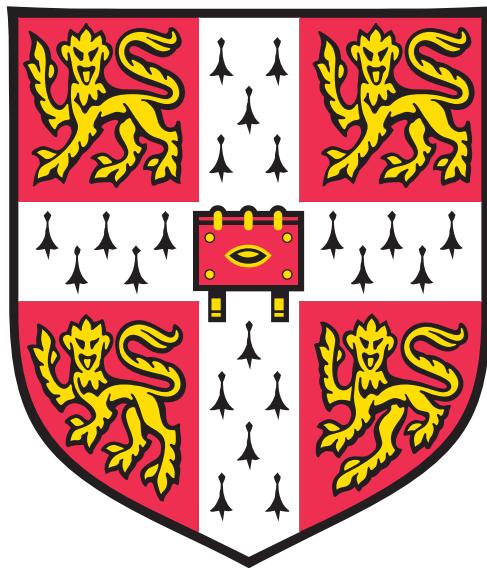# Interrogation and engineering of RAD51:BRC repeat interactions

Teodors Pantelejevs

Supervisor: Dr Marko Hyvönen

Department of Biochemistry

University of Cambridge

Submitted for the degree of Doctor of Philosophy

Declaration

This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration except as specified in the Acknowledgements and the main text. It is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution.

I further state that no part of my dissertation has already been submitted, or is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University or similar institution.

It does not exceed the prescribed word limit of 60,000 words for the Degree Committee of the Faculty of Biology.

Acknowledgments

Summary

The maintenance of genetic information is a fundamental function of every organism. The DNA in every human cell endures thousands of lesions per day, and eukaryotes have developed a sophisticated response to tackle this damage. Double strand breaks (DSBs) are the most toxic type of DNA damage, potentially leading to chromosomal rearrangements and tumorigenesis. Homologous recombination (HR) is a repair pathway that uses a homologous template to faithfully repair the DSBs. Central to HR is a recombinase protein RAD51, which forms a nucleoprotein filament with the broken ends, and allows efficient homology search. RAD51 is a promising therapeutic target in oncology.

BRCA2 is a crucial mediator of RAD51 function and interacts with RAD51 through a set of 8 BRC repeats. A BRC repeat consists of two modules, an FxxA and an LFDE module, each containing hot-spot residues that mediate binding. In this work, I set out to utilise these for the pharmacological targeting of human RAD51. First, using biophysical methods, I investigated a novel, high-affinity chimeric repeat BRC8-2 composed of FxxA and LFDE modules from different natural repeats. I determined the X-ray crystallographic structure of its complex, which allowed me to uncover the determinants of high affinity binding.

I applied these findings to the design of novel macrocyclic peptide inhibitors of RAD51. A method for cysteine stapling of recombinantly produced peptides was optimised to yield the correct product of high purity. The optimised stapling methodology provides a promising general approach for recombinant production and evaluation of cysteine-stapled peptides. Peptides were rationally designed in a structure-guided manner, and a variety of stapling architectures were evaluated for binding. The resulting purified molecules exhibit high potency and are stable in serum. Further crystallographic studies shed light on how the stapling moiety affects the peptide binding mode. The derived peptides serve as novel modalities for targeting Rad51 protein-protein interactions and can inform subsequent development of therapeutic drugs.

In addition to the human proteins, I investigated the BRC:RAD51 interactions in orthologs from *Leishmania infantum*, the causative agent of leishmaniasis. Using crystallography and biophysical methods, I uncovered novel features of BRC repeat binding. The presented work expands our understanding of the structural determinants of homologous recombination.

Abbreviations

ADP	Adenosine diphosphate
AID	Activation-induced cytidine deaminase
ATM	Ataxia-telangiectasia mutated (kinase)
ATP	Adenosine diphosphate
ATP-PNP	Adenylyl-imidodiphosphate
BIR	Break-induced replication
BTR	BLM, TopoIIIa, RMI1, and RMI2 (complex)
CD	Circular dichroism
dHJ	Double Holliday junction
DSB	Double-strand break
DSBR	double-strand break repair
DVP	Divinylpyrimidine
DVT	Divinyltriazine
EM	Electron microscopy
EMSA	Electrophoretic mobility shift assay
FP	Fluorescence polarisation
GB1	Protein G B1 domain
GTP	Guanosine triphosphate
HR	Homologous recombination
IC_{50}	Half-maximum inhibitory concentration
IMAC	Immobilised metal affinity chromatography
ITC	Isothermal titration calorimetry
K_D	Dissociation constant
MOA	Mechanism of action
MRN	Mre11, Rad50 and Nbs1 (complex)
NF	Nucleoprotein filament
NHEJ	Non-homologous end joining
NTD	N-terminal domain
OE	Oligomerisation epitope
PPI	Protein-protein interaction
RPA	Replication protein A
RPC	Reversed-phase chromatography
RTx	Radiation therapy
SDSA	Synthesis-dependent strand annealing
SPP	Solid-phase peptide synthesis
SPR	Surface plasmon resonance
TCEP	Tris(2-carboxyethyl)phosphine
TEV	Tobacco etch virus (protease)

