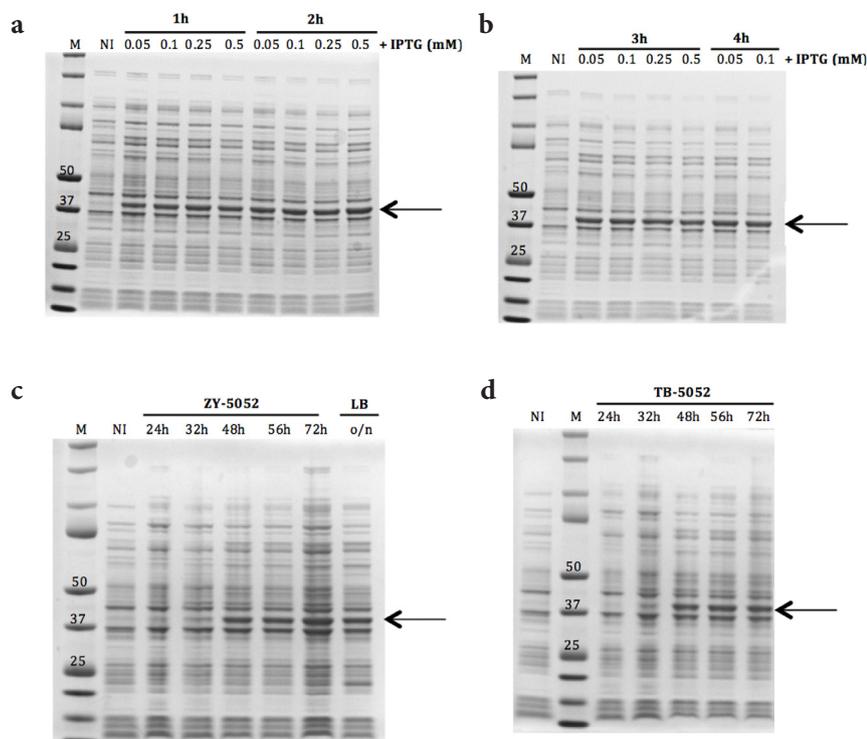
*E. coli* XL1-Blue cells were transformed with pET15b-His-hRAD51. Four different clones were selected for vector amplification. The DNA plasmid was extracted and quantified. Sequencing of plasmid DNA showed that only two pET15b-His-hRAD51 DNA plasmids had the right sequence (clone 1 and 2), whereas DNA plasmids from the other two clones (3 and 4) presented point mutations (data not shown). Hence, the vector pET15b-His-hRAD51 from clone 1 was amplified (Mini- and Maxi-Prep, kit Qiagen).

#### *Expression trials*

*E. coli* Rosetta2(DE3)pLysS cells were transformed with pET15b-His-hRAD51 (clone 1) for the expression of the protein of interest. The optimization of the protocol for hRAD51 expression was pursued running different expression trials changing the following parameters: *E. coli* growth medium (LB medium; ZYM-5052 medium and TB-5052), temperature of *E. coli* cells growth and induction (20 °C and 37 °C), IPTG concentration (0.05 mM, 0.1 mM, 0.25 mM and 0.5 mM), time of induction (2 h, 3 h, 4 h, 24 h, 32 h, 48 h, 56 h, 72 h). The optical density (OD<sub>600</sub>) of the bacterial culture was measured at the end of the induction to characterize bacteria growth under each condition. The amount of soluble protein, extracted under each set of conditions, was evaluated through SDS-PAGE gel of the lysed *E. coli* bacteria pellet.

In Rosetta2(DE3)pLysS/pET15b cells grown in LB medium, the highest amount of His-hRAD51 was detected after 3 h and 4 h of induction at 37 °C or after overnight induction at 20 °C, irrespective of the IPTG concentration used (Figure 3.10a, b and c). Thus, the time of induction plays a key role for the expression of His-hRAD51. The final OD<sub>600</sub> for Rosetta2(DE3)pLysS/pET15b cells grown in LB medium was 3.16 after 4 h and 37 °C, while OD<sub>600</sub> was 3.3 after overnight induction at 20 °C.

In auto-induction media, ZYM-5052 or TB-5052, the maximum expression was observed after 72 h of incubation, whereas only a small amount was produced after 24 h or 32 h (Figure 3.10c and d). Thus, the expression of protein was dependent on the time of incubation. The final OD<sub>600</sub> calculated for Rosetta2(DE3)pLysS/pET15b grown in ZYM-5052 and TB-5052 after 72 h was 14 and 18.5, respectively.



**Figure 3.10:** SDS-PAGE analysis of pET15b-His-hRAD51 expression in cells grown in different culture media. (a) and (b) samples from cells grown in LB medium at 37 °C for 1, 2, 3 and 4 h using different IPTG concentrations. Only the sample of cells grown in LB medium at 37 °C and induced at 20 °C overnight using 0.5 mM IPTG is reported in panel (c); (c) samples from cells grown in auto-induction medium ZYM-5052 for 24, 32, 48, 56 and 72 h after inoculation; (d) samples from cells grown in auto-induction medium TB-5052 for 24, 32, 48, 56 and 72 h after inoculation. The black arrows indicate the full-length His-hRAD51.

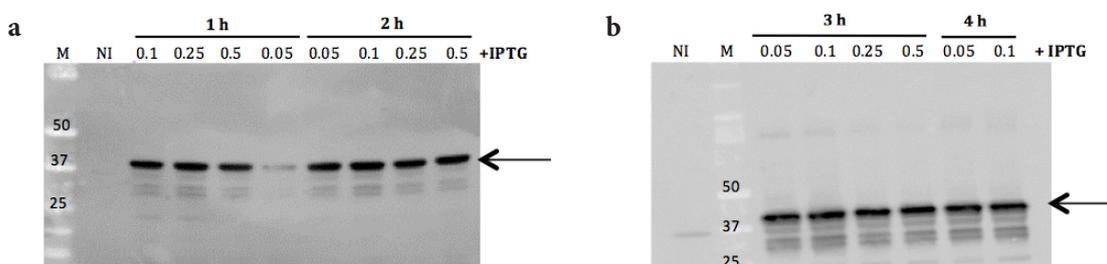
The SDS-PAGE gels and bacteria growth parameters showed no relevant differences in the expression of His-hRAD51 in LB medium or in auto-induction media. However, the bacterial cells concentration, measured as optical density at 600 nm, was much higher in ZYM-5052 or TB-5052 medium than in LB medium.

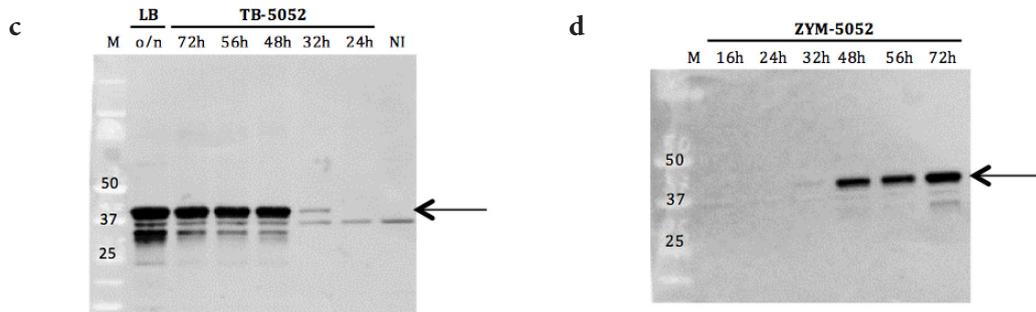
### 3.2.1 Degradation trials

Western blot analyses were performed to evaluate if degradation of His-hRAD51 protein also occurred in *E. coli* Rosetta2DE3pLysS cells grown under the different conditions described above (Figure 3.10).

In the samples from cells grown in LB medium, a band at 40 kDa, corresponding to His-hRAD51, was present in all samples. However, the products of protein digestion (bands at ~37 kDa, ~35 kDa and ~33 kDa) were present as well, irrespective of the induction time and IPTG concentration. A low intensity band was observed at 37 kDa also in the “not induced” samples, because the polyclonal antibody against hRAD51 most likely recognizes also RecA, the endogenous bacterial homologous of hRAD51. The proteolytic digestion events are time dependent, in fact, degradation is more evident after 3 h, 4 h or overnight (Figure 3.11a, b and c) induction.

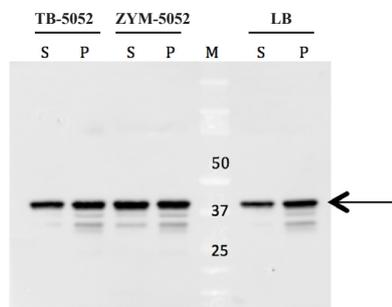
Conversely, in bacterial cells grown in auto-induction media at 20 °C, the degradation events were less evident, in fact, in both media, only a degradation band at 35 kDa was detected (Figure 3.11c and d). An increase of the temperature of bacteria growth and induction enhances protein degradation.





**Figure 3.11:** Western blot analysis of expression trials samples in different culture media using a polyclonal antibody against *hRAD51*. **(a)** and **(b)** samples from cell grown and induced in LB medium at 37 °C for 1, 2, 3 and 4 h using different IPTG concentrations. Only the sample of cells grown in LB medium at 37 °C and induced at 20 °C overnight using 0.5 mM IPTG is reported in panel **(c)**; **(c)** samples from cells grown in auto-induction medium ZYM-5052 after 24, 32, 48, 56 and 72 h after inoculation; **(d)** samples from cells grown in auto-induction medium TB-5052 after 24, 32, 48, 56 and 72 h after inoculation. The black arrows indicate the full-length His-*hRAD51*.

In order to further test the effects induced by the degradation process, western blot analyses were performed on soluble and insoluble protein fractions extracted from bacteria cells grown in TB-5052, in ZYM-5052 and LB media under the conditions where the highest protein expression was previously observed. The western blot analysis (Figure 3.12) showed an expected band at 40 kDa, corresponding to His-*hRAD51* in both soluble and insoluble protein fractions, after cell lysis. A degradation band at 35 kDa was detected in all insoluble protein fractions and in the soluble fraction of cells grown in ZYM-5052 medium. Products of digestion were, conversely, not detected in supernatant fractions of cells grown in LB and TB-5052 media.



**Figure 3.12:** Western blot analysis using a polyclonal antibody against *hRAD51* to investigate the degradation events in soluble and insoluble protein fractions after cell lysis. *M*: marker; *S*: supernatant fraction after cell lysis; *P*: insoluble fraction after cell lysis. Cells were grown in TB-5052 and ZYM-5052 media at 20 °C for about 72 h and in LB medium at 37 °C and successively induced overnight at 20 °C. The black arrow indicates the full-length His-*hRAD51*.

These expression and degradation trials allowed selecting the optimal parameters for the expression of the full-length His-hRAD51 protein. Considering overexpression yield and level of protein degradation, and knowing that higher temperatures of induction can promote the association of overexpressed protein, the following optimal conditions for the expression of His-hRAD51 were identified: growth of Rosetta2(DE3)pLysS/pET15b-His-hRAD51 and induction at 20 °C in TB-5052 medium for 72 h.

The protocol for the purification of hRAD51 was optimized with two different chromatographic steps: an affinity chromatography (His-Trap column) followed by an anion exchange chromatography (ResQ column).

#### *Affinity chromatography*

The total protein extract obtained from *E. coli* Rosetta2(DE3)pLysS-His-hRAD51 culture, grown in TB-5052 at 20 °C for about 72 h, was loaded onto a His-Trap column. The elution chromatogram (Figure 3.13) is an asymmetric peak obtained by setting a linear imidazole gradient (50 mM - 500 mM).

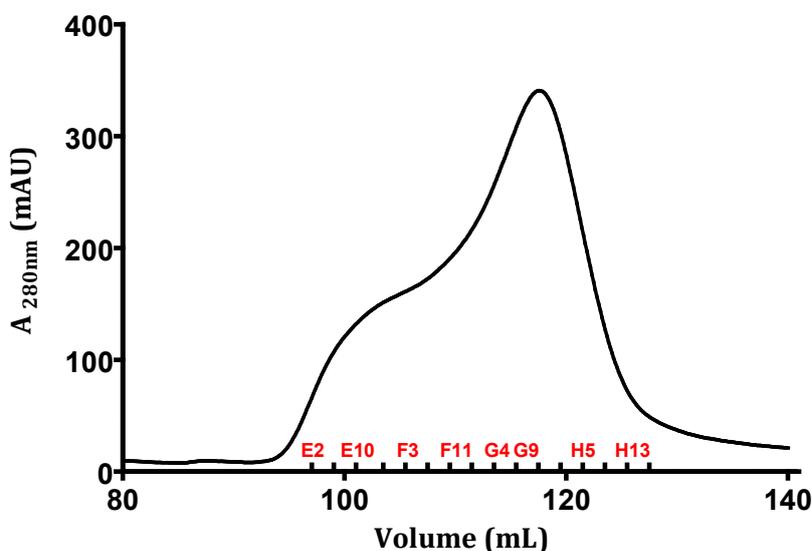
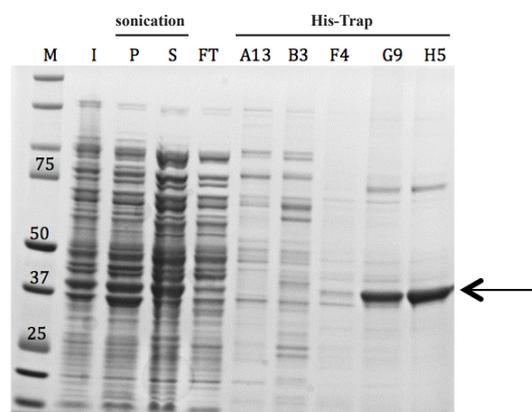


Figure 3.13: Chromatographic profile of the affinity chromatography elution.

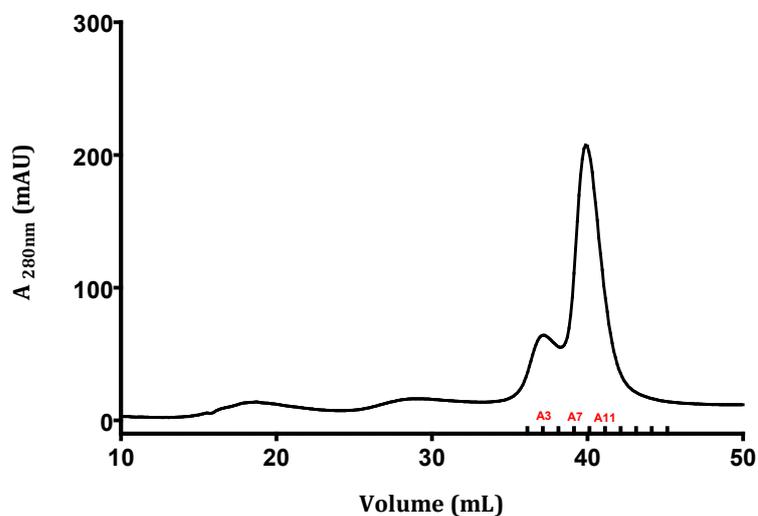
Fractions corresponding to the expression and the first purification step of His-hRAD51 were analyzed by SDS-PAGE (Figure 3.14). A band at 40 kDa, corresponding to His-hRAD51 protein, was detected in the insoluble and soluble protein fractions after cell lysis. A small amount of recombinant protein was also present in the column flow through. In the fractions of the main elution peak (G9 and H5) the band corresponding to the heterologous protein was present, whereas only a little amount of His-hRAD51 was detected in the shoulder (F4).



**Figure 3.14:** SDS-PAGE analysis of His-hRAD51 expression and first step of purification. I: induced sample; P: insoluble fraction after cell lysis; S: supernatant fraction after cell lysis; FT: flow through; A13 and B3: fractions corresponding to the peak eluted after a wash step with 50 mM imidazole; F4-H5: fractions corresponding to the shoulder and main peak of the chromatographic elution (His-Trap column). The black arrow indicates the full-length His-hRAD51.

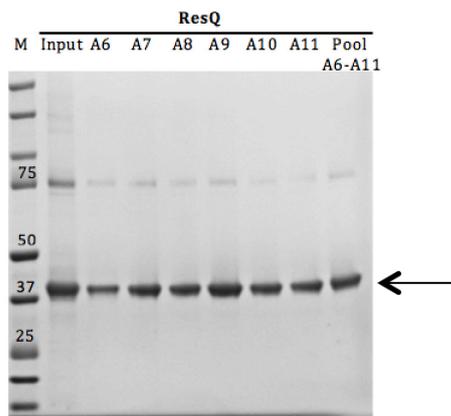
#### *Anion exchange chromatography*

The pooled fractions from the main His-Trap column elution peak were loaded onto a ResQ column. The elution (Figure 3.15), performed with a linear gradient from 200 mM to 1M KCl, generated a chromatogram with two main peaks (24 % and 50% of the linear gradient).



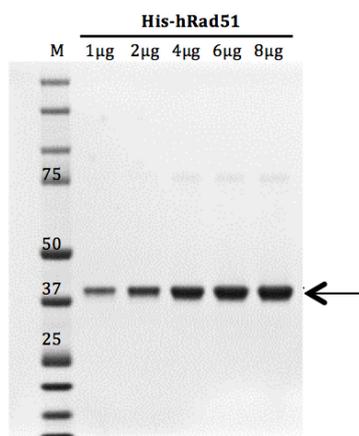
**Figure 3.15:** Chromatographic profile of the anion exchange column elution.

SDS-PAGE analysis (Figure 3.16) of the fractions eluted after anion exchange chromatography showed that the main elution peak (A6-A11 fractions) contained predominantly pure His-hRAD51 (40kDa).