Table of Contents

1 Introduction

Genomic integrity is critical for the survival of all forms of life, and successful repair of DNA lesions is an essential function of the cell. Excessive DNA damage and a compromised ability to repair are both a cause and a symptom of numerous pathological states. Crucially, genomic instability is a hallmark of cancer.[1] It has been estimated that a single human cell experiences tens of thousands of DNA lesions per day, of which double-strand breaks (DSBs) are the most severe type of genotoxic damage that can result in irreversible genomic rearrangements, aneuploidy and cell death.[2] The aetiology of DSBs is multi-faceted, and they can arise both as a result of exogenous and endogenous factors. Ionising radiation (IR) and genotoxic chemicals are examples of external agents that can cause DSBs either by directly breaking the DNA, or by generating reactive oxygen species that then react with the nucleic acid. DSBs can arise from stalled replication forks that are generated during normal DNA replication or when the replisome encounters a DNA lesion or modification.[3] DSBs also happen as a result of normal programmed cellular function, and in such instances are introduced by endogenous nucleases, for example, in meiotic recombination, where they are essential for the formation of chiasmata between homologous chromosomes, leading to expanded genetic diversity, or during V(D)J recombination that generates combinatorial immunoglobulin repertoires for the adaptive immunity.[4,5]

1.1 Double strand break repair

Eukaryotes have evolved a sensitive and highly organised response to DNA damage, which senses genotoxic events and elicits an appropriate repair cascade.[2] In eukaryotes, DSB repair can happen via a number of pathways, namely, homologous recombination (HR), non-homologous end joining (NHEJ)[6], microhomology-mediated end joining (MMEJ)[7] and single strand annealing (SSA)[8]. The relative preference of each repair pathway varies depending on the species, cell type, nature and source of the DSB, and is cell-cycle dependent, because the availability of a template sister chromatid is limited to S and G2 phases of cell cycle. Cyclin-dependent kinases thus have a critical role in pathway selection.[9] Moreover, the precise nature of the DSB influences the choice of pathway. For example, DSBs originating from stalled replication forks are one-ended and cannot be repaired by NHEJ and MMEJ, and are thus processed through HR, whereas in the case of IR-induced DSBs, there is competition between the pathways.[10,11]

1.2 DSB repair by homologous recombination

HR is the most faithful DSB repair pathway, as it employs a DNA template that is homologous to the broken locus in order to resynthesise the DNA at the double-strand break, and thus restores the original nucleotide sequence.[12] This is in contrast to NHEJ, MMEJ and SSA, which can introduce small but potentially detrimental changes to the genome. A mitotic sister chromatid is the preferred DNA donor template for HR, but repair can also proceed using the corresponding homologous chromosome or other homologous loci in the genome, which can lead to the loss of heterozygosity.[13,14] Due to the requirement for a sister chromatid, HR happens predominantly during S and G2 phases of cell cycle.[15]

Homologous recombination has been traditionally divided into three stages: pre-synapsis, synapsis and post-synapsis. Pre-synapsis is the preparative stage, during which the double strand break is recognised and processed, rendering it capable of searching for homologous sequences that will be used as repair template. During synapsis, the processed DSB ends sample genomic dsDNA for matching sequences, and, once a region of sufficient homology is found, form a long-lasting complex via strand exchange. During post-synapsis, the broken DNA end is re-synthesised and intact duplex products are restored.

1.3 Pre-synapsis

An end of a double strand break is first recognised by the MRN complex, consisting of the three proteins Mre11, Rad50 and Nbs1.[16] The MRN complex is the primary sensor of broken DNA ends in homologous recombination, and initiates the downstream biochemical cascade by recruiting other repair proteins to the break site, as well as exhibiting its own biochemical activity (**Figure 1**).

Mre11 is the central member of the MRN complex that interacts with the other two protein components and with DNA, and together with Rad50 forms a hetero-tetrameric MR core complex, which is conserved across all domains of life. In addition to recognising the DNA break ends, Mre11, through its N-terminal domain, exhibits both endonuclease and 3'-5' dsDNA exonuclease activities. This dual nuclease function allows Mre11 to process the DSB into short 3' ssDNA overhangs, first by nicking a DNA strand distally from the break end, and then by cleaving the strand in 3'-5'direction.[17] ATP binding to Rad50 initially induces a closed conformation on Mre11, which inhibits its exonuclease activity, leaving it function as an nicking endonuclease only. Subsequent hydrolysis of ATP by the Rad50 ATP-binding cassette

changes the conformation of the complex and renders exonuclease motifs accessible to permit short range resection.[18]

The resulting overhangs of up to 300 bp are then bulk-resected in the 5'-3' direction either by the exonuclease ExoI, or alternatively by the BTR-DNA2 complex, both recruited by MRN, producing single-stranded overhangs ranging from several hundred to tens of thousands of bases in length (**Figure 1**).[19,20,21] In BTR-DNA2-mediated resection, the RecQ-like BLM helicase, a member of the BTR complex, unwinds the 5' strand of a short-range resected end in an ATP-dependent manner, and feeds it to DNA2 for nucleolytic degradation.[22] ExoI, on the other hand, can process the DSB end in the absence of BLM helicase-mediated unwinding. However, both MRN and BLM stimulate ExoI activity by increasing its affinity for DNA.[19] For more detail on end resection, the reader is referred to a review by Symington.[23]

DNA end resection is the key step for committing the break for repair by HR, as resection by the MRN-CtIP complex causes the dissociation of Ku70-Ku80 from the break end, thereby inhibiting NHEJ.[24] It is therefore at this step that there exist many levels of regulation to favour one pathway or the other. Through Nbs1, the MRN complex recruits CtIP, on which many of these regulatory signals act. Upon T847 phosphorylation by CDKs during the S/G2 phases of the cell cycle, CtIP associates with MRN and stimulates the endonuclease activity of Mre11, leading to DSB end resection and thus favours DNA repair by HR.[25,26]

Nbs1 is involved in the nuclear translocation of the MRN complex, as it is the only component with a nuclear localisation signal.[27] Nbs1 also activates ATM kinase, a master regulator of HR, and recruits it to the DSB, which triggers the downstream activation of many repair factors to the break site.[28] ATM kinase has a profound role in inducing checkpoint signalling in response to DNA damage. For example, it phosphorylates and inhibits the MDM2 E3 ligase, leading to the accumulation of p53 tumour suppressor protein, and causes cell cycle arrest, therefore allowing sufficient time for the DNA damage repair.[29]
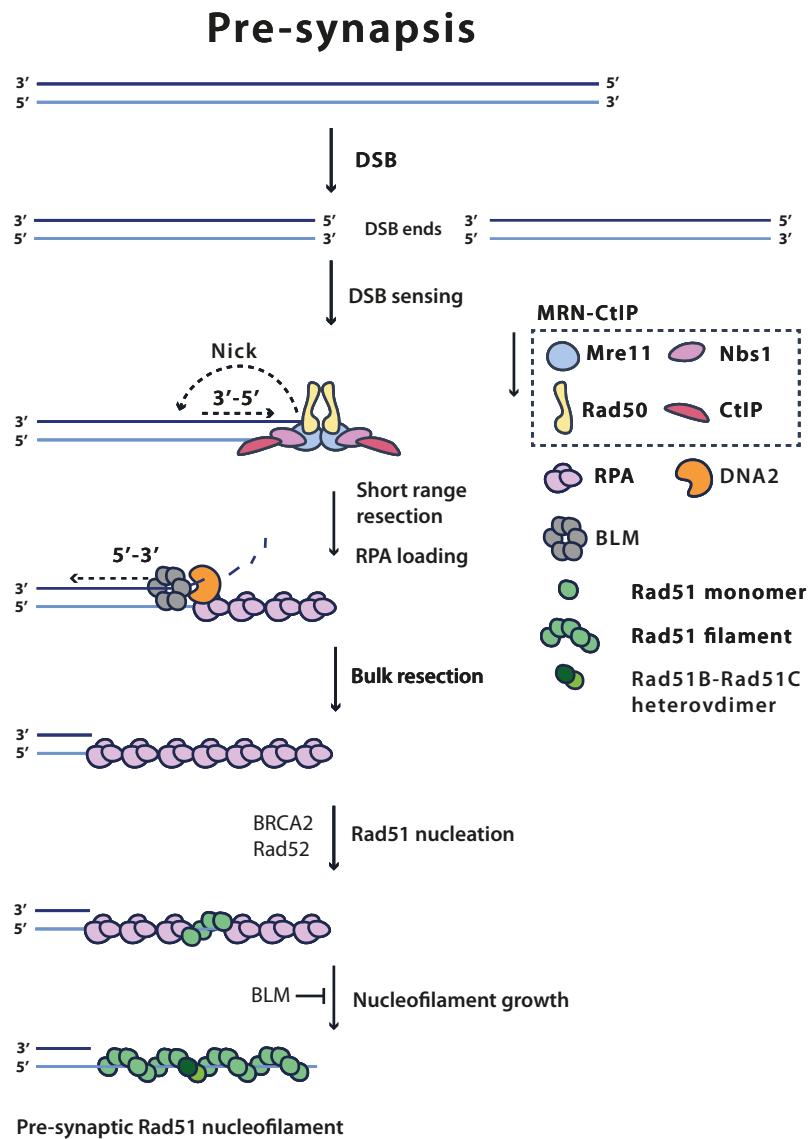
Pre-synapsis

Figure 1. Pre-synapsis. During the pre-synaptic stage of homologous recombination, a double-strand break end is processed to yield a RAD51-coated 3' ssDNA overhang. Many of the protein involved in this complex pathway have been omitted from the figure for simplicity.