**Figure 3.16:** SDS-PAGE analysis of fractions eluted from the anion exchange chromatography. *M*: marker; *input*: loading input of the ResQ chromatographic column; *A6-A11*: fractions of the ResQ column elution. The black arrow indicates the full-length His-hRAD51.

The purity of the obtained protein was further tested by concentrating the eluted protein and successively loading incremental quantities in a SDS-PAGE gel (Figure 3.17). Only one band corresponding to His-hRAD51 can be detected in all samples.



**Figure 3.17:** SDS-PAGE analysis of increasing quantities (1, 2, 4, 6 and 8 μg) of purified His-hRAD51. The black arrow indicates the full-length His-hRAD51.

The final protein yield using the pET15b-His-hRAD51 vector was 15-20 mg per liter of bacteria culture.

### **3.2.2 His-hRAD51 characterization**

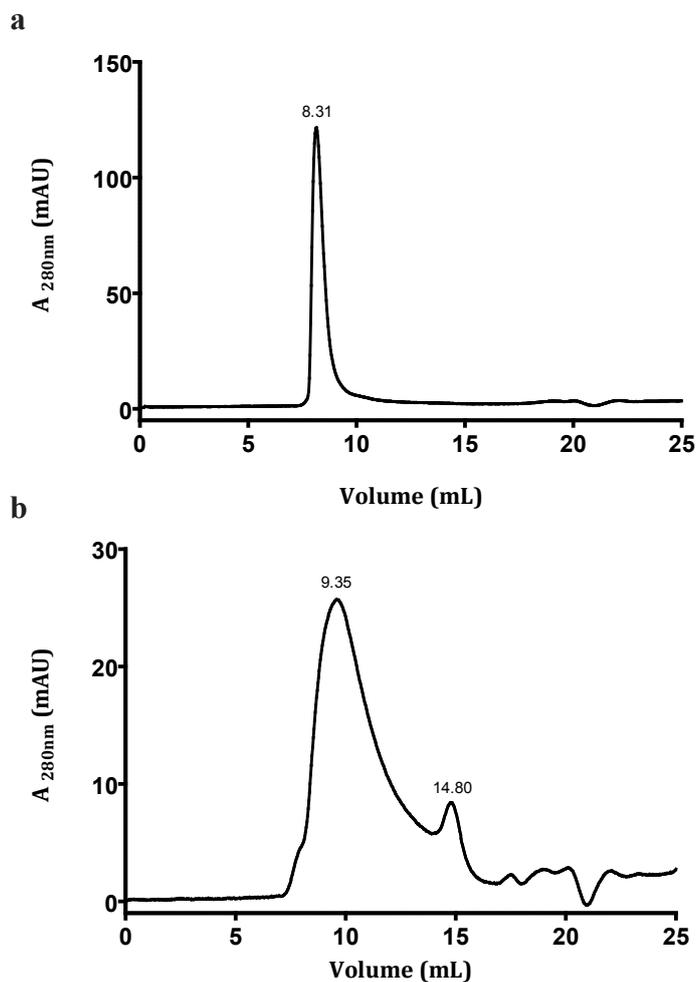
Purified His-hRAD51 was analyzed by gel filtration, dynamic light scattering and electron microscopy to assess the aggregation state of the purified protein.

#### *Size exclusion chromatography*

Size exclusion chromatographic experiments were performed in order to evaluate the state of His-hRAD51 oligomerization in the absence or presence of BRC4.

The size exclusion chromatogram of the purified His-hRAD51 corresponds to the elution of a high molecular weight oligomer (peak in the column dead volume; Figure 3.18a). One hour incubation at 37 °C of His-hRAD51 with BRC4 peptide in a 1:4 molar ratio suggests the progressive dissociation of His-hRAD51 fibril into smaller oligomers up to the monomeric form (narrow peak eluted at 14.8 mL; Figure 3.18b).

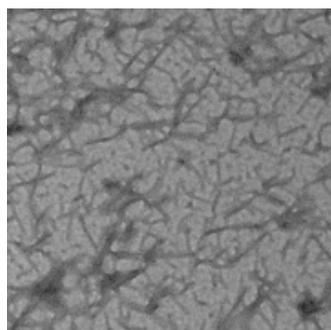
The intrinsic tendency of hRAD51 to self-assemble into heptameric rings (fibrils), through a polymerization motif 85-GGFTTATE-91 located at the end of the N-terminal region of the protein, is already known<sup>91</sup>. Likewise, the tendency of hRAD51 to monomerize in the presence of BRCA2 (BRC4) has already been reported in literature<sup>97</sup>. The latter mechanism, which is activated when the stoichiometric ratio between the two proteins (hRAD51:BRC4) is higher than 1:1, seems to be critical in the disassembly process of the hRAD51 fibril after DNA repair has been terminated<sup>97</sup>.



**Figure 3.18:** Superdex200 chromatogram of hRAD51 before (a) and after (b) incubation with BRC4 (1 h in 1:4 ratio).

### *Electron Microscopy*

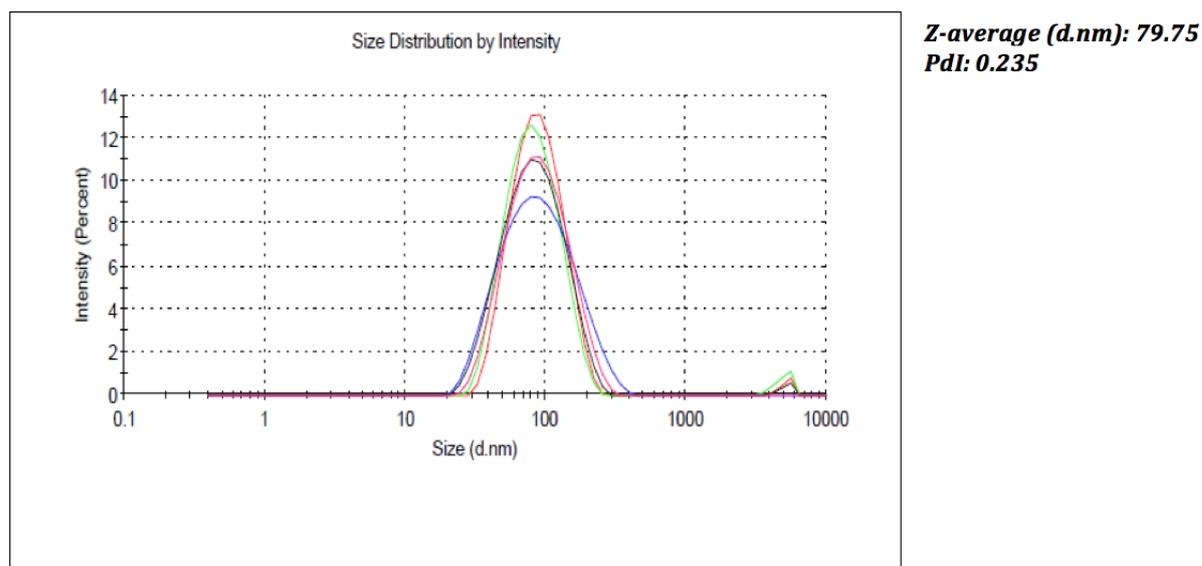
The electron microscopy (EM) image of freshly purified His-hRAD51, reported in Figure 3.19, shows the fibrillar structure of His-hRAD51. This result confirms the natural tendency of hRAD51 to self-aggregate into fibrils



**Figure 3.19:** EM-negative stain image of His-hRAD51 fibrillar structure.

*Dynamic Light Scattering (DLS)*

The His-hRAD51 oligomer size was determined by dynamic light scattering (Figure 3.20). Data collected at five successive time points and expressed in terms of the intensity distribution showed a particle size with a mean diameter of 79.75 nm at 25 °C. The polydispersity index (PDI) calculated was 23%. DLS suggests that size distribution of recombinant His-hRAD51 is monodisperse. Based on DLS data reported by Esashi *et al.*<sup>98</sup> for the monomeric hRAD51, the purified His-hRAD51, has a molecular diameter fourteen fold higher than hRAD51 monomer.



**Figure 3.20:** Particle size analysis of hRAD51 through DLS analysis. Data collected through five successive measurements.

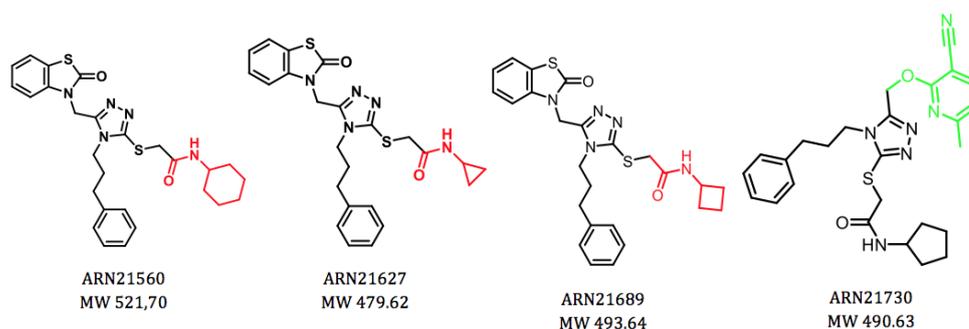
## Biophysical characterization of hRAD51-hit molecules interaction through Surface Plasmon Resonance (SPR), Nuclear Magnetic Resonance (NMR) spectroscopy and Microscale Thermophoresis (MST)

### 3.3 Biophysical characterization of hRAD51-small molecules interaction

The computational unit of our group, through a virtual screening campaign carried on the only available crystal structure of hRAD51 in complex with BRC4 (PDB 1N0W), has pursued the identification of potential inhibitors of the hRAD51-BRC4 interaction.

Compounds selected by virtual screening were synthesized by the chemistry group and further tested in biochemical and in cells assays.

One promising class of inhibitors contained a key triazole scaffold (Figure 3.21) in their structures. The chemistry unit performed a further structure activity relationship (SAR) analysis of the initial compounds. All the synthesized triazole derivative compounds were tested in biochemical assays designed to determine their ability to affect the hRAD51-BRC4 interaction<sup>99</sup>. The most promising inhibitors were also tested in combination with olaparib in nonmutated BRCA2 cells by our collaborators from the University of Bologna. The simultaneous administration of the two molecules (new inhibitors and olaparib) allows to trigger pharmacologically induced synthetic lethality showing anticancer activity comparable to the one observed when olaparib is administered to BRCA2-defective cells<sup>99</sup>.



**Figure 3.21:** Triazole compounds identified by virtual screening protocol on X-ray crystal structures of hRAD51-BRC4 (PDB: 1N0W).

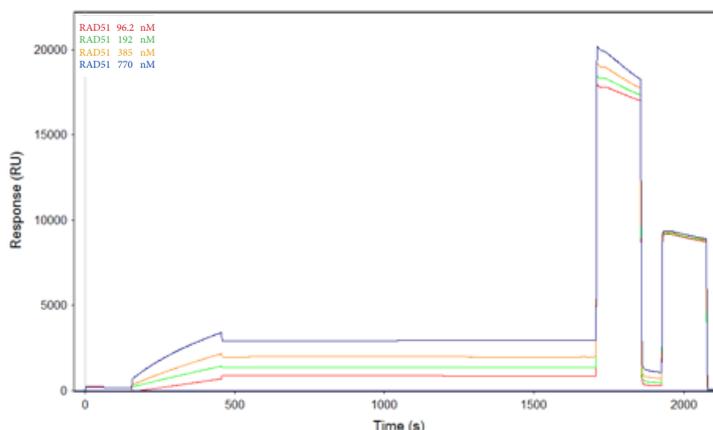
In order to confirm that the selected triazole inhibitors specifically target the protein-protein hRAD51-BRCA2 interaction, biophysical analyses were performed on the purified recombinant His-hRAD51 protein.

### 3.3.1 Surface Plasmon Resonance (SPR)

The most challenging step in a SPR experiment is frequently the setting up of a protocol for immobilization of the recombinant protein on a specific SPR sensor chip (see Appendix A4).

The first attempt that was performed was the immobilization of the His-RAD51 protein on a high density COOH5 chip, but the protein precipitated in the acetate buffer used to promote the interaction of His-hRAD51 to the sensor surface.

The second trial was performed using a His-Cap chip that is characterized by a oligoethylene glycol surface with nitrilotriacetic acid attached, which is activated by addition of nickel ions. His-hRAD51 binds to the sensor chip surface exploiting the hexahistidine tail located at the N terminal region of the protein. Four different concentrations of the protein were tested to achieve an optimal level of immobilization (Figure 3.22). Injecting His-hRAD51 770 nM, a response of 3000 RU was obtained, showing a stable immobilization of the protein on the sensor surface.



**Figure 3.22:** Protein immobilization optimization using a His-Cap chip. Four different concentrations of His-hRAD51 were tested: 96.2 nM (red curve), 192 nM (green curve), 385 nM (orange curve), 770 nM (blue curve). The His-Cap chip surface was regenerated with EDTA (350 mM) and imidazole (1 M).

Unfortunately, even though the experiment was successfully set up, the low solubility in aqueous buffer of the triazole molecules did not allow us to calculate their affinity and kinetic constant for the binding to His-hRAD51. The aqueous solubility of the triazole compounds, which have an average molecular weight of 490 Da, has been analytically calculated to be around 10  $\mu\text{M}$ . The minimal stock concentration that we needed to perform SPR experiments and to detect the maximal response predicted (40 RU) was approximately 500  $\mu\text{M}$ . The latter concentration was estimated from the  $\text{EC}_{50}$  values (about 50  $\mu\text{M}$ ) calculated for these molecules in the biochemical assays. At 500  $\mu\text{M}$  concentration triazole compounds were not soluble.

We further attempted to characterize the binding of triazole compounds to His-hRAD51 using different techniques, such as MST and NMR that usually require lower compounds concentrations. However, also in these cases, the low solubility of these molecules did not allow to achieve the minimal concentration in solution required to detect the interaction with hRAD51.

Given the solubility issues encountered with some of the compounds under investigation, we characterized the triazole compounds (identity, purity and solubility) in buffer solutions by NMR spectroscopy. Moreover, their aggregation state was verified by SPAM Filter<sup>100</sup> experiments. Each triazole compound has been tested at 100  $\mu\text{M}$  and 200  $\mu\text{M}$  theoretical concentrations. All the triazoles showed very low solubility in buffer (from 1  $\mu\text{M}$  up to 10  $\mu\text{M}$ ) and no aggregation effects were observed for this class of compounds. Due to the low solubility of the triazoles, it was not possible to detect their binding to His-hRAD51 by NMR.

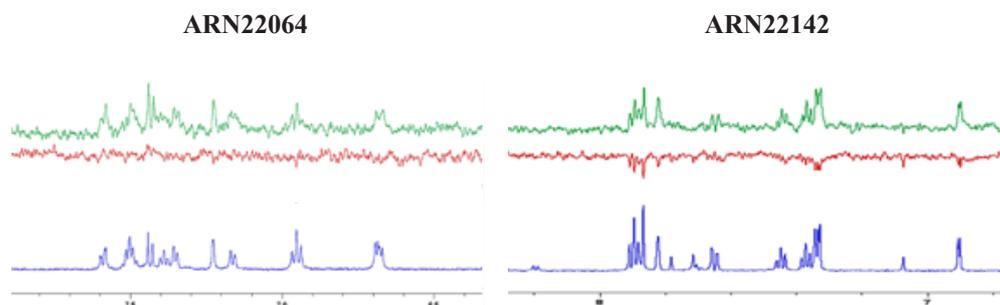
Unfortunately, the solubility issues that we encountered *in vitro* with this class of compounds has become a critical issue for the overall project, thus the urgency to investigate a different class of compounds became at this point quite pressing.

Starting from different scaffolds and performing different structure activity relationship (SAR) studies, our computational and medicinal chemistry groups have recently identified other two classes of potential hRAD51-BRCA2 inhibitors: quinolinone pyrazoline derivative and the carboxyphenyl furane derivative, which have a higher solubility in aqueous solutions than triazoles. ARN22064, a representative of the quinolinone pyrazoline derivative class has already been reported in a publication of Acker *et al.*<sup>101</sup>, while the ARN22142, a representative of the carboxyphenyl furane derivative class has been newly synthesized in our group. The NMR characterization of ARN22064 and ARN22142 confirms a good solubility for both compounds (higher than 200  $\mu\text{M}$ ) in buffer solution while aggregation effects were not observed.

### **3.3.2 NMR WaterLOGSY experiments**

WaterLOGSY<sup>102</sup> experiments were set up to test if the two new classes of compounds were *in vitro* able to bind the recombinant full-length His-hRAD51 protein.

WaterLOGSY experiments exploit the transfer of the bulk water magnetization to the chemical compound interacting with the protein (see Appendix A2.1). In a WaterLOGSY experiment the NMR signals of a chemical compound in solution are negative but, when the molecule interacts with a protein, its NMR signals become positive. We have tested ARN22064 and ARN22142 at 100  $\mu\text{M}$  concentration in the absence and presence of 1  $\mu\text{M}$  full-length His-hRAD51. Figure 3.23 shows the <sup>1</sup>H 1D NMR spectra (bottom blue) of the aromatic portions of ARN22064 (left) and ARN22142 (right) and the two WaterLOGSY experiments in the absence (red) and presence (green) of His-hRAD51: both molecules bind the protein, i.e. their NMR WaterLOGSY signals become positive in the presence of His-hRAD51.