The resulting ssDNA overhang is immediately coated with replication protein A (RPA), a ubiquitous stabilising factor that prevents the ssDNA from acquiring inhibitory secondary structure, prevents it from annealing with other homologous ssDNA, and promotes long-range strand resection by ExoI or BTR-DNA2 (**Figure 1**).[20,30] Moreover, it protects the ssDNA from nucleolytic cleavage.[31] Eventually RPA is displaced from the ssDNA by RAD51, which is the principal recombinase catalysing the strand exchange reaction and the primary focus of this dissertation. *In vitro*, RPA has been shown have both inhibitory and stimulatory effects on RAD51 loading. If RPA is added after RAD51 nucleation has been allowed to happen, RPA stimulates further growth of RAD51 nucleofilaments by eliminating inhibitory secondary structure, and thus promotes strand exchange.[32] However, in a cellular context, RPA is more

abundant and has higher affinity for ssDNA than RAD51, and therefore is rapidly bound to nascent ssDNA prior RAD51 loading.[33] Such RPA binding inhibits nucleation by competing for the same ssDNA and sterically precluding RAD51 binding, implying that additional mediators are necessary for the exchange of the two proteins.[34]

In budding yeast, RAD51 loading and displacement of RPA is mediated by RAD52. Rad52 interacts with RPA, DNA and RAD51, and has been proposed to seed RAD51 nucleation events on RPA-coated DNA.[34] Once a nucleation event happens, RAD51 nucleofilament extension proceeds readily due to the cooperative interactions between RAD51 protomers. *S. cerevisiae* Rad52 is essential for homologous recombination and loss-of-function mutants display dramatic hypersensitivity to DNA-damaging agents.[35] Interestingly, in mammals, loss of RAD52 is tolerated and does not exhibit a strong phenotype, indicating the existence of another gene with similar function.[36] Indeed, nucleation of RAD51 on ssDNA and the concomitant displacement of RPA in mammals is dependent on stimulation by BRCA2, a crucial mediator that is not found in *S. cerevisiae*.[37] In BRCA2-deficient human cells, RAD52 is able to compensate for loss of BRCA2 activity, but synthetic lethality arises when both gene products are depleted.[38] The two proteins therefore have overlapping function as stimulators of RAD51 loading. This explains why some eukaryotes that have a BRCA2 ortholog, such as *C. elegans* and *D. melanogaster*, do not have a RAD52 homolog at all.[39] BRCA2 and its function are discussed in detail later in this work.

In addition to the mediator proteins RAD52 and BRCA2, seven RAD51 paralogs that share limited homology to RAD51 have been identified in vertebrates: RAD51B, RAD51C, RAD51D, XRCC2, XRCC3 and SWSAP1, along with the meiotic recombinase DMC1. RAD51 paralogs have arisen as a result of gene duplication events and mutagenesis or knock-out of each of them sensitises cells to DNA damage or confers embryonic lethality to the organism.[40,41,42] RAD51 paralogs act as RAD51 mediators downstream of BRCA2-stimulated nucleofilament formation.[43] For example, RAD51B and RAD51C form a heterodimeric complex that lacks recombinase activity but is capable *in vitro* of integrating into the RAD51 nucleofilament, exerting a stabilising effect on it and conferring enhanced strand exchange activity, while preventing RAD51 disassembly by anti-recombinogenic helicases (**Figure 1**).[44]

1.4 Synapsis and post-synapsis

RAD51 binding to ssDNA results in a pre-synaptic nucleofilament (NF): an oligomeric, filamentous assembly that is capable of searching for a double-stranded homologous template.[45,46] During homology search, the pre-synaptic nucleofilament probes the genomic

DNA for matching sequences by forming dynamic, transient interaction with the duplex that are stabilised when a sufficient number of base-paired contacts arise.[47]

Once a template match is found, strand exchange proceeds, and the presynaptic NF displaces the strand that is (+) to the ssDNA overhang and anneals to the complementary (-) template strand, forming a short-lived synaptic complex intermediate which contains all three strands within the nucleofilament. The structural details of the synaptic complex are poorly understood. Eventually, the (+) template strand becomes completely displaced, leaving a heteroduplex post-synaptic nucleofilament. The resulting total structure, containing the newly formed heteroduplex and the displaced strand is called a displacement loop (D-loop, **Figure 2**).