**Figure 3.23:**  $^1\text{H}$  1D NMR spectra (blue) of ARN22064 and ARN22142. WaterLOGSY experiments in the absence (red) and presence (green) of His-hRAD51 show the binding of the two compounds to His-hRAD51. Compounds concentration 100  $\mu\text{M}$ , His-hRAD51 concentration 1  $\mu\text{M}$ .

### 3.3.3 Microscale Thermophoresis (MST)

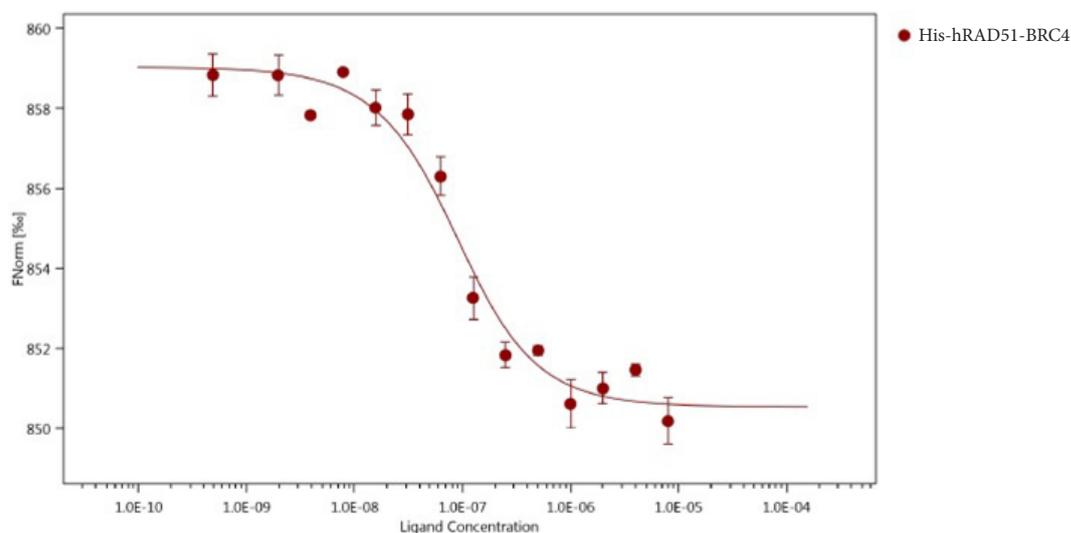
MST experiments were then performed in order to characterize and quantify the binding affinity of ARN22064 and ARN22142 for the full-length His-hRAD51 protein and to compare these data with the binding affinity of BRC4 for the protein (see Appendix A3).

#### *His-hRAD51-BRC4 binding*

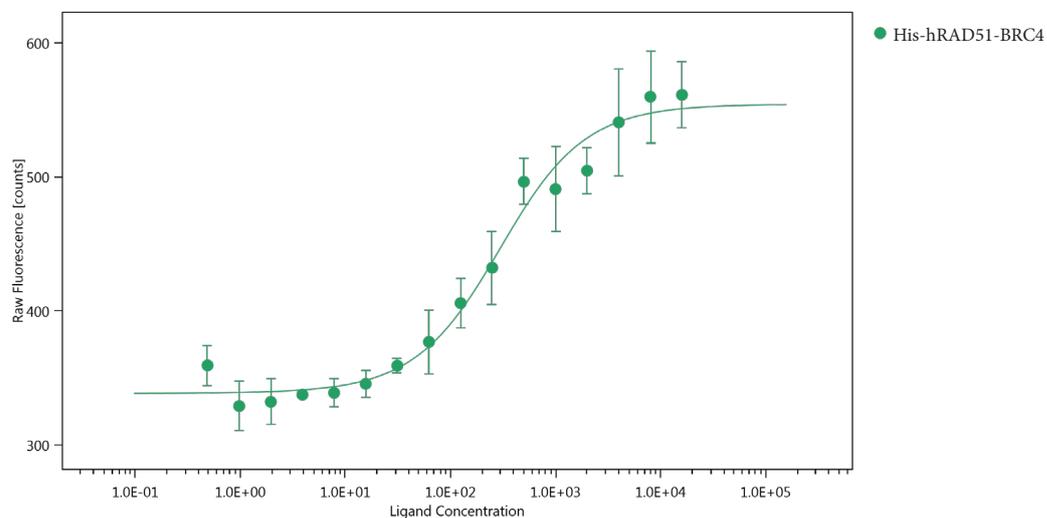
In a first set of MST experiments His-hRAD51 was labelled with the NT-647 dye, which specifically binds the hexahistidine tail at the N terminal of His-hRAD51. NT-647 is a red fluorescent dye with excitation and emission maxima at approximately 650 and 670 nm, respectively. A constant concentration of NT-647-labelled His-hRAD51 (100 nM) was used to characterize the protein binding to the BRC4 peptide. A  $K_d$  of  $63.30 \pm 17.98$  nM (Figure 3.24) was determined.

The obtained affinity data were further confirmed by using a constant concentration of NT-495-labelled His-hRAD51 (80 nM) (Figure 3.25). NT-495 is a blue fluorescent dye with excitation and emission maxima at approximately 493 and 521 nm, which is covalently bound to primary amines, mainly lysine residues of His-hRAD51 through amine coupling reactions.

The physiological hRAD51-BRC4 interaction is, as expected, strong.



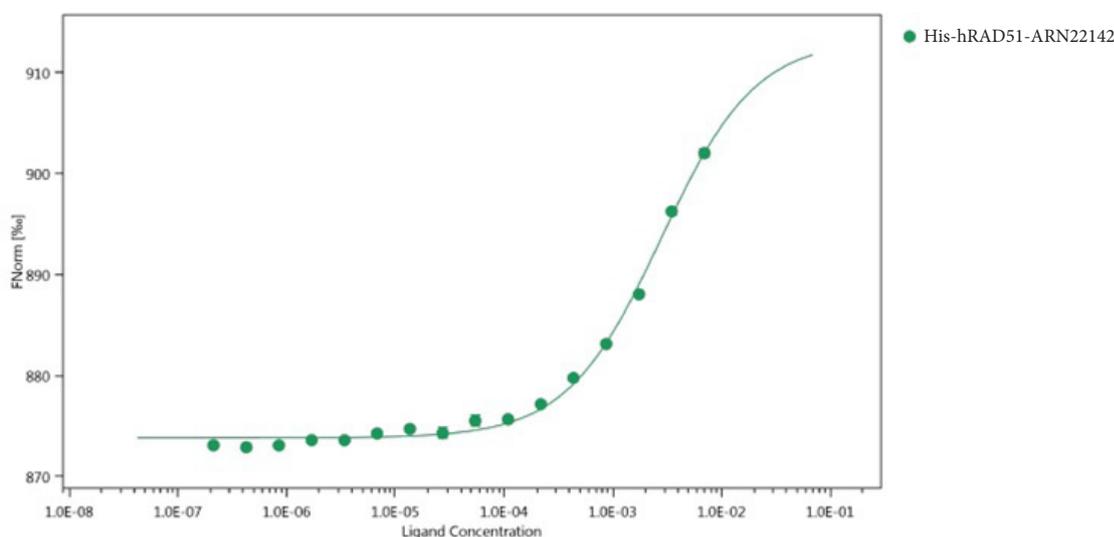
**Figure 3.24:** MST analysis of His-hRAD51-BRC4 binding. Titration curve of (NT-647)-His-hRAD51 (100 nM) with increasing concentrations of the BRC4 peptide. Sigmoidal fitting curves were obtained using the Affinity Analysis software of Nanotemper Technologies. MST data are the average of three replicates.



**Figure 3.25:** MST analysis of His-hRAD51-BRC4 binding. Titration curve of (NT-495)-His-hRAD51 (80 nM) with increasing concentrations of the BRC4 peptide. Sigmoidal fitting curves were obtained using the Affinity Analysis software of Nanotemper Technologies. MST data are the average of three replicates.

### *His-hRAD51-ARN22142 binding*

In a second set of experiments, a constant concentration of NT-647-labelled His-hRAD51 (100 nM) was used to characterize the binding of the ARN22142 compound to the protein. Although the solubility of ARN22142 (6 mM) prevented us from using higher compound concentration and therefore to obtain a complete binding curve, it was possible to estimate a minimal  $K_d$  value of  $2.71 \pm 0.30$  mM from the fitting of the collected data (Figure 3.26).

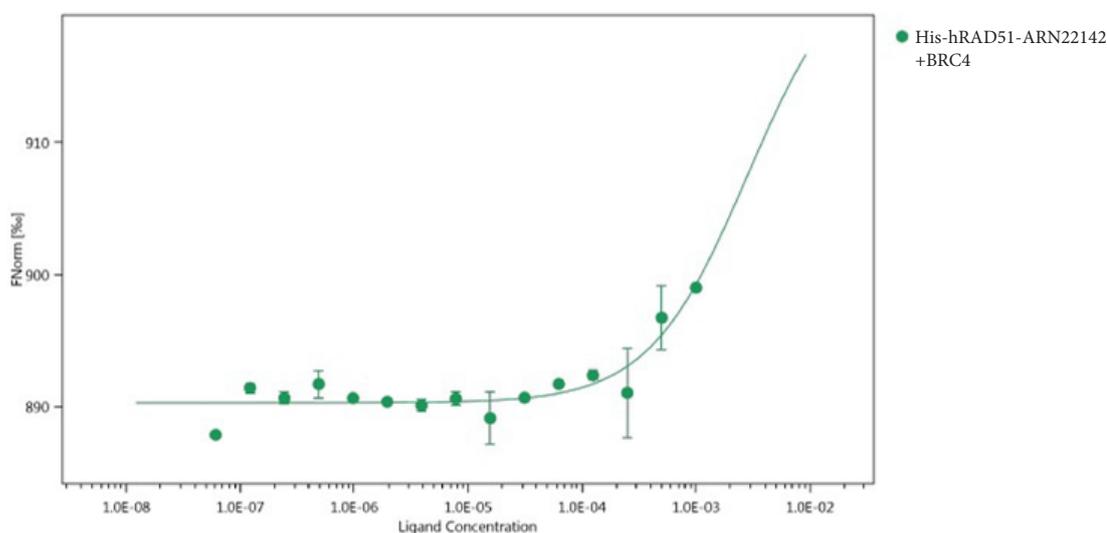


**Figure 3.26:** MST analysis of His-hRAD51-ARN22142 binding. Titration curve of (NT-647)-His-hRAD51 (100 nM) with increasing concentrations of ARN22142. Sigmoidal fitting curves were obtained using the Affinity Analysis software of Nanotemper Technologies. MST data are the average of two replicates.

A displacement assay was set up in order to verify if ARN22142 was able to prevent or interfere with His-hRAD51-BRC4 interaction. NT-647-labelled His-hRAD51 was pre-incubated with the highest soluble ARN22142 concentration in order to promote His-hRAD51-ARN22142 complex formation. BRC4 was successively titrated into the complex to verify its ability to displace the compound. From previous MST experiments, 16  $\mu$ M BRC4 was able to saturate His-hRAD51 (100nM) binding sites (see Figure 3.24 and 25); using the same BRC4 concentration to displace ARN22142 from His-hRAD51, no binding was detected (data not shown).

Thus, another displacement trial using a higher BRC4 concentration (1 mM) was

performed. Figure 3.27 shows that BRC4 binds His-hRAD51. The data obtained were considered sufficient evidence of displacement even though the binding curve was not complete. The complete binding curve would have required a significant amount of peptide without providing substantial additional information.

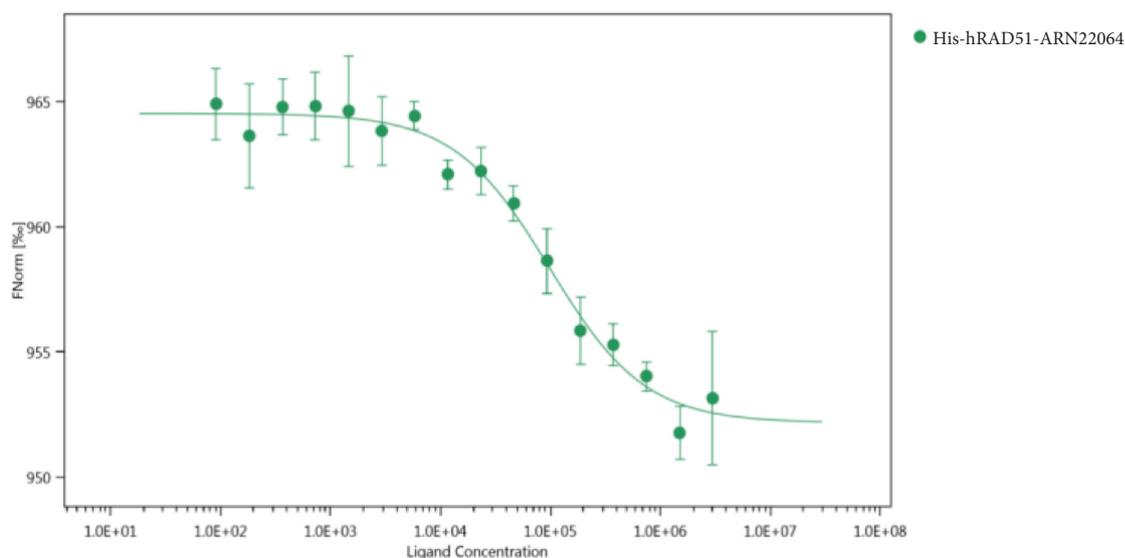


**Figure 3.27:** MST displacement assay. Titration curve of (NT-647)-His-hRAD51 (100 nM)-ARN22142 complex with increasing concentration of BRC4. Sigmoidal fitting curves were obtained using the Affinity Analysis software of Nanotemper Technologies. MST data are the average of three replicates.

As reported above, ARN22142 weakly interacts with His-hRAD51 ( $K_d$   $2.71 \pm 0.30$  mM). Since pre-incubation of hRAD51 with ARN22142 induces a significant positive shift of the BRC4-hRAD51 titration curve and therefore a decreased binding affinity of BRC4 for hRAD51, we assume that BRC4 displaces ARN22142 that binds hRAD51 in the same pocket. These data confirm the virtual screening analysis initially performed with this compound on the hRAD51-BRC4 crystal structure available (1N0W).

### *His-hRAD51-ARN22064 binding*

A constant concentration of NT-495-labelled His-hRAD51 (80 nM) was used to characterize the binding of the ARN22064 compound to the protein (Figure 3.28). A  $K_d$  of  $98.81 \pm 16.99 \mu\text{M}$  was determined, suggesting that ARN22064 compound weakly interacts with His-hRAD51.



**Figure 3.28:** MST analysis of ARN22064 binding to His-hRAD51. Titration curve of (NT-495)-His-hRAD51 (80 nM) with increasing concentrations of the ARN22064. Sigmoidal fitting curves were obtained using the Affinity Analysis software of Nanotemper Technologies. MST data are the average of three replicates.

The BRC4-His-hRAD51 binding (binding affinity in the nanomolar range) is stronger compared to the bindings of the two synthesized compounds to the protein (binding affinities in the micro and millimolar ranges). This result is quite straightforward comparing a “physiological” protein-protein interaction to a synthetic molecule-protein interaction. Nevertheless, as reported above, the preventive binding of one of the molecules is able to weaken the protein-protein interaction therefore negatively affecting the target His-hRAD51-BRC4 interaction.

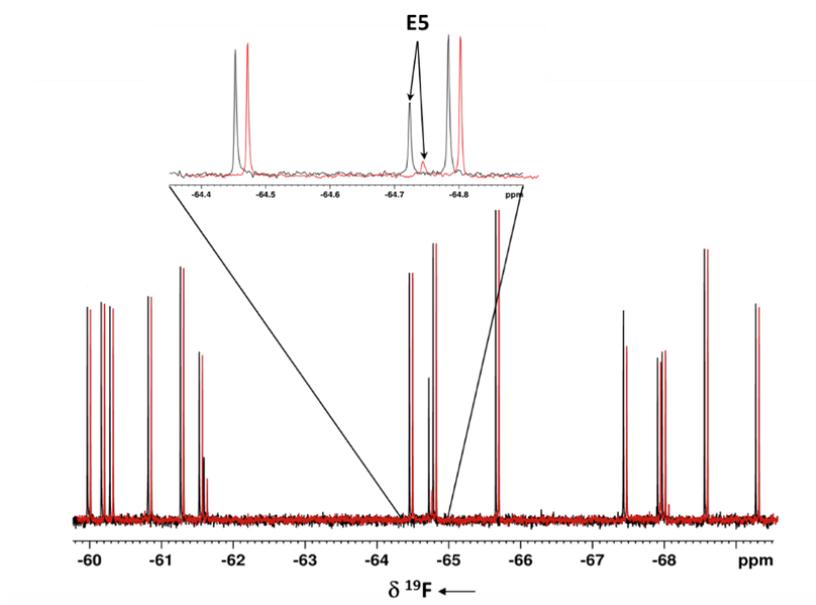
### **3.4 Biophysical characterization of hRAD51-fragments interaction**

#### **3.4.1 NMR Fragment Based screening**

In parallel to the classical drug-like screening approach, above described, we decided to also pursue a fragment based approach (FBA) through NMR spectroscopy (Appendix A2). In Italian Institute of Technology (IIT) we have a home made local Environment of Fluorine (LEF) library<sup>103</sup> constituted by about 900 fluorinated compounds with a molecular weight ranging from 100 to 350 Da and solubility in buffer higher than 400  $\mu\text{M}$ .

We performed a preliminary direct  $^{19}\text{F}$ -NMR ligand-binding screening against full-length His-hRAD51, using a portion of 185 fragments of our LEF Library. To identify the possible binders, we recorded transverse  $R_2$  filter fluorine experiments. The transverse relaxation rate  $R_2$  is a very sensible parameter for these studies, due to the large Chemical Shift Anisotropy (CSA) of  $^{19}\text{F}$  nucleus and to the large exchange contribution. The observed response ( $R_{\text{obs}}$ ) in the  $R_2$  filter experiments results in a line-broadening of the compound  $^{19}\text{F}$ -NMR signal when it binds to the target protein.

The fragments have been screened in mixtures of 20-25 compounds each, at the concentration of 40  $\mu\text{M}$ , in the presence and absence of 1  $\mu\text{M}$  full-length His-hRAD51. The knowledge of the chemical shift of the molecules forming the mixtures has allowed us to easily identify the binders. An example of the  $^{19}\text{F}$ -NMR screenings performed is reported in Figure 3.29 where the  $^{19}\text{F}$ -NMR signals of compound mixtures in the absence of protein (black) and in presence of His-hRAD51 (red) are reported. Most of the signals do not show any significant change in the presence of the protein and therefore they do not interact with His-hRAD51. On the contrary, in the zoomed region, compound E5 shows a strong broadening of its  $^{19}\text{F}$  NMR signals in the presence of His-hRAD51, suggesting that E5 binds the protein. Out of the 185 tested fragments, ten hits were identified.

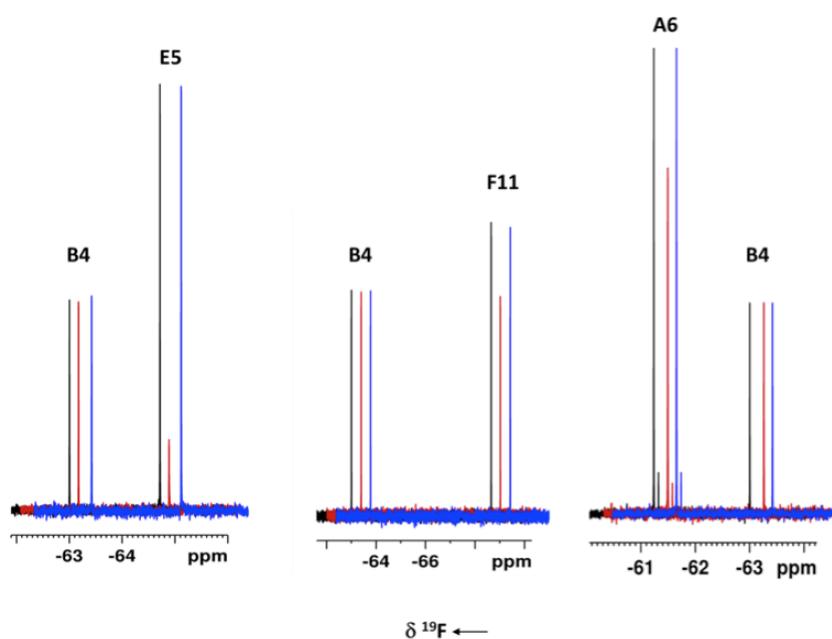


**Figure 3.29:**  $^{19}\text{F}$ -NMR screening of compound mixtures in the presence (red) and in the absence of His-hRAD51 (black).

The binding was also confirmed testing the identified binders (hits) as single compounds in the same experimental conditions described above in the presence of a non-binder as internal negative control.

In order to verify that the identified hits bind His-hRAD51 in the same pocket as BRC4,  $^{19}\text{F}$ -NMR competition binding experiments were performed. The hits were tested in the absence and in the presence of protein and after the addition of 20  $\mu\text{M}$  BRC4. The presence of competition between the compounds and BRC4 results in a sharpening of the  $^{19}\text{F}$ -NMR signal of the hit after BRC4 addition to the compound-His-hRAD51 mixture. Out of the ten hits identified by FBA, three fragments have been displaced by BRC4 (Figure 3.30). The NMR signals of compound B4 (negative control) do not show any variation in the presence of His-hRAD51 (red) as well as after the addition of BRC4 (blue) confirming that it does not interact with any of the two proteins. On the other end, the  $^{19}\text{F}$  signal of the three compounds E5, F11 and A6, shows a

significant line broadening effect and a decrease in signal intensity in the presence of His-hRAD51 (red). The latter signals become sharp again after the addition of BRC4 suggesting a competition between the compound and the peptide for the binding to His-hRAD51.

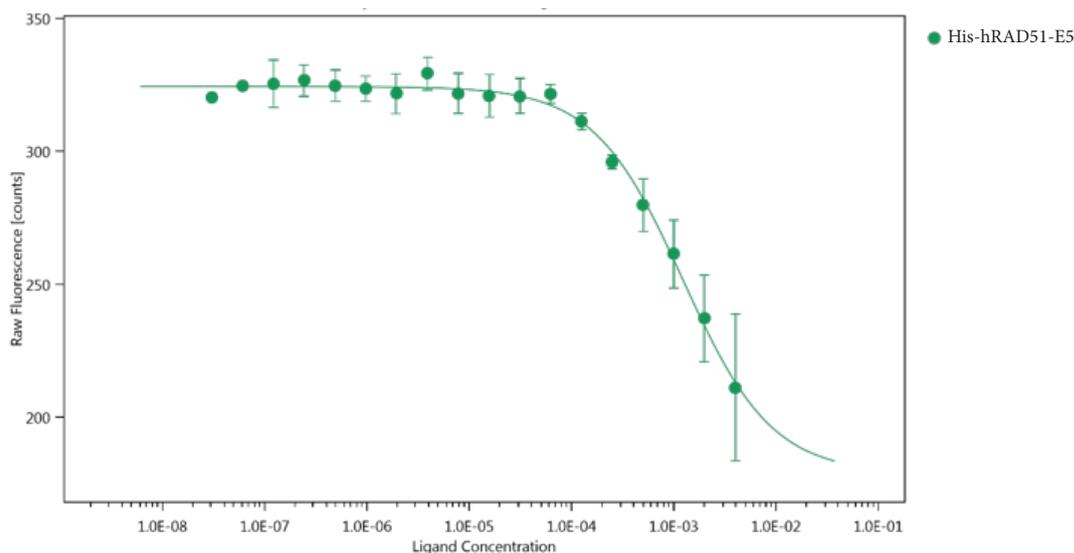


**Figure 3.30:**  $^{19}\text{F}$ -NMR competition binding experiment. E5 (left), F11 (middle) and A6 (right) were tested in the absence (black) and in the presence (red) of His-hRAD51, and after addition of BRC4 (blue) to the compound-His-hRAD51 mixture. B4 was used as negative control.

### 3.4.2 Microscale Thermophoresis analysis of the hit fragments

Microscale Thermophoresis analyses were performed in order to characterize the interaction between the fragments E5, F11 and A6 and His-hRAD51. A constant concentration of NT-647 labelled His-hRAD51 (50 nM) was used to determine the compounds binding to the protein. Addition of F11 and A6 fragments to the protein did not induce any MST signal change, suggesting that these fragments do not interact with His-hRAD51 at the concentration tested (4 mM). On the contrary, MST analysis of the E5 fragment binding to His-hRAD51 clearly

confirmed the interaction of the small compound with the protein. A  $K_d$  of  $1.23 \pm 0.13$  mM was calculated, suggesting that E5 weakly interacts with His-hRAD51 (Figure 3.31).

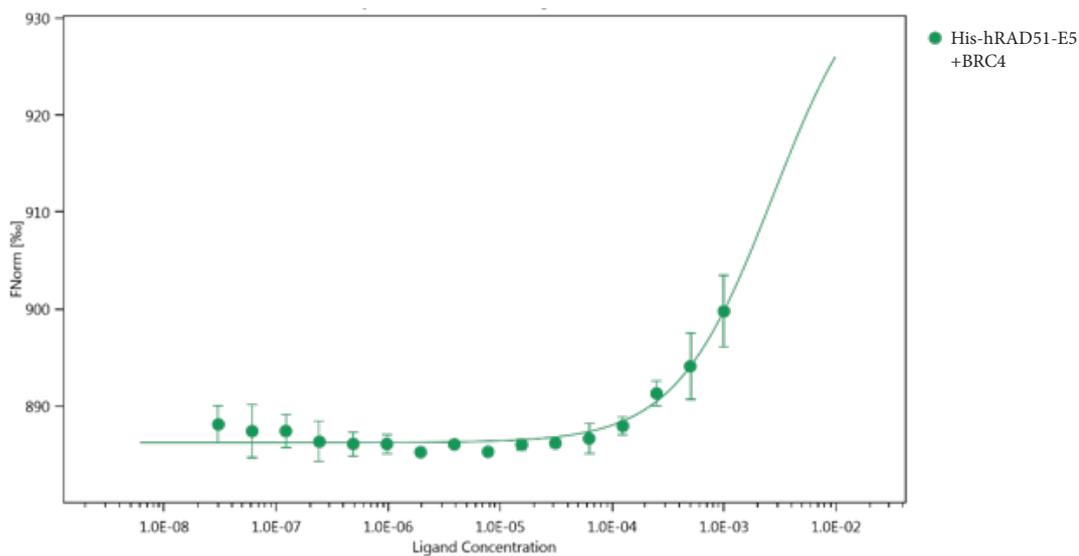


**Figure 3.31:** MST analysis of His-hRAD51-E5 binding. Titration curve of (NT-647)-His-hRAD51 (100 nM) with increasing concentrations of E5. Sigmoidal fitting curves were obtained using the Affinity Analysis software of Nanotemper Technologies. MST data are the average of three replicates.

A displacement assay was set up to verify if the E5 fragment was able to prevent or interfere with His-hRAD51-BRC4 binding. The titration of BRC4 into the complex His-hRAD51-E5 determined the displacement of the fragment as shown in Figure 3.32, confirming NMR data.

A significant amount of peptide would have been necessary in order to obtain the complete binding, without providing substantial additional information therefore we considered the reported data as clear evidence of the displacement.

The data show a decrease of the binding affinity of the peptide for His-hRAD51 compared to the same experiment in the absence of fragment E5, suggesting that the fragment most likely binds in the same protein pocket as BRC4.



**Figure 3.32:** MST displacement assay. Titration curve of (NT-647)-His-hRAD51 (100 nM)-E5 complex with increasing concentration of BRC4. Sigmoidal fitting curves were obtained using the Affinity Analysis software of Nanotemper Technologies. MST data are the average of three replicates.

### **X-ray structural characterization of hRAD51-BRCA2**

In order to structurally characterize the full-length hRAD51, the His-hRAD51 protein purified as described above, was concentrated up to 5 - 10 mg/mL for crystallization screenings, which were carried out as described in the Material and Methods section. The conditions tested up to now did not allow to obtain crystals with good diffraction patterns. Even though from DLS data the purified His-hRAD51 sample is monodisperse, the fibrillar nature of the protein can be one of the possible reasons that make the crystallization procedure quite challenging.

In order to promote protein crystallization, we looked for a condition able to produce a monomeric, stable form of hRAD51. As we also reported above, the interaction of His-hRAD51 with BRC4 in higher stoichiometry than 1:1 induces hRAD51 monomerization. Moreover, it has been reported that mutation of phenylalanine 86 residue in glutamic acid (F86E), in hRAD51, can determine the production of monomeric protein, since it removes one of the key residues for protomer-protomer interaction (FXXA)<sup>51</sup>.

Therefore, hRAD51 or hRAD51 F86E in complex with BRC4 can be suitable samples to obtain the crystal structure of the full-length protein and, at the same time, to characterize the interactions between the two proteins.

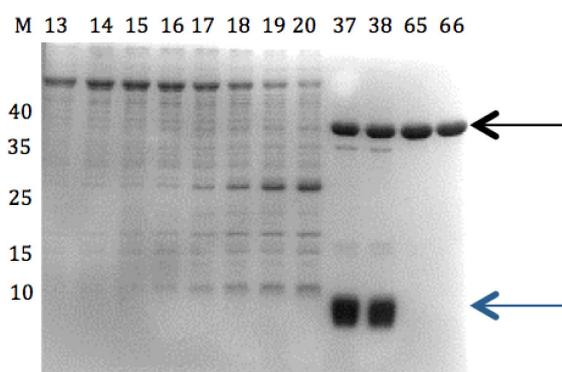
### **3.5 Expression and purification of hRAD51-BRC4 complex**

The hRAD51-BRC4 complex was obtained by co-expressing the two proteins, i.e. by simultaneously transforming the two cDNAs with *E. coli* Rosetta2(DE3)pLysS cells in the presence of kanamycin and ampicillin antibiotics (Materials and Methods). The protocol for the purification of the hRAD51-BRC4 complex was optimized with two chromatographic steps: an affinity chromatography (His-Trap column) and a size exclusion chromatography (Superdex200 column).

*Affinity chromatography*

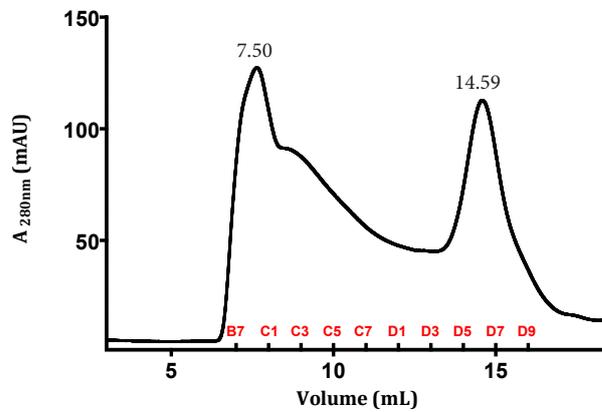
The total bacterial culture proteins extract, obtained after cells sonication, was loaded onto a His-Trap column. A linear gradient of imidazole (10 mM - 500 mM) in the elution buffer led to three elution chromatographic peaks.

Fractions eluted from the affinity chromatography were analyzed by SDS-PAGE gel (fig. 3.33). While the first eluted peak (fractions 13-20) contained mainly non-specific bound proteins, the second one (fractions 37 and 38) contained both hRAD51 and BRC4, and the last peak (fractions 65 and 66) was mainly constituted by hRAD51 alone.



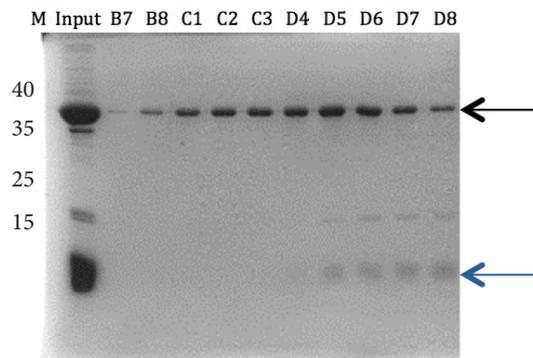
**Figure 3.33:** SDS-PAGE of fractions eluted from the affinity chromatography. 13-20: fractions corresponding to the first eluted peak of the His-Trap column. 37-38: fractions corresponding to the second eluted peak of the His-Trap column. 65-66: fractions corresponding to the third eluted peak of the His-Trap column. The black arrow indicates the full-length hRAD51. The blue arrow indicates the BRC4 peptide.

The fractions containing the hRAD51-BRC4 complex were pooled, concentrated and successively loaded onto a Superdex200 column. The resulting chromatogram (Figure 3.34) showed a heterogeneous population with two dominant peaks: one in the column dead volume (7.5 ml) and the second one at the predicted molecular weight for the monomeric hRAD51-BRC4 complex (14.59mL).



**Figure 3.34:** Chromatogram of the Superdex200 column elution profile.

The eluted fractions were analyzed by SDS-PAGE (Figure. 3.35). As expected, the monomeric peak (fractions D5-D8) corresponds to the hRAD51-BRC4 complex, while the rest of the chromatogram elution can be ascribed to hRAD51 alone in different oligomerization states.



**Figure 3.35:** SDS-PAGE analysis of fractions eluted from gel filtration. Input: input loading sample; B7-D8: fractions of Superdex200 column elution. The black arrow indicates the full-length hRAD51. The blue arrow indicates the BRC4 peptide.

The co-expression of hRAD51 and BRC4 allowed obtaining a heterogeneous pool of hRAD51 with different molecular weights and the heterodimeric hRAD51-BRC4 complex was only partially formed.

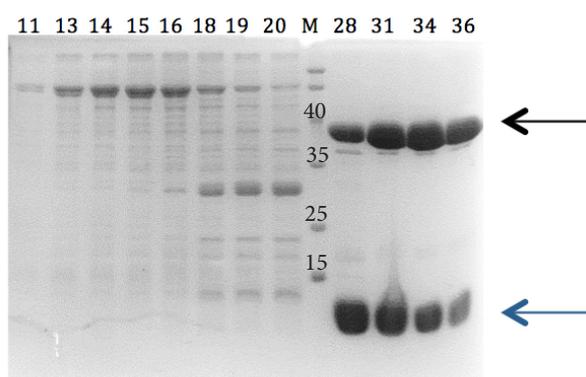
### 3.6 Expression and purification of hRAD51 F86E-BRC4 complex

A further attempt to obtain a homogeneous hRAD51-BRC4 complex led us to express and purify the hRAD51 F86E-BRC4 complex. The two proteins were co-expressed as described above for the non-mutated co-expressed complex. The purification protocol was again optimized in two chromatographic steps: an affinity chromatography (His-Trap column) and a size exclusion chromatography (Superdex200 column).

#### *Affinity chromatography*

The total protein extracted from a *E. coli* bacterial culture was loaded onto a His-Trap column. The elution was performed using a linear gradient from 10 mM to 500 mM imidazole resulting in the elution of two separate peaks.