The annealed 3' overhang within the D-loop serves as a primer for DNA synthesis by DNA polymerases, which use the homologous DNA as template, further unwinding the template duplex and extending the D-loop in a displacement synthesis step.[48] For this to happen, the motor protein Rad54 removes terminal RAD51 protomers from the invading 3' strand in an ATP-dependent manner to allow polymerase binding to the priming region.[49,50] Following D-loop formation, homologous recombination can proceed two main pathways: double strand break repair (DSBR), synthesis-dependent strand-annealing (SDSA) (**Figure 2**).
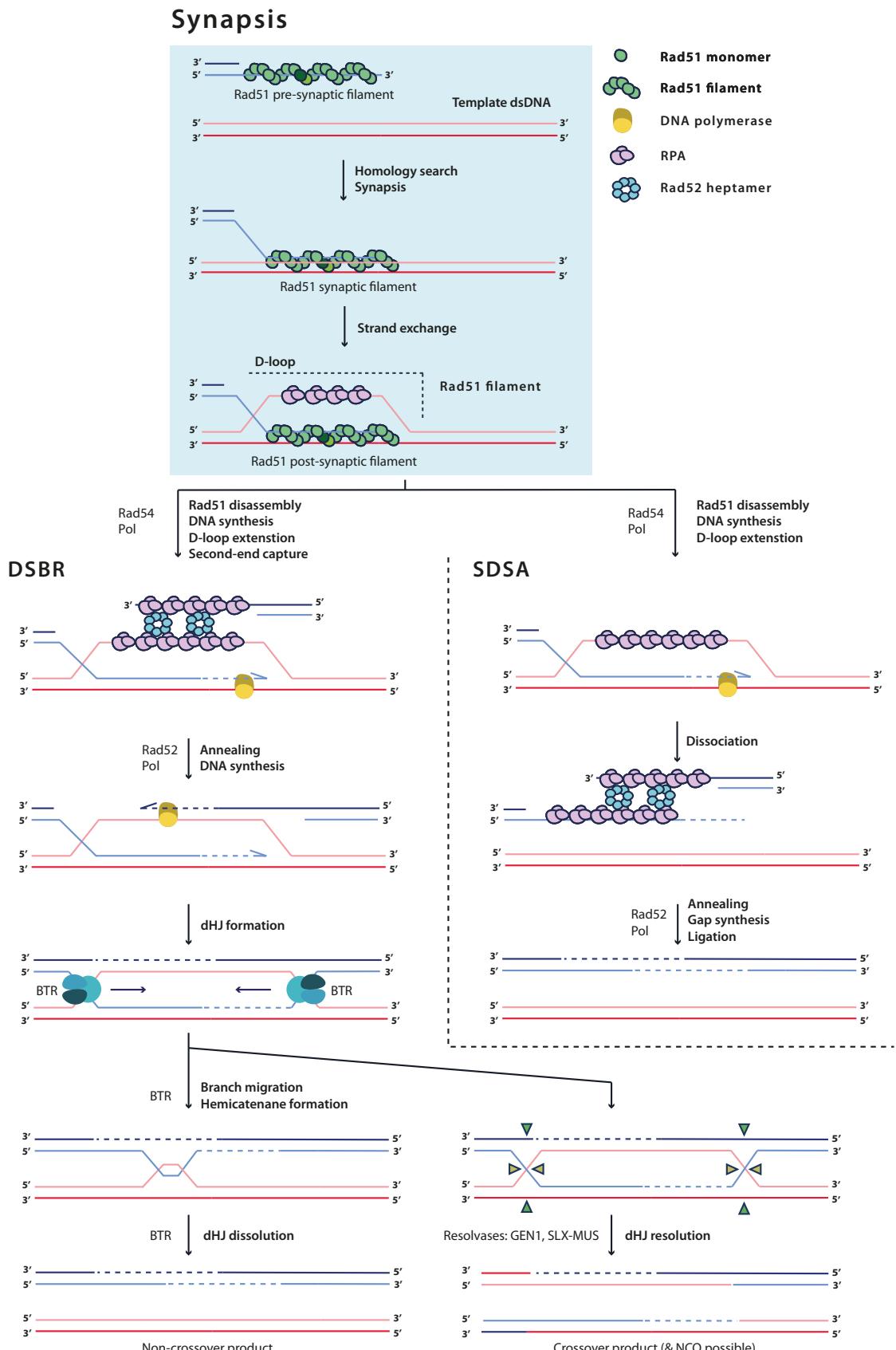
***Figure 2**.* Synapsis and post-synapsis. A generalised schematic depicting the pre-synaptic stage involving homology search and strand invasion (blue background, second end not shown), followed by the two main post-synaptic pathways of HR-mediated DSB repair, DSBR and SDSA. Some of the critical proteins involved in the pathway are provided.

1.4.1 DSBR

The critical step that differentiates between the post-synaptic pathways is the treatment of the second DSB end, so far excluded from the discussion. In DSBR, the second end of the DSB is similarly processed to yield a 3′ overhang, which is then coated with RPA. Then, in a step called "second end capture", it is annealed by Rad52 to the displaced template strand within the extended D-loop, which now has an extensive displaced ssDNA strand (**Figure 2**).[51] Annealing of RPA-coated ssDNA is another function of Rad52 besides its pre-synaptic mediator activity.[51] This results in a second heteroduplex product that can serve as a priming region for a DNA synthesis in the opposite direction.[51] The elongation of the two strands is eventually halted and the newly synthesised 3′ extensions at both ends are ligated to the corresponding 5′ of the opposite ends, resulting in the formation of a double Holliday junction (dHJ).