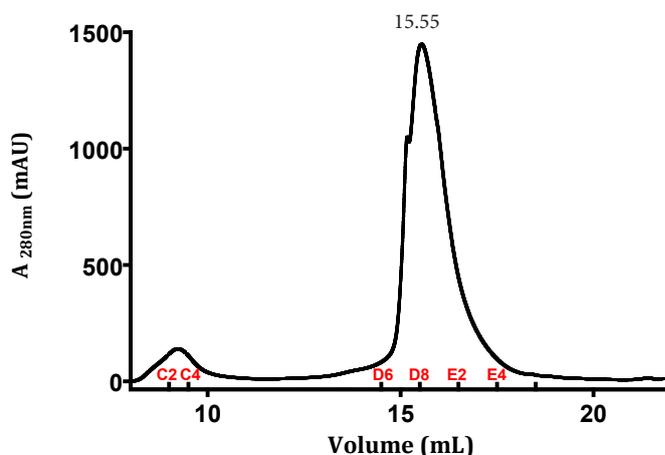
Fractions corresponding to the two eluted peaks were analyzed by SDS-PAGE (Figure 3.36). Only in the second peak, eluted at the highest imidazole concentration (fractions 28-36), the two proteins, hRAD51 and BRC4, are simultaneously present.



**Figure 3.36:** SDS-PAGE analysis of fractions eluted from the affinity chromatography. M: marker; 11-20: fractions corresponding to the first eluted peak of the His-Trap column; 28-36: fractions corresponding to the second eluted peak of the His-Trap column. The black arrow indicates the full-length hRAD51. The blue arrow indicates the BRC4 peptide.

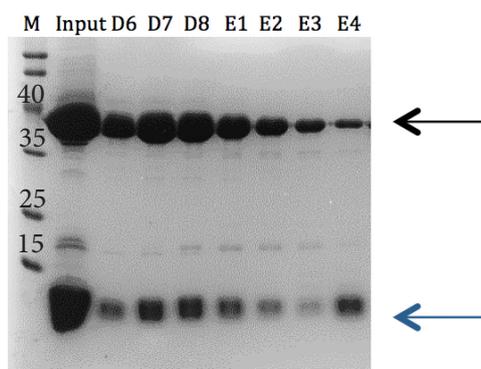
*Size exclusion chromatography*

The fractions containing the hRAD51 F86E-BRC4 complex were pooled, concentrated and successively loaded onto a Superdex200 column. The chromatogram reported in Figure 3.37 evidences the presence of a small peak eluted in the dead volume of the column (fractions C2-C4), and a main peak eluted at 15.55 mL.



**Figure 3.37:** Chromatogram corresponding to the Superdex200 column elution profile.

Fractions corresponding to the main eluted peak were analyzed by SDS-PAGE (Figure 3.38) confirming that the hRAD51 F86E-BRC4 complex was pure and that the F86E mutation was critical to obtain a hRAD51 monomeric protein co-expressed with BRC4.



**Figure 3.38:** SDS-PAGE analysis of fractions eluted from Superdex200 column. M: marker; Input: input loading; D6-E4: fractions of Superdex200 column elution. The black arrow indicates the full-length hRAD51. The blue arrow indicates the BRC4 peptide.

### **3.7 Crystallization trials of hRAD51 F86E-BRC4 complex**

Different crystallization trials were performed to obtain crystals from the purified hRAD51 F86E-BRC4 complex.

The first trial was performed concentrating the complex up to 26 mg/mL and using three different crystallization screening kits in 96 wells plates (see Materials and Methods).

Since this first protein batch resulted to be too concentrated, i.e. significant protein precipitation in most of the crystallization wells was observed, the successive protein batches were concentrated only up to 13.5 mg/mL.

The protein storage buffer was also modified in order to improve complex stability. The buffer of the storage solution (Hepes, Tris, etc.) was modified as well as the NaCl concentration in the buffer.

Different crystallization trials were performed using several commercial kits (Materials and Methods). Basic crystallization kits for sparse matrix screenings were taken into consideration as well as more specific kits for biological macromolecules or macromolecular complexes.

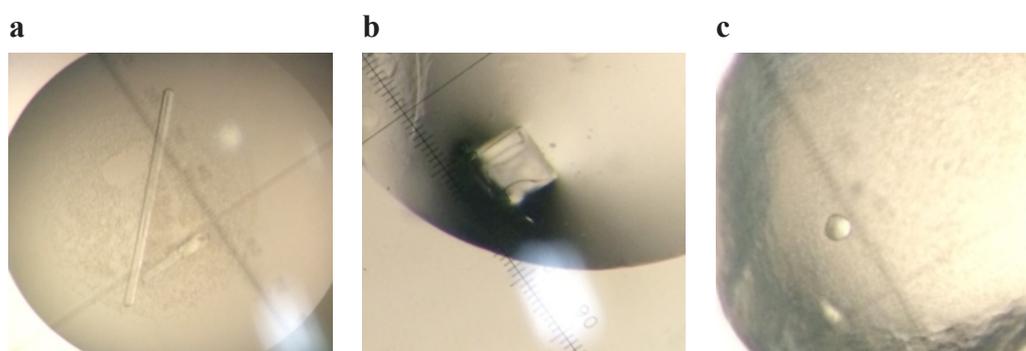
Several weeks after sample plating large single crystals appeared in the following conditions:

- 0.1 M MES monohydrate pH 6.5, 1.6 M Magnesium sulfate heptahydrate (crystal Screen 2, condition 68, Hampton Research; Figure 3.39a)
- 0.2 M Potassium chloride, 0.05 M HEPES pH 7.5, 35% v/v Pentaerythritol propoxylate (5/4 PO/OH) (Index screen, condition 56, Hampton Research; Figure 3.39b)
- 0.2 M Sodium sulfate decahydrate, 20% w/v Polyethylene glycol 3,350 (PEG/Ion screen, condition 33, Hampton Research; Figure 3.39c)

We do not know if the obtained crystals are salts or protein crystals, even though the long time taken for crystals formation (several weeks) and the crystals shapes suggest that they are protein crystals rather than salts. Moreover, even though the crystals morphology and sizes

are promising, we will find out if the crystals have enough long range internal order to allow the collection of good diffraction patterns to solve the complex structures only at the end of December, when they will be taken to the synchrotron for data collection.

We are anyway simultaneously screening a variety of different conditions to look for completely different crystallization set up that may result in crystals with different morphology and eventually diffraction patterns.



**Figure 3.39:** Optical microscope pictures of crystals grew from **(a)** crystal Screen 2 kit, condition 68 (Hampton Research); **(b)** Index screen, condition 56 (Hampton Research) and **(c)** PEG/Ion screen, condition 33 (Hampton Research).

---

---

Chapter four

# Conclusions and perspectives

---

---

Human RAD51 (hRAD51) is an ATP dependent recombinase that plays a key role in the homologous recombination pathway via interaction with BRCA2. Mutations in BRCA2 protein impair its ability to bind hRAD51, leading to incompetent HR repair pathway. Recently, FDA has approved olaparib, a PARPi, for the treatment of ovarian cancer in BRCA2 defective patients. This therapeutic approach is based on the concept of synthetic lethality, which is a lethal phenotype deriving from a combination of two gene mutations. Thus, in BRCA2 defective oncology patients, the administration of PARPi allows the simultaneous inhibition of two independent repair mechanisms, which result in an increase in cancer cells sensitivity to chemo- or radio-therapy.

The key idea of our project is to trigger the synthetic lethality process through the administration of two different small molecules. This strategy can be used for the treatment of cancer in patients carrying nonmutated genes or when PARPi resistance is developed in BRCA2 mutated patients. Indeed, in the latter case, resistance is often due to restoration of nonmutated BRCA2 or overexpression of hRAD51, which is reported in several tumors (i.e. breast, ovarian and prostate cancers).

Herein we optimized a protocol for the expression and purification of full-length hRAD51 in order to characterize the binding of hRAD51 and BRCA2 as well as of hRAD51 and inhibitor molecules through biophysical techniques.

Initially, *E. coli* BL21(2DE3)pLysS cells were transformed using the pGEX-4T-1-GST-hRAD51 construct. Once the expression protocol was optimized, the protein was purified by two chromatographic steps: an affinity chromatography (GST-Trap column) and a size exclusion chromatography (Superdex200 column). When applied onto a Superdex200 column, GST-hRAD51 was eluted in the column dead volume, suggesting that it was purified as an

oligomer with high molecular weight. This is in agreement with the tendency of hRAD51 to self-aggregate. However, several issues occurred with this protocol: a significant amount of recombinant protein was present in the insoluble fraction after cell lysis, GST-hRAD51 was produced in low yield (1 mg of pure protein per liter of bacterial culture) and the recombinant protein was not pure enough for further biophysical analyses. HPLC-MS analyses also showed that purity and yield of GST-hRAD51 were significantly affected by proteolytic degradation mediated by endogenous proteases (ClpX and HslU-V proteases) of bacterial cells. Attempts to inhibit or decrease the proteolytic degradation did not allow to avoid GST-hRAD51 fragmentation. The expression and purification strategy was modified in order to produce a full-length pure hRAD51 protein in high yield.

The gene coding for hRAD51 was subcloned from the pGEX-4T-1 vector to the pET15b vector to obtain a recombinant protein with a His-tag at the N terminal region. *E. coli* Rosetta2(DE3)pLysS cells were transformed using the new construct. The protocol for His-hRAD51 expression was optimized growing bacterial cells in TB-5052 medium for 72 h at 20 °C. The protein was then purified with two different chromatographic steps: an affinity chromatogram (His-Trap column) and an anion exchange chromatography (ResQ column) that allowed to obtain a pure hRAD51 in high yield (20 mg of recombinant protein per liter of bacterial culture). As expected, when applied onto a Superdex200 column, the protein was eluted in the column dead volume. DLS analysis suggested that the size distribution of the recombinant hRAD51 is monodisperse calculating a hydrodynamic radius of about 40 nm, a value fourteen times higher than the one of the monomeric form<sup>98</sup>. Furthermore, EM (negative staining) investigations on the pure recombinant protein confirmed the intrinsic tendency of hRAD51 to self-aggregate into fibrillar structures of high molecular weight.

Compounds selected through a virtual screening campaign carried on the available crystal structure of hRAD51 in complex with BRC4 (PDB 1N0W) were synthesized by the chemistry group and further tested in biochemical and in cells assays. One promising class of hRAD51-BRCA2 inhibitors was constituted by molecules with a key triazole scaffold<sup>99</sup>. These compounds, tested on HR-proficient (BxPC-3) and -deficient (Capan-1) cells, showed their ability to trigger synthetic lethality when administered in combination with a PARPi (olaparib). The EC<sub>50</sub> calculated for these molecules was in the micromolar ( $\mu\text{M}$ ) range. The attempt of characterizing the binding of triazoles inhibitors to hRAD51 through biophysical techniques failed because of the low solubility ( $2\ \mu\text{M}$  -  $10\ \mu\text{M}$ ) of these class of molecules in aqueous buffer. In cell-based experiments, the solubility issue was probably overcome by the presence of carrier proteins (i.e. albumin) or other factors present in cells and in cells growth medium. Further SAR studies carried on the triazole compounds showed that improving the potency of these molecules was quite challenging. The low solubility of this class of compounds together with the issue of low potency effectively moved the group to look into new classes of inhibitors. The medicinal chemistry unit of the group identified two new classes of compounds: dihydropyrazole and carboxyphenyl furane molecules, respectively. The leading hit molecules, belonging to these two new classes of compounds ARN22064 and ARN22142, were investigated by nuclear magnetic resonance spectroscopy (NMR) and microscale thermophoresis (MST) to assess their binding to hRAD51. The binding of ARN22064 to hRAD51 and of ARN22142 to hRAD51, were characterized by a binding affinity of  $K_d = 98.81 \pm 16.99\ \mu\text{M}$  and of  $K_d = 2.71 \pm 0.30\ \text{mM}$ , respectively. As expected, these interactions are weaker than the physiological one between hRAD51 and BRC4 that we estimated to be  $K_d = 63.30 \pm 17.98\ \text{nM}$ . Nevertheless, a displacement assay, performed by MST, showed that ARN22142 was able to interfere with the

hRAD51-BRC4 protein-protein interaction, weakening the latter interaction. These data fully support our primary goal that is preventing hRAD51-BRCA2 interaction rather than displacing BRCA2 after binding. Even though we are still working to improve the affinity of the selected compounds for hRAD51, we do not aim at completely inhibiting hRAD51-BRC4 interaction since it would result in toxicological effects targeting also the HR pathway in non-pathological cells. We aim at modulating the HR pathway through inhibition of the hRAD51-BRCA2 interaction relying on the fact that DSBs take place more often in cancer cells rather than in normal cells. The in-cell based experiments using pancreatic adenocarcinoma cells to verify the ability of the two compounds, ARN22064 and ARN22142, to impair the HR pathway are in process. These data, together with the obtained biophysical results, will highlight how these molecules specifically target the protein-protein hRAD51-BRCA2 interaction.

The search for a new class of hRAD51-BRCA2 inhibitors was also carried out in parallel through a NMR fragment based drug discovery (FBDD) approach. An in house library of fluorinated compounds (MW ~ 350 Da) was screened by NMR for the identification of molecules able to bind the full-length hRAD51 protein. The identified hits were then tested by NMR for their ability to be displaced by BRC4 in their binding to the hRAD51 protein. Out of 185 fragments initially screened, three fragments, E5, A6 and F11, were selected through NMR experiments as hRAD51 binders able to be displaced by BRC4. The binding affinity of these fragments for hRAD51 was determined by MST. The MST analysis yielded a  $K_d$  of  $1.23 \pm 0.13$  mM for the E5 fragment binding hRAD51. The E5 fragment weakly binds hRAD51. Nevertheless, E5 is able to affect the hRAD51-BRC4 interaction leading to a reduced BRC4 affinity for hRAD51. Titration of A6 and F11 did not induce any signal change in MST binding curves, suggesting that these two fragments do not bind hRAD51 at the tested concentration (4

mM). Taking into account that the scaffold of the identified E5 fragment is quite different from the classes of molecules proposed by the classical drug-discovery process, we plan to pursue the optimization of this fragment and the successive development of a new molecule through a fragment based approach supported by NMR spectroscopy.

As already mentioned, in the PDB the only crystal structure of the hRAD51-BRC4 complex available (1N0W), is missing the N terminal portion of hRAD51 and a flexible polypeptide chain links the C terminal of hRAD51 and BRC4. The structure of hRAD51 recombinase filaments, during catalysis of DNA-strand exchange, was recently solved by cryo-EM<sup>58</sup> and image processing with the support of the already published structures (1N0W, 1SZP and 1B22). Although these findings provided new structural insights on the full length hRAD51 protein, they did not clarify how hRAD51 and BRCA2 interact and in particular, how the N terminal region of hRAD51 can be involved in the interaction with the BRCA2 protein.

We did pursue the crystallization of the full-length hRAD51 as well as of the protein in complex with BRC4. The expression and purification protocols of the two complexes were optimized. hRAD51 was still partially aggregated in the co-expressed wild type hRAD51-BRC4 complex, while it was completely monomeric in the co-expressed hRAD51 F86E-BRC4 complex, thus suggesting that mutation F86E was essential to obtain a monomeric hRAD51 protein. Several crystallization screenings were performed both on the full-length hRAD51 and on the co-expressed hRAD51-BRC4 and hRAD51 F86E-BRC4 complexes. Large, single crystals have been obtained from three crystallization screening conditions for the hRAD51 F86E-BRC4 complex. Data will be collected at a synchrotron source in the next weeks.

BRC3, one of the BRCA2 domains that shows the highest affinity for hRAD51, interacts with the N terminal domain of hRAD51<sup>104</sup>. No structural details on this critical interaction are

yet available. We have recently cloned a BRC3-BRC4 construct that we will be co-expressed with the hRAD51 F86E protein and successively used in crystallization screenings.

Crystallization trials are also in progress for hRAD51 F86E in complex with the recently identified soluble hit inhibitors.

The present study has allowed the identification of hit molecules able to interfere with hRAD51-BRC4 starting from two different drug discovery strategies: a virtual screening and a fragment based approach. Biophysical analyses of hRAD51-BRC4 binding ( $K_d$  in the nanomolar range) shows how it is challenging to identify molecules able to prevent this interaction. Nevertheless, MST analyses of some of the tested molecules show their ability to interfere with hRAD51-BRC4 binding. In addition, structural information on the optimized full-length hRAD51 F86E-BRC4 complex and full-length hRAD51 F86E-inhibitors will clarify how these molecules are located with respect to the FXXA hRAD51 motif of protomer-protomer interaction. These data will be critical for the medicinal chemistry group to synthesize new lead molecules able to prevent the hRAD51-BRCA2 binding.

The structure of the full-length hRAD51 F86E in complex with BRC4 and BRC3-BRC4 will also clarify the role of the N terminal region of hRAD51 in the interaction with BRCA2. These findings will eventually allow the identification of new hotspots in the protein-protein interaction while providing new details about the HR pathway mechanism.

The goal of this multidisciplinary project is the identification of small molecules, inhibitors of the hRAD51-BRCA2 interaction, to be administrated in combination with PARPi in patients carrying nonmutated genes or when PARPi resistance is developed. We are confident that our results will contribute to extend the possibility of exploiting this approach to other gene pairs involved in diverse synthetic lethality pathways.