Holliday junctions are DNA secondary structure elements that contain two homo-duplex and two hetero-duplex dsDNA arms, joined together at a single eight-stranded intersection. To restore intact linear chromosomes after nascent end synthesis, the dHJs are either dissolved or resolved (**Figure 2**).[52] Dissolution entails the convergent migration of the two HJs towards each other and their eventual collapse into a hemicatenane structure, catalysed by the BLM helicase activity within the BTR complex (also called "dissolvasome", STR in yeast), followed by the unlinking of the interwound strands by topoisomerase IIIα, also part of the BTR complex.[53,54] Dissolution results exclusively in non-crossover recombination products, with the DSB ends re-connected and intact template restored. Non-crossover products produced by dissolution are preferred to crossover in mitotic HR in order to avoid loss-of-heterozygosity (LOH), which can arise when a homologous chromosome, rather than a sister chromatid, is used as template for repair.[52] BTR-mediated dissolution is therefore the predominant pathway for dHJ processing and occurs at the early stages of cell cycle.[52]

Resolution, on the other hand, acts as a back-up pathway later in mitosis, to ensure that no residual links between chromatids exist before chromatid segregation.[52] Resolution is catalysed by nucleolytic resolvase enzymes, such as mammalian GEN1 or the SLX-MUS complex, that nick the dHJs at symmetric positions across the junction, allowing the non-covalent separation of two nicked homo-duplexes.[52] Because resolvases can cut at two different orientations at each of the holiday junctions, a range of crossover and non-crossover products are possible, which can be lead to detrimental loss of heterozygosity when a homologous chromosome, rather than a sister chromatid, is used as template.[55]

1.4.2   SDSA

During SDSA, the first captured end is synthesised in the absence of second end capture, and double holiday junctions are not formed (**Figure 2**). The second end is also resected but remains passive. Instead, the extended invading strand is eventually displaced by the branch migration activity of the BLM helicase within the BTR complex and can re-anneal to the second end.[56] The resulting ssDNA gaps are filled, and the nicks ligated, forming exclusively non-crossover products. SDSA is the dominant HR pathway for repairing mitotic DSBs in mammalian cells, but also occurs in meiotic cells.[57] Because a dHJ is not formed during SDSA, the pathway results in exclusively non-crossover products .

1.5   The RAD51 recombinase

The eukaryotic Rad51 gene was first identified and mapped in a screen for *Saccharomyces cerevisiae* mutants that are sensitive to X-ray radiation.[58] Its involvement in the maintenance of genomic integrity stemming from its recombinase activity, as well as its functional and structural homology to bacterial RecA, have since been extensively characterised.[59] Today, RAD51 is recognised as a ubiquitous mediator of homologous recombination that is central for this mode of DNA repair. It is essential for cellular survival in higher eukaryotes, for example, RAD51-null chicken DT40 cells accumulate chromosomal breaks and eventually die.[60] Embryonic lethality arises as a result of defective Rad51 in mammals.[61] The RAD51/RecA family of proteins stems from a common evolutionary origin and is highly conserved: eukaryotic RAD51 orthologs have >60% sequence identity, which increases to >97% when only mammalian sequences are examined.[62] Gene duplication events in eukaryotes have resulted in RAD51 paralogs with divergent or specialised function, such as the meiosis-specific recombinase DMC1 or the RAD55-RAD57 paralogs, which are mediators of RAD51 function.[63,64]

1.5.1   Biochemical characteristics of RAD51

The first extensive biochemical characterisation of purified human RAD51 (*Hs*RAD51) *in vitro* was reported by Baumann *et al*. not long after the identification of the human ortholog.[65] The study showed that human RAD51 binds both ssDNA and dsDNA *in vitro* with comparable affinities, and has a slightly higher affinity for the single-stranded substrate than the duplex. The stoichiometry of DNA binding corresponds to 3 nucleotides per RAD51 protomer, and is conserved across species.[65] The ATPase activity of *Hs*RAD51 is orders of

magnitude lower than that of RecA, and it is modulated by DNA binding. [65] Steady-state ATP turnover by RAD51 increases from 0.03 molecules per minute in the absence of DNA, to 0.05 and 0.16 ATP/min when dsDNA and ssDNA is added, respectively.[65] RAD51 catalyses the strand exchange between a ssDNA molecule and a homologous dsDNA template, leading to the formation of joint molecules, a process that is ATP/Mg^{2+}-dependent, but also happens in the presence of other nucleotides, such as GTP, UTP, dATP or the non-hydrolysable ATPyS, as shown in the same study.[65] However, given the higher abundance of ATP in the cell, it is unlikely that the other nucleotides contribute significantly to NTPase activity.

RAD51 polymerisation on DNA occurs in two phases: a rate-limiting nucleation phase and a faster growth phase. As a result of RAD51 forming oligomeric species in solution, initial nuclei are formed by the binding of short RAD51 oligomers, with a minimum length 2-3 protomers, sufficient as a stable nucleating species.[66] Gradual coverage of DNA happens through the formation of many such nucleation sites, followed by their extension through short-range growth and eventual joining.[66]

The mechanism of homology search had long eluded detailed examination due to the dynamic features of the process being obscured by bulk ensemble measurements. Single molecule optical trap experiments with RecA, the bacterial homolog of RAD51, revealed that efficient homology search involves hopping of the pre-synaptic NF between spatially proximate but sequence-distant segments of the randomly coiled dsDNA, in a process termed 'intersegmental contact sampling'.[67] It is likely that during these non-specific dsDNA binding events, the pre-synaptic nucleofilament samples the duplex sequence via base-flipping, where the duplex Watson-Crick hydrogen-bonding interactions are transiently disrupted and matched with the pre-synaptic NF.[47]