---

---

# Appendix

---

---

## **A1 Fragment based drug discovery (FBDD)**

The FBDD is based on the Fragment Based Approach (FBA) a screening methods developed in the 80's-90's to overcome some drawbacks of the High Throughput Screening. A potent inhibitor can be considered as the combination of fragments that bind different parts of the protein<sup>105</sup>; the affinity of the whole molecule can be considered as the sum of its single parts<sup>106</sup>. The FBA advantage relies on the easiness in investigating the chemical space of a binding site using small molecules rather than bigger chemical entities. Indeed, Hann and coworkers<sup>107</sup> have demonstrated that the probability of randomly finding lead compounds strongly decrease with increasing molecular complexity, because small molecules can better fit the different binding pockets. Covering a chemical space can be better achieved by screening small fragment libraries rather than larger ones such as the one constituted by complex molecules. The hit rate for screening fragments is however higher than for larger molecules.

Currently, there is not a unique definition for a fragment. Fragments are typically compounds that follow the rule of three (RO3)<sup>108</sup>:  $MW \leq 300$  Da; No. HBD (Hydrogen Bond Donor)  $\leq 3$ ; No. HBA (Hydrogen Bond Acceptor)  $\leq 3$ ; No. ClogP (calculated LogP)  $\leq 3$ . However, some fragment hits do not adhere to these restrictions indicating that this rule can be applied only as a guideline. The generation of a fragment library is based on the application of filters (i.e. RO3) to compounds from commercial or in house collections.

The low size of the fragments allows to check their properties in buffer solution before testing them. This can avoid issues related to false positive due to the aggregation of the compounds in solutions. On this purpose, Dalvit and coworkers<sup>100</sup> developed a NMR approach to study the solubility, purity and aggregation state of the small molecules (SPAM), which has also been applied to the research work reported in this thesis.

Due to their chemical simplicity and reduced functionality, active compounds (hits),

coming from FBA, usually show weak affinity for the target protein in the micro- milli-molar range. Fragments need therefore to be screened at high concentration; sensitive detection methods and structural information are also required to identify active fragments and improve their potency. The most used techniques are NMR, X-ray crystallography, SPR and fluorescence spectroscopy. Each approach presents both advantage and disadvantages.

In this research work the FBDD was investigated by NMR, therefore only NMR-based screenings will be described in more details.

## A2 NMR-based screenings

Diversity and robustness of NMR-based screening methods make these techniques highly interesting a tool for drug discovery. Only 5% of a fragment needs to interact with the partner to be revealed as a NMR hit, consequently NMR can detect hits with solubilities lower than their potencies<sup>109</sup>. In addition, NMR studies are characterized by high reproducibility and by the opportunity to check the quality of fragments (i.e. purity, solubility and aggregation state).

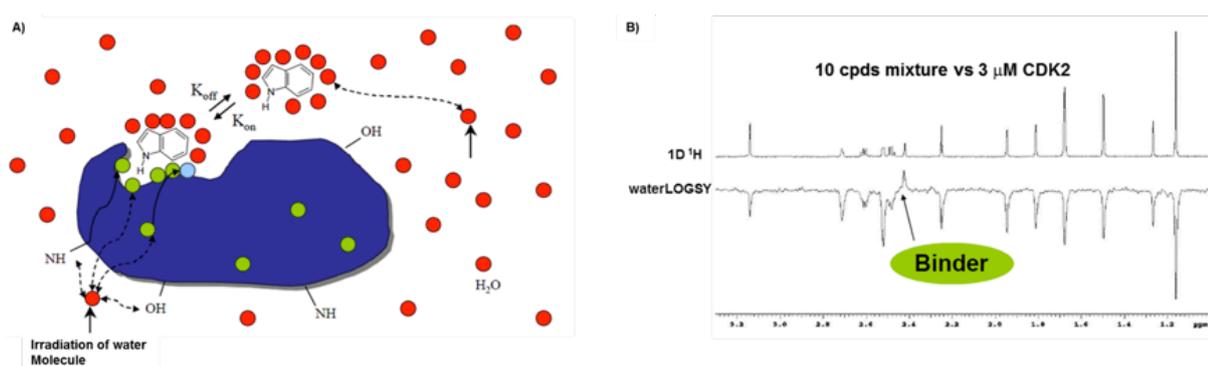
In NMR-based screenings two main approaches can be distinguished: “target-observed” and “ligand-observed” based techniques. The former provides structural information about the protein-ligand binding and allows the detection of the binding affinity, but it requires large amount of labeled proteins and it is limited by the size of the target proteins. The “ligand-observed” based technique can be applied to a broader range of targets and unlabeled proteins at low concentration are used. Ligand-observed methods show changes in the NMR spectrum of a ligand upon interaction with the target protein. The most used ligand-observed methods are: WaterLOGSY<sup>102</sup>, <sup>1</sup>H transverse relaxation time ( $T_{2p}$ )<sup>110</sup>, STD<sup>111-113</sup> and <sup>19</sup>F NMR-based experiments<sup>114</sup>.

In this thesis WaterLOGSY and <sup>19</sup>F NMR-based experiments were used in order to investigate the binding of fragments (from in house collection), alone or in mixtures to hRAD51.

### A2.1 WaterLOGSY

WaterLOGSY (Water-Ligand Observation with Gradient SpectroscopyY) is based on the transfer of magnetization from the protein to the small ligand (Figure A1) by using the bulk water magnetization. The water is selectively irradiated for a long period or selectively excited and followed by a long mixing time. During the irradiation time or the mixing time, the water magnetization is transferred to the protein via intermolecular NOE (Nuclear Overhauser Effect) and chemical exchange. The stored magnetization on the protein is transferred via intermolecular

NOE to the bound fragment. This NOE is negative (signals are conventionally displayed as positive in the WaterLOGSY spectra) due to the long tumbling correlation time of the protein. In addition to this mechanism one has to consider the transfer of bulk water magnetization to the fragment free in solution. This intermolecular NOE is positive (signals are conventionally displayed as negative in the WaterLOGSY spectra) due to the short tumbling correlation time of the fragment not complexed with the receptor. Therefore, the WaterLOGSY effect on a ligand is the weighted average of these two NOE effects. Owing to the fact that the ligand is not in high excess with respect to the protein and that the negative NOE is significantly larger in absolute value with respect to the positive NOE, the signals in the WaterLOGSY spectrum of a ligand in the presence of the protein are positive (Negative NOE). Thus, it is easily possible to discriminate the binders from non-binders. This method can also be applied to evaluate the possible aggregation of molecules.



**Figure A1:** Image adapted from Dalvit et al.<sup>102</sup>

**A)** The protein is shown with the buried cavities and the active binding site. Ligand is shown in the bound and free states. Excitation of bulk water (circles) is shown with a solid arrow and some of different magnetization transfer pathways are shown with dashed lines.

**B)** One-dimensional reference (upper) and WaterLOGSY (lower) spectra recorded for a 10-compound chemical mixture in the presence of  $10 \mu M$  cdk2: positive and negative signals in the lower spectrum identify cdk2 interacting (binder) and not interacting molecules, respectively.

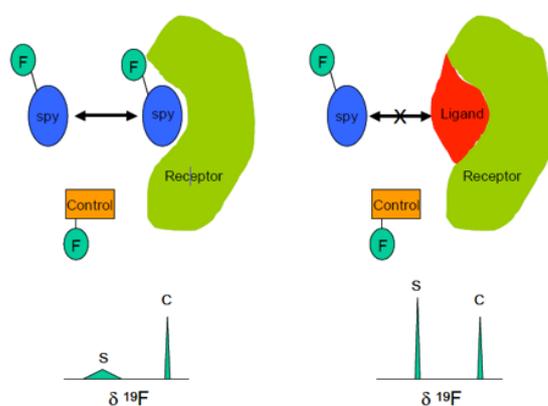
## A2.2 <sup>19</sup>F NMR-based experiments

In a recent study Dalvit<sup>115</sup> demonstrated as <sup>19</sup>F NMR, due to its very large dynamic range, is the most sensitive technique for detecting weak binders. <sup>19</sup>F NMR, therefore, allows to test the compounds at concentration orders of magnitude lower than their affinities and to identify very weak inhibitors (mM) even if their solubility is low. <sup>19</sup>F nucleus presents several advantages: 100% natural abundance, high sensitivity, favorable transverse relaxation and large dispersion of chemical shifts that allows the screening of large chemical mixtures without issues of signals overlapping. By <sup>19</sup>F NMR spectroscopy it is possible to perform direct and competition (FAXS)<sup>116</sup> binding screenings and functional biochemical screenings (FABS)<sup>117</sup>. Both binding experiments are based on R<sub>2</sub> filter fluorine experiments. The transverse relaxation rate R<sub>2</sub> is a very sensible parameter for these studies, due to the large Chemical Shift Anisotropy (CSA) of <sup>19</sup>F nucleus and to the large exchange contribution. For a fragment weakly interacting with the receptor, the observed response (R<sub>obs</sub>), in the R<sub>2</sub> filter experiments is given by the equation:

$$R_{2,obs} = P_b R_{2,b} + P_f R_{2,f} + P_b (P_f)^2 \frac{4\pi(\delta_{free} - \delta_{bound})^2}{K_{-1}}$$

where the R<sub>2,obs</sub> and R<sub>b</sub> are transverse relaxation rates in the free and bound states, respectively. P<sub>b</sub> and P<sub>f</sub> are the fraction of ligand bound and free, respectively. The last term  $(\delta_{free} - \delta_{bound})^2$  is the exchange term where  $\delta_{bound}$  and  $\delta_{free}$  are the isotropic chemical shifts of the fluorine resonance of molecules in the bound and free states, respectively, and 1/K<sub>-1</sub> is the residence time  $\tau_{res}$  of the ligand bound to the protein. The latter is responsible for the line-broadening and the consequent intensity decrease of the NMR signal observed when a compound binds to the target protein. The compounds can be screened in large mixtures or as single compounds. For each sample, two <sup>19</sup>F R<sub>2</sub> filter experiments are recorded; one in the absence and one in presence of

the target protein. A scheme of the FAXS experiment is shown in Figure A2. In the presence of a competitor, the line broadening of the  $^{19}\text{F}$  NMR signal of the binder becomes sharp again. FAXS experiments results very useful not only to screen non fluorinated compounds, but also to confirm the binding pocket of the identified hits: the displacement of the hits by the natural substrate and/or a known inhibitor suggests a competition event for the same binding site.

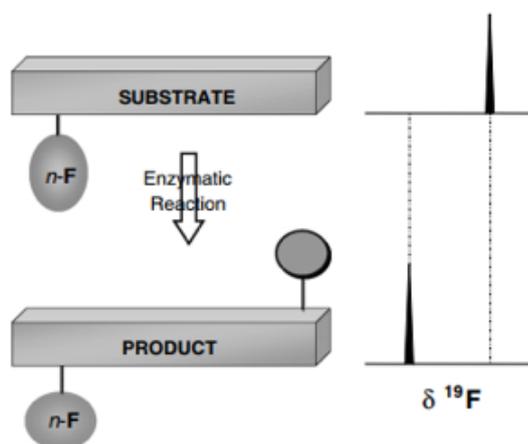


**Figure A2:** Schematic diagram of the FAXS experiment. The broad signal of the spy molecule becomes sharp in the presence of a competitive ligand due to the displacement of the spy molecule from the receptor. (Taken from Dalvit et al.<sup>114</sup>).

### A2.3 n-Fluorine Atoms for Biochemical Screenings (n-FABS)

n-FABS (n-Fluorine Atoms for Biochemical Screening) is a substrate-based screening assay in which the substrate or the cofactor of an enzyme is labelled with a fluorine moiety (Figure A3).  $^{19}\text{F}$ -NMR spectroscopy is then used to detect and quantify the signal intensities of the substrate and of the enzymatic reaction products. The enzymatic modification of the substrate changes the electronic cloud of the F moiety resulting in different isotropic chemical shifts for the product and substrate fluorine signals. The high sensitivity of fluorine to the chemical environment allows the insertion of the F moiety even far from the reaction site. It is very important to notice that when using  $^{19}\text{F}$ -NMR, only the signals of the fluorinated molecules appear in the spectrum

and there is no interference between the enzymatic reaction and the detection system.



**Figure A3:** Schematic representation of FABS experiment. (Taken from Dalvit et al.<sup>114</sup>).

In the present study, WaterLOGSY and both direct-binding and FAXS experiments were performed on an in house fluorinated fragment library and on compounds identified by virtual screening campaigns on the X-ray structure of the hRAD51-BRC4 complex (PDB: 1N0W). For triazole compounds it was not possible to obtain any NMR spectrum, because the solubility of the small molecules was too low (<10  $\mu\text{M}$ ) to detect any signal.

---

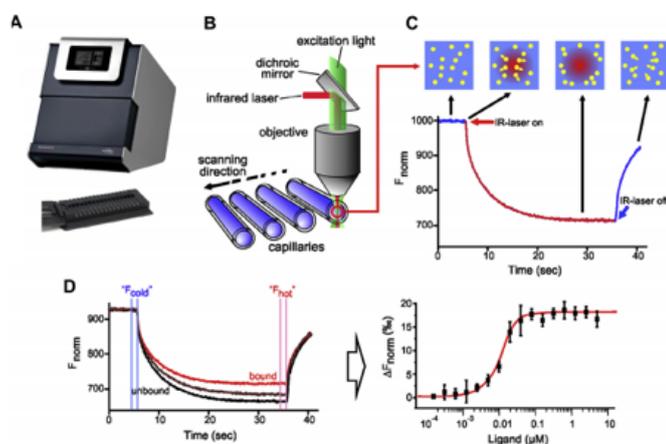
### A3 Microscale thermophoresis (MST)

Microscale thermophoresis is a powerful tool to quantify biomolecular interactions. The measurement method is based on the directed movement of fluorescent molecules along a temperature gradient, an effect called “thermophoresis”. A local temperature difference  $\Delta T$ , induced by an infrared (IR) laser localized on glass capillaries containing the fluorescent molecules, leads to a local change in molecule concentration ( $c_{hot}/c_{cold}$ ; depletion or enrichment), quantified by the Soret coefficient  $ST$ :

$$\frac{c_{hot}}{c_{cold}} = e^{-ST\Delta T}$$

Thermophoresis depends on changes in size or on charge and solvation of shell molecules. As the buffer is kept constant, alterations in the thermophoretic depletion or enrichment can only arise from changes in the size, charge or solvation entropy of the fluorescent molecules. As the binding of a non-fluorescent ligand to a labeled molecule changes at least one of these properties, the binding can be quantified by measuring the change in thermophoresis at different ligand concentrations.

The instrument records the thermophoretic traces arising from the binding of a potential interactor mixed with the labelled molecule. Sixteen capillaries are simultaneously analyzed, each one containing the labelled molecule at constant concentration and the potential interactor at increasing concentrations. For analysis, the change in thermophoresis is expressed as a normalized fluorescence ( $\Delta F_{norm}$ ), which is plotted against molecular concentration to obtain a binding curve and consequently the  $K_d$  value of the interaction.

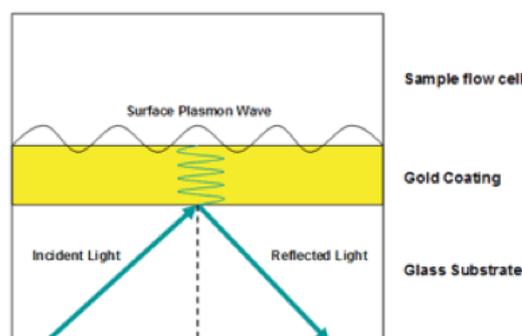


**Figure A4:** Schematic representation of MST setup and experiments. **A** Monolith NT.115 instrument (NanoTemper Technologies GmbH); **B** Representation of MST optics. IR laser scans up to 16 capillaries with a total volume of 4  $\mu\text{L}$ ; **C** Signal from a MST experiment: a constant initial fluorescence from samples is detected. When laser is on, the T-jump is observed, due to a rapid change of in fluorophore properties in different temperature conditions. Subsequently, the thermophoretic of the molecules is detected in the temperature gradient. When laser is off, an inverse T-jump effect is detected; **D** Typical binding experiment in which it is possible to recognize the thermophoretic traces of molecules in bound and unbound state. Subsequently, the fluorescence traces are normalized and plotted vs ligand concentration to derive the  $K_d$  value; (Taken from Jerabek-Willemsen et al.<sup>118</sup>).

In the present study the protein His-hRAD51 was labelled with two different dyes: the blue dye NT-495 that directly binds the amine group of lysine residues and the red dye NT-647 that labels biomolecules on hexahistidine tails. The choice of two different dyes arises from the intrinsic fluorescence of some compounds in blue or red light that interfere with MST analysis. The molecules arising from virtual screening analysis (triazoles, ARN22142 and ARN22064) and from NMR-based screenings (E5, A6 and F11) were analyzed through MST in order to investigate their ability to bind hRAD51.

#### A4 Surface Plasmon Resonance (SPR)

SPR is a physical phenomenon due to the interaction between light and matter. When a plane polarized light beam hits the interface metal-dielectric with an angle equal to an incident angle, called SPR angle, in condition of total internal reflection (TIR) it is partially adsorbed, generating a surface plasmon wave (Figure A5). The SPR angle depends on different factors, such as the thickness of the metal film or change in the refractive index due to the mass increase on the surface of the metal layer. In SPR instruments, the sensor consists of a glass substrate and a thin gold coating. The high sensitivity of modern SPR sensors allows to detect the binding of small molecules with protein targets. The former are dissolved in the buffer, the latter are usually immobilized through different chemical linkages to the surface of the biosensor.



**Figure A5:** Schematic representation of a sensor in a SPR instrument.  
(Taken from <https://www.fortebio.com/pioneer-spr-principles.html>).