ATP hydrolysis is not required for strand exchange to happen, and hydrolysis-incompetent *Hs*RAD51 K133R mutant can rescue DNA repair function in RAD51-null chicken DT40 cells.[68] Moreover, the same mutant can form pre-synaptic NF and catalyse strand exchange more efficiently *in vitro* than the wild-type protein.[69] Similarly, Ca^{2+} ions, which inhibit ATPase activity, have been shown to enhance the strand-exchange activity of RAD51.[70] It has therefore been suggested that ATP hydrolysis functions in the dynamic disassembly and turnover of RAD51 from the nucleofilament, rather than the strand-exchange reaction itself.[71]

RAD51/RecA family recombinases have to dissociate from the post-synaptic dsDNA nucleofilament product to allow the execution of subsequent steps of recombination, such as DNA synthesis and Holliday junction resolution. *In vitro*, bacterial RecA rapidly dissociates from dsDNA as a result of ATP hydrolysis, whereas eukaryotic RAD51-dsDNA

nucleofilaments stay associated after ATP has been converted to ADP.[66] Additional protein co-factors are required for post-synaptic disassembly.

1.5.2 Rad51 structure

The human RAD51 gene product is a 37 kDa protein that consists of two globular domains, a C-terminal ATPase domain and a smaller N-terminal domain, which are linked by a flexible linker called the oligomerisation epitope (OE) (***Figure 3***A). The ATPase domain has a classic Rossmann fold common to many nucleotide-binding proteins, such as helicases or ATP synthases,[72,73] containing an extended core β-sheet that is surrounded by α-helices at both faces, resulting in a three-layered sandwich tertiary architecture. The central β-sheet consists of mixed parallel (β3,β2,β4,β5,β1,β6) and anti-parallel (β7,β8,β9) strands. Two flexible DNA-binding loops, L1 and L2, protrude from one side of the ATPase domain and interact with DNA upon formation of a nucleofilament. The L2 loop is significantly larger and is only partially resolved in structural models due to conformational flexibility (see **Section 1.5.4**, **Figure 6**C).

Near the DNA binding loops on the ATPase domain, a nucleotide-binding site is formed by Walker A (also called P-loop) and Walker B motifs that interacts with the nucleotide di- and tri- phosphates in a metal-ion dependent fashion and make up the catalytic centre for ATP hydrolysis (***Figure 3***B). The interaction with nucleotides is further stabilised by additional contacts between the nucleobase and the ATPase domain outside of the Walker A motif, which determines the specificity for ATP as catalytic co-factor.[74]

The N-terminal domain of human RAD51 is smaller than the ATPase and composed entirely of α-helices. It interacts with both ssDNA and dsDNA, however, its exact functional role has not been conclusively elucidated.[75]
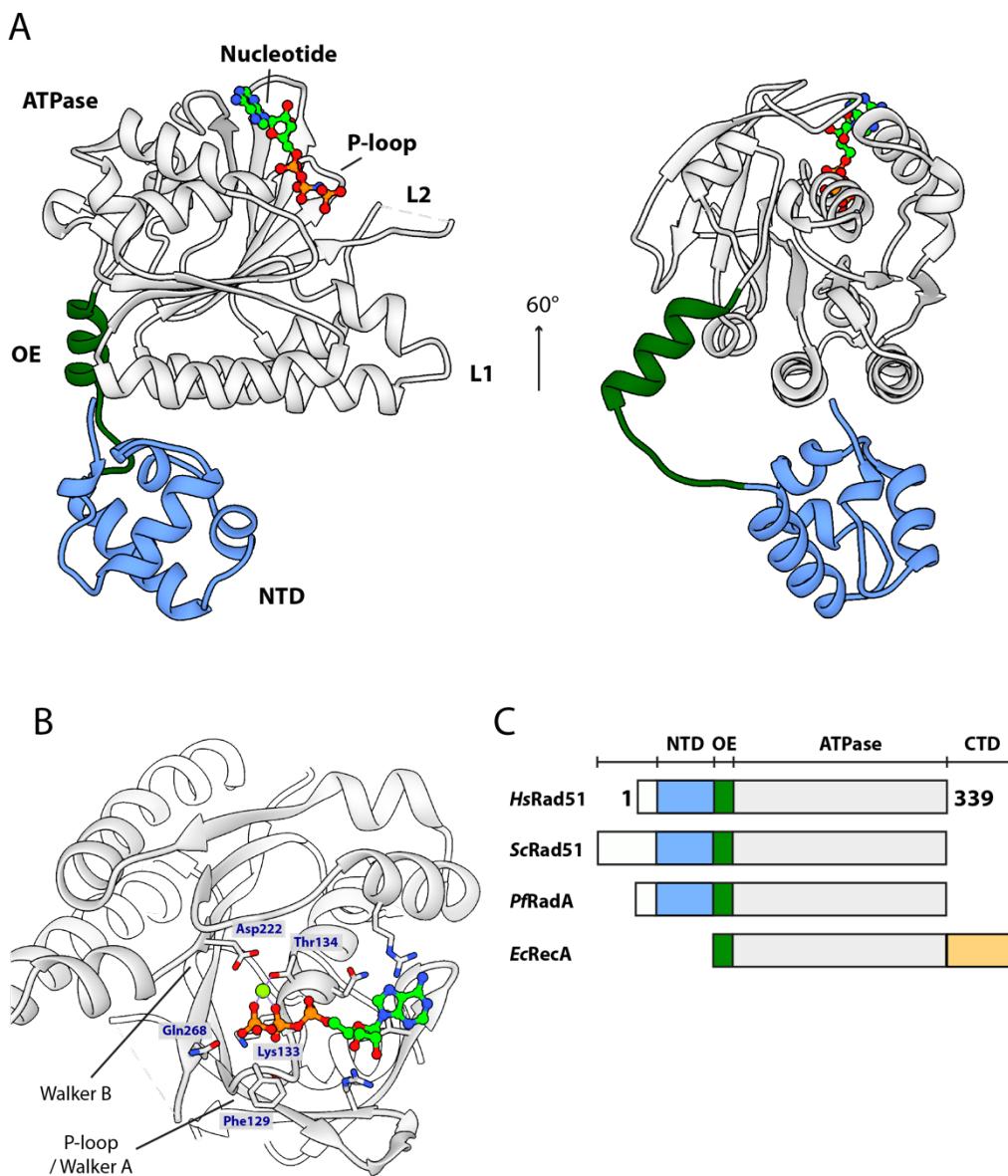
Figure 3. Tertiary structure of a RAD51 protomer. (**A**) A ribbon representation of a RAD51 monomer showing the different domains: N-terminal domain (NTD, blue), oligomerisation epitope (OE, green), ATPase domain (grey). Nucleotide (AMP-PNP) depicted in light green bound to the ATP site. L1, L2 – DNA binding loops. Electron density lacking for L2. Monomer extracted from a cryo-EM structure of the active ssDNA-bound nucleofilament (PDB ID: 5H1B).[76] (**B**) RAD51 ATP-binding site with the nucleotide cofactor bound. Triphosphate is seen binding the P-loop (also called Walker A motif) and an Mg^{2+} ion cofactor. Additional interaction through the adenine base (right) determine specificity for ATP (PDB: 4A6X).[74] (**C**) Comparison of domain organisation in different RAD51 orthologs.