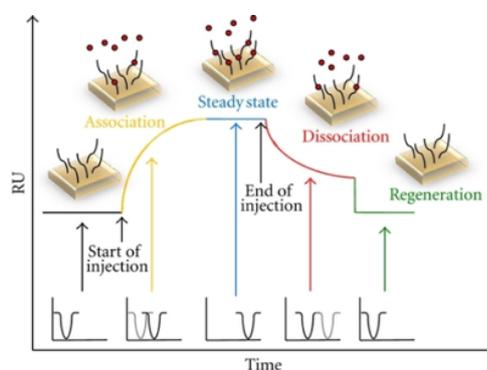
Injecting analytes over the chip, they can bind the protein target, causing a change in the SPR angle, proportional to the mass of bound ligand. Once the macromolecule is immobilized onto the chip ( $R_{\text{ligand}}$ ), the maximum response ( $R_{\text{max}}$ ) achievable during the binding of the ligand can be calculated according to the following equation:

$$R_{\text{max}} = \frac{MW_{\text{analyte}} * R_{\text{ligand}} * n}{MW_{\text{ligand}}}$$

where  $MW_{\text{analyte}}$  and  $MW_{\text{ligand}}$  correspond to the molecular weight of analyte and ligand, respectively.  $n$  represents the stoichiometry of the reaction.

$R_{\max}$  values higher than theoretic  $R_{\max}$  are due to non-specific binding or to a super-stoichiometric interaction.

The result from a SPR experiment is a sensogram, a plot of the change in SPR angle (Response) against the time that contains kinetic information of the binding detected (Figure A6). In a sensogram, the association phase in which the analyte binds to the macromolecule can be distinguished from the plateau, when the equilibrium is achieved (steady state). Finally, after the injection is stopped, the molecule starts to dissociate (dissociation phase). By fitting the curves corresponding to the association and dissociation phases, it is possible to calculate the kinetic constants  $k_{\text{on}}$  and  $k_{\text{off}}$  of the biomolecular interaction. The ratio of these two values provides the equilibrium constant  $K_d$ .



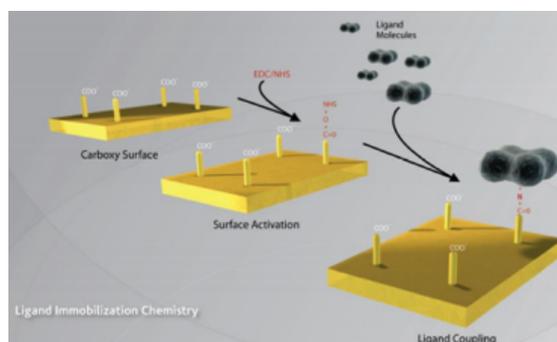
**Figure A6:** Representation of a sensogram in which it is possible discriminate the different phases during a SPR experiment. (Taken from Ritzefeld et al.<sup>119</sup>).

This technique allows to obtain measurements in real time and kinetics. In addition, the labelling of the molecules is not required. However, the immobilization of the proteins on the sensor surface can affect, in specific cases, the measured binding constants, since the immobilization may hinder peculiar binding pockets of the molecules on the protein.

Different sensor chips are available for the immobilization of the biomolecule of interest. Biosensors are low, medium or high capacity two or three dimensional surface chemistry that

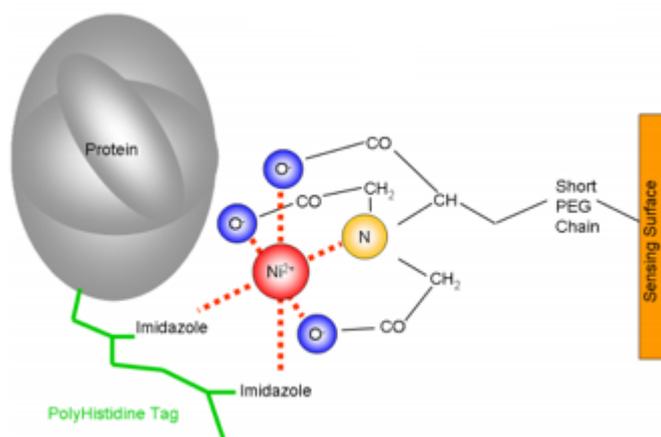
can comprise different spacer (i.e. polyethylene glycol, dextran, carboxylated polysaccharide hydrogel) provided with different functional groups (i.e. carboxylic acids, nitrilotriacetic acid (NTA) functional groups, biotin functional groups) that allow to immobilize biomolecules exploiting different chemistries.

In the present study two different immobilization strategies were performed. In one case His-hRAD51 was immobilized via amine coupling, exploiting the presence of lysine residues on the protein surface available to form amine bonds (Figure A7). The protein immobilization step exploits electrostatics interactions between a positively charged His-hRAD51 onto the electronegative surface. However, since this step includes dilution of the protein in a solution with a pH below its isoelectric point, His-hRAD51 precipitated.



**Figure A7:** Immobilization of a biomolecule via amine coupling onto a carboxyl surface chemically activated. (Taken from <http://www.alfatest.it/keyportal/uploads/coo-chips-20130924.pdf>).

A second immobilization trial of His-hRAD51 was performed using a His-Cap biosensor that employs the nitriloacetic acid (NTA)-nickel technique for protein attachment exploiting the His-tag binding (Figure A8). This method has the advantage of working in the optimal protein buffer condition and of providing a directed immobilization of His-tagged proteins. The protein was successfully immobilized, as shown in the results and discussion section.



**Figure A8:** Schematic representation of a His-tag biomolecule on a sensing surface exploiting the NTA-nickel technique; (Taken from <http://www.alfatest.it/keyportal/uploads/hiscap-chips.pdf>).

- 5' dRP:** 5' deoxyribose phosphate moiety
- ADP:** adenosine diphosphate
- A-NHEJ:** Alternative-Non Homologous End Joining pathway
- AP:** APyrimidinic or APurinic bases
- ATM:** Ataxia-Telangiectasia Mutated kinase
- ATP:** Adenosine TriPhosphate
- ATR:** Ataxia Telangiectasia and Rad3-related protein
- BER:** Base Excision Repair
- BLM/DNA2:** Bloom Syndrome RecQ Like Helicase/DNA replication helicase nuclease 2
- bp:** Base pairs
- BRCA1:** BReast Cancer-Associated 1
- BRCA2:** BReast Cancer-Associated 2
- CDK:** Cyclin Dependent Kinase
- CHK1:** CHeckpoint Kinase 1
- CHK2:** CHeckpoint Kinase 2
- ClogP:** Calculated LogP
- C-NHEJ:** Classical Non-Homologous End-Joining
- CSA:** Chemical Shift Analysis
- CSB:** Chicago Sky Blue
- CtBP:** C-terminal Binding Protein
- CtIP:** C-terminal binding protein Interacting Protein
- DBD:** DNA Binding Domain
- DDR:** DNA Damage Response
- dHJ:** Double Holliday Junction
- DIDS:** 4,4'-Diisothiocyanostilbene-2,2'-DiSulfonic acid
- DLS:** Dynamic Light Scattering
- DMSO:** Dimethyl Sulfoxide
- DNA-PKc:** DNA Protein Kinases
- DSB:** Double Strand Breaks
- DSBR:** Double Strand Break Repair
- DTT:** DiThioThreitol

**EDTA:** EthyleneDiamineTetraacetic Acid  
**EM:** Electron Microscopy  
**ERCC1:** Excision Repair Complementing 1  
**Exo1:** Exonuclease 1  
**FABS:** Fluorine Atoms for Biochemical Screening  
**FANCF:** Fanconi Anemia group F protein  
**FAXS:** Fluorine chemical shift Anisotropy and eXchange for Screening  
**FBA:** Fragment Based Approach  
**FBDD:** Fragment Based Drug Discovery  
**FDA:** Food and Drug Administration  
**G1:** Gap 1 phase  
**G2:** Gap 2 phase  
**GST:** Glutathione S-Transferase  
**HBA:** Hydrogen Bond Acceptor  
**HPLC-MS:** High Performance Liquid Chromatography –Mass Spectrometry  
**HR:** Homologous Recombination  
**ICL:** Interstrand CrossLink  
**IPTG:** IsoPropyl  $\beta$ -D-1-ThioGalactopyranoside  
**IR:** Ionizing Radiation  
**ITC:** Isothermal Calorimetry  
**LB:** Luria Bertani  
**LEF:** Local Environment Fluorine  
**M:** Mitosis  
**MDC1:** Mediator of DNA damage Checkpoint protein 1  
**MES:** 2-(N-Morpholino)EthaneSulfonic acid  
**MMEJ:** Microhomology-Mediated End-Joining  
**MMR:** MisMatch Repair  
**MRN complex:** Mre11-Rad50-Nbs1 complex  
**MST:** MicroScale Thermophoresis  
**MYC:** MYC proto-oncogene protein  
**NER:** Nucleotide Excision Repair

**NHEJ:** Non-Homologous End Joining  
**NLS:** Nuclear Localization Signal  
**NMR:** Nuclear Magnetic Resonance  
**NOE:** Nuclear Overhauser Effect  
**NTA:** NitriloTriacetic Acid  
**OD:** Optical Density  
**PARP:** Poly(ADP)Ribose Polymerase  
**PARPi:** PARP inhibitors  
**PCNA:** Proliferating Cell Nuclear Antigen  
**PCR:** Polymerase Chain Reaction  
**PDB:** Protein Data Bank  
**PEG:** PolyEthylene Glycol  
**PFG:** Pulsed Field Gradient  
**PIKKs:** PhosphatidyInositol 3-Kinase-like protein Kinase family  
**PNKP:** PolyNucleotide Kinase 3' Phosphatase  
**pRB:** RetinoBlastoma protein  
**RFC:** Replication Factor C  
**RO3:** Rule Of three  
**ROS:** Reactive Oxygen Species  
**RPA:** Replication Protein A  
**RU:** Response Unit  
**S:** Synthesis phase  
**SAR:** Structure Activity Relationship  
**SCID:** Severe Combined Immunodeficiency  
**SDS-PAGE:** Sodium Dodecyl Sulphate - PolyacrylAmide Gel Electrophoresis  
**SDSA:** Synthesis-Dependent Strand Annealing  
**SOC:** Super Optimal Catabolite repression  
**SPAM:** Solubility, Purity and Aggregation of the Molecule  
**SPR:** Surface Plasmon Resonance  
**SSA:** Single Strand Annealing  
**SSB:** Single Strand Breaks

**SSBR:** Single Strand Break Repair

**STD:** Saturation Transfer Difference

**TB:** Terrific Broth

**TDP2:** Tyrosyl-DNA Phosphodiesterase 2

**TopBP1:** Topoisomerase-Binding Protein1

**UV light:** Ultraviolet light

**WaterLOGSY:** Water-Ligand Observation with Gradient Spectroscopy

**WEE1:** WEE1-like protein kinase

**XLf:** XRCC4 Like Factor

1. Hoeijmakers, J. H. J. DNA Damage, Aging, and Cancer. *N. Engl. J. Med.* **361**, 1475–1485 (2009).
2. Harper, J. W. & Elledge, S. J. The DNA damage response: ten years after. *Mol. Cell* **28**, 739–745 (2007).
3. Jackson, S. P. & Bartek, J. The DNA-damage response in human biology and disease. *Nature* **461**, 1071–1078 (2009).
4. Jiricny, J. The multifaceted mismatch-repair system. *Nat. Rev. Mol. Cell Biol.* **7**, 335–346 (2006).
5. Lindahl, T. & Barnes, D. E. Repair of endogenous DNA damage. *Cold Spring Harb. Symp. Quant. Biol.* **65**, 127–133 (2000).
6. Moldovan, G.-L. & D’Andrea, A. D. How the fanconi anemia pathway guards the genome. *Annu. Rev. Genet.* **43**, 223–249 (2009).
7. Caldecott, K. W. Single-strand break repair and genetic disease. *Nat. Rev. Genet.* **9**, 619–631 (2008).
8. West, S. C. Molecular views of recombination proteins and their control. *Nat. Rev. Mol. Cell Biol.* **4**, 435–445 (2003).
9. O’Connor, M. J. Targeting the DNA Damage Response in Cancer. *Mol. Cell* **60**, 547–560 (2015).
10. Meek, K., Dang, V. & Lees-Miller, S. P. DNA-PK: the means to justify the ends? *Adv. Immunol.* **99**, 33–58 (2008).
11. Zou, L. & Elledge, S. J. Sensing DNA damage through ATRIP recognition of RPA-ssDNA complexes. *Science* **300**, 1542–1548 (2003).
12. Schreiber, V., Dantzer, F., Ame, J.-C. & de Murcia, G. Poly(ADP-ribose): novel functions for an old molecule. *Nat. Rev. Mol. Cell Biol.* **7**, 517–528 (2006).
13. Rogakou, E. P., Pilch, D. R., Orr, A. H., Ivanova, V. S. & Bonner, W. M. DNA double-stranded breaks induce histone H2AX phosphorylation on serine 139. *J. Biol. Chem.* **273**, 5858–5868 (1998).
14. Podhorecka, M., Skladanowski, A. & Bozko, P. H2AX Phosphorylation: Its Role in DNA Damage Response and Cancer Therapy. *Journal of Nucleic Acids* (2010). doi:10.4061/2010/920161
15. Castedo, M. *et al.* Mitotic catastrophe constitutes a special case of apoptosis whose suppression entails aneuploidy. *Oncogene* **23**, 4362–4370 (2004).

- 
16. Ashworth, A. A synthetic lethal therapeutic approach: poly(ADP) ribose polymerase inhibitors for the treatment of cancers deficient in DNA double-strand break repair. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* **26**, 3785–3790 (2008).
  17. Curtin, N. J. DNA repair dysregulation from cancer driver to therapeutic target. *Nat. Rev. Cancer* **12**, 801–817 (2012).
  18. European Medicines Agency - Find medicine - Lynparza. Available at: [http://www.ema.europa.eu/ema/index.jsp?curl=pages/medicines/human/medicines/003726/human\\_med\\_001831.jsp&mid=WC0b01ac058001d124](http://www.ema.europa.eu/ema/index.jsp?curl=pages/medicines/human/medicines/003726/human_med_001831.jsp&mid=WC0b01ac058001d124). (Accessed: 2nd October 2017)
  19. Byun, T. S., Pacek, M., Yee, M., Walter, J. C. & Cimprich, K. A. Functional uncoupling of MCM helicase and DNA polymerase activities activates the ATR-dependent checkpoint. *Genes Dev.* **19**, 1040–1052 (2005).
  20. Cimprich, K. A. & Cortez, D. ATR: an essential regulator of genome integrity. *Nat. Rev. Mol. Cell Biol.* **9**, 616–627 (2008).
  21. Macheret, M. & Halazonetis, T. D. DNA replication stress as a hallmark of cancer. *Annu. Rev. Pathol.* **10**, 425–448 (2015).
  22. Bester, A. C. *et al.* Nucleotide deficiency promotes genomic instability in early stages of cancer development. *Cell* **145**, 435–446 (2011).
  23. Buisson, R., Boisvert, J. L., Benes, C. H. & Zou, L. Distinct but Concerted Roles of ATR, DNA-PK, and Chk1 in Countering Replication Stress during S Phase. *Mol. Cell* **59**, 1011–1024 (2015).
  24. Jones, R. M. *et al.* Increased replication initiation and conflicts with transcription underlie Cyclin E-induced replication stress. *Oncogene* **32**, 3744–3753 (2013).
  25. Rohban, S. & Campaner, S. Myc induced replicative stress response: How to cope with it and exploit it. *Biochim. Biophys. Acta* **1849**, 517–524 (2015).
  26. Vafa, O. *et al.* c-Myc can induce DNA damage, increase reactive oxygen species, and mitigate p53 function: a mechanism for oncogene-induced genetic instability. *Mol. Cell* **9**, 1031–1044 (2002).
  27. Sabharwal, S. S. & Schumacker, P. T. Mitochondrial ROS in cancer: initiators, amplifiers or an Achilles' heel? *Nat. Rev. Cancer* **14**, 709–721 (2014).
  28. Barazzuol, L. *et al.* Evaluation of poly (ADP-ribose) polymerase inhibitor ABT-888 combined with radiotherapy and temozolomide in glioblastoma. *Radiat. Oncol. Lond. Engl.* **8**, 65 (2013).

- 
29. Reaper, P. M. *et al.* Selective killing of ATM- or p53-deficient cancer cells through inhibition of ATR. *Nat. Chem. Biol.* **7**, 428–430 (2011).
30. Erice, O. *et al.* MGMT Expression Predicts PARP-Mediated Resistance to Temozolomide. *Mol. Cancer Ther.* **14**, 1236–1246 (2015).
31. Ceccaldi, R., Rondinelli, B. & D’Andrea, A. D. Repair Pathway Choices and Consequences at the Double-Strand Break. *Trends Cell Biol.* **26**, 52–64 (2016).
32. Srivastava, M. & Raghavan, S. C. DNA Double-Strand Break Repair Inhibitors as Cancer Therapeutics. *Chem. Biol.* **22**, 17–29 (2015).
33. Karanam, K., Kafri, R., Loewer, A. & Lahav, G. Quantitative live cell imaging reveals a gradual shift between DNA repair mechanisms and a maximal use of HR in mid S phase. *Mol. Cell* **47**, 320–329 (2012).
34. Walker, J. R., Corpina, R. A. & Goldberg, J. Structure of the Ku heterodimer bound to DNA and its implications for double-strand break repair. *Nature* **412**, 607–614 (2001).
35. Dobbs, T. A., Tainer, J. A. & Lees-Miller, S. P. A structural model for regulation of NHEJ by DNA-PKcs autophosphorylation. *DNA Repair* **9**, 1307–1314 (2010).
36. Fell, V. L. & Schild-Poulter, C. The Ku heterodimer: function in DNA repair and beyond. *Mutat. Res. Rev. Mutat. Res.* **763**, 15–29 (2015).
37. Truong, L. N. *et al.* Microhomology-mediated End Joining and Homologous Recombination share the initial end resection step to repair DNA double-strand breaks in mammalian cells. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 7720–7725 (2013).
38. Cheng, Q. *et al.* Ku counteracts mobilization of PARP1 and MRN in chromatin damaged with DNA double-strand breaks. *Nucleic Acids Res.* **39**, 9605–9619 (2011).
39. Mansour, W. Y., Rhein, T. & Dahm-Daphi, J. The alternative end-joining pathway for repair of DNA double-strand breaks requires PARP1 but is not dependent upon microhomologies. *Nucleic Acids Res.* **38**, 6065–6077 (2010).
40. Escribano-Díaz, C. *et al.* A cell cycle-dependent regulatory circuit composed of 53BP1-RIF1 and BRCA1-CtIP controls DNA repair pathway choice. *Mol. Cell* **49**, 872–883 (2013).
41. Sneeden, J. L., Grossi, S. M., Tappin, I., Hurwitz, J. & Heyer, W.-D. Reconstitution of recombination-associated DNA synthesis with human proteins. *Nucleic Acids Res.* **41**, 4913–4925 (2013).

- 
42. Lin, F. L., Sperle, K. & Sternberg, N. Model for homologous recombination during transfer of DNA into mouse L cells: role for DNA ends in the recombination process. *Mol. Cell. Biol.* **4**, 1020–1034 (1984).
43. Richardson, C., Stark, J. M., Ommundsen, M. & Jasin, M. Rad51 overexpression promotes alternative double-strand break repair pathways and genome instability. *Oncogene* **23**, 546–553 (2004).
44. Iwanaga, R., Komori, H. & Ohtani, K. Differential regulation of expression of the mammalian DNA repair genes by growth stimulation. *Oncogene* **23**, 8581–8590 (2004).
45. Chen, F. *et al.* Cell cycle-dependent protein expression of mammalian homologs of yeast DNA double-strand break repair genes Rad51 and Rad52. *Mutat. Res.* **384**, 205–211 (1997).
46. Henson, S. E. *et al.* Pir51, a Rad51-interacting protein with high expression in aggressive lymphoma, controls mitomycin C sensitivity and prevents chromosomal breaks. *Mutat. Res.* **601**, 113–124 (2006).
47. Davies, A. A. *et al.* Role of BRCA2 in control of the RAD51 recombination and DNA repair protein. *Mol. Cell* **7**, 273–282 (2001).
48. Gildemeister, O. S., Sage, J. M. & Knight, K. L. Cellular redistribution of Rad51 in response to DNA damage: novel role for Rad51C. *J. Biol. Chem.* **284**, 31945–31952 (2009).
49. Wong, A. K., Pero, R., Ormonde, P. A., Tavtigian, S. V. & Bartel, P. L. RAD51 interacts with the evolutionarily conserved BRC motifs in the human breast cancer susceptibility gene *brca2*. *J. Biol. Chem.* **272**, 31941–31944 (1997).
50. Buis, N. *et al.* Classical molecular dynamics simulations of the complex between the RAD51 protein and the BRC hairpin loops of the BRCA2 protein. *Mol. Simul.* **34**, 749–759 (2008).
51. Pellegrini, L. *et al.* Insights into DNA recombination from the structure of a RAD51-BRCA2 complex. *Nature* **420**, 287–293 (2002).
52. Venkitaraman, A. R. Cancer susceptibility and the functions of BRCA1 and BRCA2. *Cell* **108**, 171–182 (2002).
53. Carreira, A. *et al.* The BRC repeats of BRCA2 modulate the DNA-binding selectivity of RAD51. *Cell* **136**, 1032–1043 (2009).
54. Shahid, T. *et al.* Structure and mechanism of action of the BRCA2 breast cancer tumor suppressor. *Nat. Struct. Mol. Biol.* **21**, nsmb.2899 (2014).

- 
- 
55. San Filippo, J., Sung, P. & Klein, H. Mechanism of eukaryotic homologous recombination. *Annu. Rev. Biochem.* **77**, 229–257 (2008).
56. Bochkarev, A. & Bochkareva, E. From RPA to BRCA2: lessons from single-stranded DNA binding by the OB-fold. *Curr. Opin. Struct. Biol.* **14**, 36–42 (2004).
57. Chen, C. F., Chen, P. L., Zhong, Q., Sharp, Z. D. & Lee, W. H. Expression of BRC repeats in breast cancer cells disrupts the BRCA2-Rad51 complex and leads to radiation hypersensitivity and loss of G(2)/M checkpoint control. *J. Biol. Chem.* **274**, 32931–32935 (1999).
58. Xu, J. *et al.* Cryo-EM structures of human recombinase RAD51 filaments in the catalysis of DNA strand exchange. *Nat. Struct. Mol. Biol.* **24**, 40–46 (2017).
59. Luo, G. *et al.* Disruption of mRad50 causes embryonic stem cell lethality, abnormal embryonic development, and sensitivity to ionizing radiation. *Proc. Natl. Acad. Sci. U. S. A.* **96**, 7376–7381 (1999).
60. Xiao, Y. & Weaver, D. T. Conditional gene targeted deletion by Cre recombinase demonstrates the requirement for the double-strand break repair Mre11 protein in murine embryonic stem cells. *Nucleic Acids Res.* **25**, 2985–2991 (1997).
61. Barlow, C. *et al.* Atm-deficient mice: a paradigm of ataxia telangiectasia. *Cell* **86**, 159–171 (1996).
62. Brown, E. J. & Baltimore, D. ATR disruption leads to chromosomal fragmentation and early embryonic lethality. *Genes Dev.* **14**, 397–402 (2000).
63. Rijkers, T. *et al.* Targeted inactivation of mouse RAD52 reduces homologous recombination but not resistance to ionizing radiation. *Mol. Cell. Biol.* **18**, 6423–6429 (1998).
64. Rooney, S. *et al.* Leaky Scid phenotype associated with defective V(D)J coding end processing in Artemis-deficient mice. *Mol. Cell* **10**, 1379–1390 (2002).
65. Dupré, A. *et al.* A forward chemical genetic screen reveals an inhibitor of the Mre11-Rad50-Nbs1 complex. *Nat. Chem. Biol.* **4**, 119–125 (2008).
66. Knight, Z. A. Small molecule inhibitors of the PI3-kinase family. *Curr. Top. Microbiol. Immunol.* **347**, 263–278 (2010).
67. Aggarwal, M., Sommers, J. A., Shoemaker, R. H. & Brosh, R. M. Inhibition of helicase activity by a small molecule impairs Werner syndrome helicase (WRN) function in the cellular response to DNA damage or replication stress. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 1525–1530 (2011).

- 
68. Srivastava, M. *et al.* An inhibitor of nonhomologous end-joining abrogates double-strand break repair and impedes cancer progression. *Cell* **151**, 1474–1487 (2012).
69. Freschauf, G. K. *et al.* Identification of a small molecule inhibitor of the human DNA repair enzyme polynucleotide kinase/phosphatase. *Cancer Res.* **69**, 7739–7746 (2009).
70. Schlacher, K., Wu, H. & Jasin, M. A distinct replication fork protection pathway connects Fanconi anemia tumor suppressors to RAD51-BRCA1/2. *Cancer Cell* **22**, 106–116 (2012).
71. Ray Chaudhuri, A. *et al.* Replication fork stability confers chemoresistance in BRCA-deficient cells. *Nature* **535**, 382–387 (2016).
72. Wahba, L., Gore, S. K. & Koshland, D. The homologous recombination machinery modulates the formation of RNA-DNA hybrids and associated chromosome instability. *eLife* **2**, e00505 (2013).
73. Wiegmans, A. P. *et al.* Rad51 supports triple negative breast cancer metastasis. *Oncotarget* **5**, 3261–3272 (2014).
74. Ito, M. *et al.* Rad51 siRNA delivered by HVJ envelope vector enhances the anti-cancer effect of cisplatin. *J. Gene Med.* **7**, 1044–1052 (2005).
75. Budke, B., Lv, W., Kozikowski, A. P. & Connell, P. P. Recent Developments Using Small Molecules to Target RAD51: How to Best Modulate RAD51 for Anticancer Therapy? *ChemMedChem* **11**, 2468–2473 (2016).
76. Alagpulinsa, D. A., Ayyadevara, S. & Shmookler Reis, R. J. A Small-Molecule Inhibitor of RAD51 Reduces Homologous Recombination and Sensitizes Multiple Myeloma Cells to Doxorubicin. *Front. Oncol.* **4**, 289 (2014).
77. Mason, J. M. *et al.* RAD54 family translocases counter genotoxic effects of RAD51 in human tumor cells. *Nucleic Acids Res.* **43**, 3180–3196 (2015).
78. Takaku, M. *et al.* Halenaquinone, a chemical compound that specifically inhibits the secondary DNA binding of RAD51. *Genes Cells Devoted Mol. Cell. Mech.* **16**, 427–436 (2011).
79. Hühn, D., Bolck, H. A. & Sartori, A. A. Targeting DNA double-strand break signalling and repair: recent advances in cancer therapy. *Swiss Med. Wkly.* **143**, w13837 (2013).
80. Zhu, J. *et al.* A novel small molecule RAD51 inactivator overcomes imatinib-resistance in chronic myeloid leukaemia. *EMBO Mol. Med.* **5**, 353–365 (2013).

- 
81. Raderschall, E. *et al.* Elevated levels of Rad51 recombination protein in tumor cells. *Cancer Res.* **62**, 219–225 (2002).
82. Maacke, H. *et al.* DNA repair and recombination factor Rad51 is over-expressed in human pancreatic adenocarcinoma. *Oncogene* **19**, 2791–2795 (2000).
83. Maacke, H. *et al.* Over-expression of wild-type Rad51 correlates with histological grading of invasive ductal breast cancer. *Int. J. Cancer* **88**, 907–913 (2000).
84. Mitra, A. *et al.* Overexpression of RAD51 occurs in aggressive prostatic cancer. *Histopathology* **55**, 696–704 (2009).
85. Russell, J. S. *et al.* Gleevec-mediated inhibition of Rad51 expression and enhancement of tumor cell radiosensitivity. *Cancer Res.* **63**, 7377–7383 (2003).
86. Lohse, I. *et al.* BRCA1 and BRCA2 mutations sensitize to chemotherapy in patient-derived pancreatic cancer xenografts. *Br. J. Cancer* **113**, 425–432 (2015).
87. Lord, C. J. & Ashworth, A. BRCAness revisited. *Nat. Rev. Cancer* **16**, 110–120 (2016).
88. Liu, Y. *et al.* RAD51 Mediates Resistance of Cancer Stem Cells to PARP Inhibition in Triple-Negative Breast Cancer. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* **23**, 514–522 (2017).
89. Montoni, A., Robu, M., Pouliot, É. & Shah, G. M. Resistance to PARP-Inhibitors in Cancer Therapy. *Front. Pharmacol.* **4**, (2013).
90. Hwang, T. L. & Shaka, A. J. Water Suppression That Works. Excitation Sculpting Using Arbitrary Wave-Forms and Pulsed-Field Gradients. *J. Magn. Reson. A* **112**, 275–279 (1995).
91. Shin, D. S. *et al.* Full-length archaeal Rad51 structure and mutants: mechanisms for RAD51 assembly and control by BRCA2. *EMBO J.* **22**, 4566–4576 (2003).
92. Aihara, H., Ito, Y., Kurumizaka, H., Yokoyama, S. & Shibata, T. The N-terminal domain of the human Rad51 protein binds DNA: structure and a DNA binding surface as revealed by NMR1 | Edited by P. E. Wright. *J. Mol. Biol.* **290**, 495–504 (1999).
93. Rohrwild, M. *et al.* HslV-HslU: A novel ATP-dependent protease complex in *Escherichia coli* related to the eukaryotic proteasome. *Proc. Natl. Acad. Sci. U. S. A.* **93**, 5808–5813 (1996).
94. Wojtkowiak, D., Georgopoulos, C. & Zylicz, M. Isolation and characterization of ClpX, a new ATP-dependent specificity component of the Clp protease of *Escherichia coli*. *J. Biol. Chem.* **268**, 22609–22617 (1993).

- 
95. Sousa, M. C., Kessler, B. M., Overkleeft, H. S. & McKay, D. B. Crystal structure of HslUV complexed with a vinyl sulfone inhibitor: corroboration of a proposed mechanism of allosteric activation of HslV by HslU. *J. Mol. Biol.* **318**, 779–785 (2002).
96. Martinez, A. *et al.* Expression of recombinant human phenylalanine hydroxylase as fusion protein in *Escherichia coli* circumvents proteolytic degradation by host cell proteases. Isolation and characterization of the wild-type enzyme. *Biochem. J.* **306** ( Pt 2), 589–597 (1995).
97. Nomme, J. *et al.* Inhibition of filament formation of human Rad51 protein by a small peptide derived from the BRC-motif of the BRCA2 protein. *Genes Cells Devoted Mol. Cell. Mech.* **13**, 471–481 (2008).
98. Esashi, F., Galkin, V. E., Yu, X., Egelman, E. H. & West, S. C. Stabilization of RAD51 nucleoprotein filaments by the C-terminal region of BRCA2. *Nat. Struct. Mol. Biol.* **14**, nsmb1245 (2007).
99. Falchi, F. *et al.* Synthetic Lethality Triggered by Combining Olaparib with BRCA2–Rad51 Disruptors. *ACS Chem. Biol.* **12**, 2491–2497 (2017).
100. Dalvit, C., Caronni, D., Mongelli, N., Veronesi, M. & Vulpetti, A. NMR-based quality control approach for the identification of false positives and false negatives in high throughput screening. *Curr. Drug Discov. Technol.* **3**, 115–124 (2006).
101. Acker, T. M. *et al.* Structure-activity relationships and pharmacophore model of a noncompetitive pyrazoline containing class of GluN2C/GluN2D selective antagonists. *J. Med. Chem.* **56**, 6434–6456 (2013).
102. Dalvit, C. *et al.* Identification of compounds with binding affinity to proteins via magnetization transfer from bulk water\*. *J. Biomol. NMR* **18**, 65–68 (2000).
103. Vulpetti, A., Hommel, U., Landrum, G., Lewis, R. & Dalvit, C. Design and NMR-based screening of LEF, a library of chemical fragments with different local environment of fluorine. *J. Am. Chem. Soc.* **131**, 12949–12959 (2009).
104. Galkin, V. E. *et al.* BRCA2 BRC motifs bind RAD51–DNA filaments. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 8537–8542 (2005).
105. Nakamura, C. E. & Abeles, R. H. Mode of interaction of .beta.-hydroxy-.beta.-methylglutaryl coenzyme A reductase with strong binding inhibitors: compactin and related compounds. *Biochemistry (Mosc.)* **24**, 1364–1376 (1985).

- 
106. Jencks, W. P. On the attribution and additivity of binding energies. *Proc. Natl. Acad. Sci.* **78**, 4046–4050 (1981).
107. Hann, M. M., Leach, A. R. & Harper, G. Molecular complexity and its impact on the probability of finding leads for drug discovery. *J. Chem. Inf. Comput. Sci.* **41**, 856–864 (2001).
108. Congreve, M., Carr, R., Murray, C. & Jhoti, H. A ‘rule of three’ for fragment-based lead discovery? *Drug Discov. Today* **8**, 876–877 (2003).
109. Jhoti, H., Williams, G., Rees, D. C. & Murray, C. W. The ‘rule of three’ for fragment-based drug discovery: where are we now? *Nat. Rev. Drug Discov.* **12**, 644 (2013).
110. Dalvit, C. *et al.* High-Throughput NMR-Based Screening with Competition Binding Experiments. *J. Am. Chem. Soc.* **124**, 7702–7709 (2002).
111. Mayer, M. & Meyer, B. Characterization of Ligand Binding by Saturation Transfer Difference NMR Spectroscopy. *Angew. Chem. Int. Ed.* **38**, 1784–1788 (1999).
112. Mayer, M. & Meyer, B. Group epitope mapping by saturation transfer difference NMR to identify segments of a ligand in direct contact with a protein receptor. *J. Am. Chem. Soc.* **123**, 6108–6117 (2001).
113. Meyer, B. & Peters, T. NMR spectroscopy techniques for screening and identifying ligand binding to protein receptors. *Angew. Chem. Int. Ed Engl.* **42**, 864–890 (2003).
114. Dalvit, C. Ligand- and substrate-based <sup>19</sup>F NMR screening: Principles and applications to drug discovery. *Prog. Nucl. Magn. Reson. Spectrosc.* **4**, 243–271 (2007).
115. Dalvit, C. NMR methods in fragment screening: theory and a comparison with other biophysical techniques. *Drug Discov. Today* **14**, 1051–1057 (2009).
116. Dalvit, C. *et al.* A General NMR Method for Rapid, Efficient, and Reliable Biochemical Screening. *J. Am. Chem. Soc.* **125**, 14620–14625 (2003).
117. Dalvit, C., Ardini, E., Fogliatto, G. P., Mongelli, N. & Veronesi, M. Reliable high-throughput functional screening with 3-FABS. *Drug Discov. Today* **9**, 595–602 (2004).
118. Jerabek-Willemsen, M. *et al.* MicroScale Thermophoresis: Interaction analysis and beyond. *J. Mol. Struct.* **1077**, 101–113 (2014).
119. Ritzefeld, M. & Sewald, N. Real-Time Analysis of Specific Protein-DNA Interactions with Surface Plasmon Resonance. *J. Amino Acids* **2012**, 816032 (2012).