A comprehensive discussion of RAD51 structure and function inevitably involves looking at different orthologs of the protein. Human RAD51 (*Hs*RAD51) has ostensibly been the focus of most research, however, many published experiments have been conducted with orthologs from other species. Most notable examples are: *Saccharomyces cerevisiae* Rad51

(*Sc*Rad51), *Escherichia coli* RecA (*Ec*RecA), *Pyrococcus furiosus* and *Methanococcus voltae* RadA (*Pf*RadA, *Mv*RadA). While their functional and structural properties remain broadly similar, there are differences worth noting. First, the domain organisation in bacterial RecA proteins is different from eukaryotes and archaea (***Figure 3*C**), reflecting the phylogenetic distance between the different domains of life. The N-terminal domain as seen in eukaryotes and archaea is absent in bacteria, while the oligomerisation epitope is similarly found N-terminal to the ATPase domain. Instead, a C-terminal domain of similar size to eukaryotic NTD, but different evolutionary origin and different fold, is present in bacteria and is believed to have arisen by convergent evolution to exert a similar function. Secondly, putative disordered extensions of varying length are observed in different homologs at the protein termini. For example, *Sc*Rad51 has a long, disordered N-terminal extension of unknown function. For a note on Rad51 nomenclature, see [*].

1.5.3 Rad51 oligomerisation

It has been demonstrated by structural and biophysical methods that various RAD51 orthologs and paralogs form soluble oligomers in the absence of DNA both *in vitro* and *in vivo*. The shape and assembly size of free RAD51 oligomers differs between species: for example, bacterial *Ec*RecA forms closed, helicase-like hexameric rings[77] while archaeal RadA from *Pyrococcus furiosus* forms closed heptameric ring dimers[78] (**Figure 4A**, left) as evidenced by crystallography, electron microscopy and DLS. Human and other eukaryotic RAD51 orthologs, on the other hand, can form linear oligomers of varying assembly size. For example, yeast Rad51 crystallises as a linear filament (**Figure 4A**, right).[79] It has been shown that increasing concentration of purified HsRAD51 shifts the distribution of oligomeric species towards higher molecular weight *in vitro*.

---

[*] **A note on nomenclature**. In this work, depending on the context, "RAD51" in uppercase refers to the human protein and is used interchangeably with *Hs*RAD51, or to the protein family as a whole. If other non-human species are used, an appropriate prefix, such as "*Sc*", is applied to a lowercase "Rad51". Archaeal and bacterial Rad51 homologs are exclusively called "RadA" and "RecA", respectively. Surrogate systems based on thermostable archaeal RadA are described later in the text. These contain a RadA root and a prefix that denotes which protein the system mimics, e.g. HumRadA22 is a *Pf*RadA protein that mimics human RAD51.
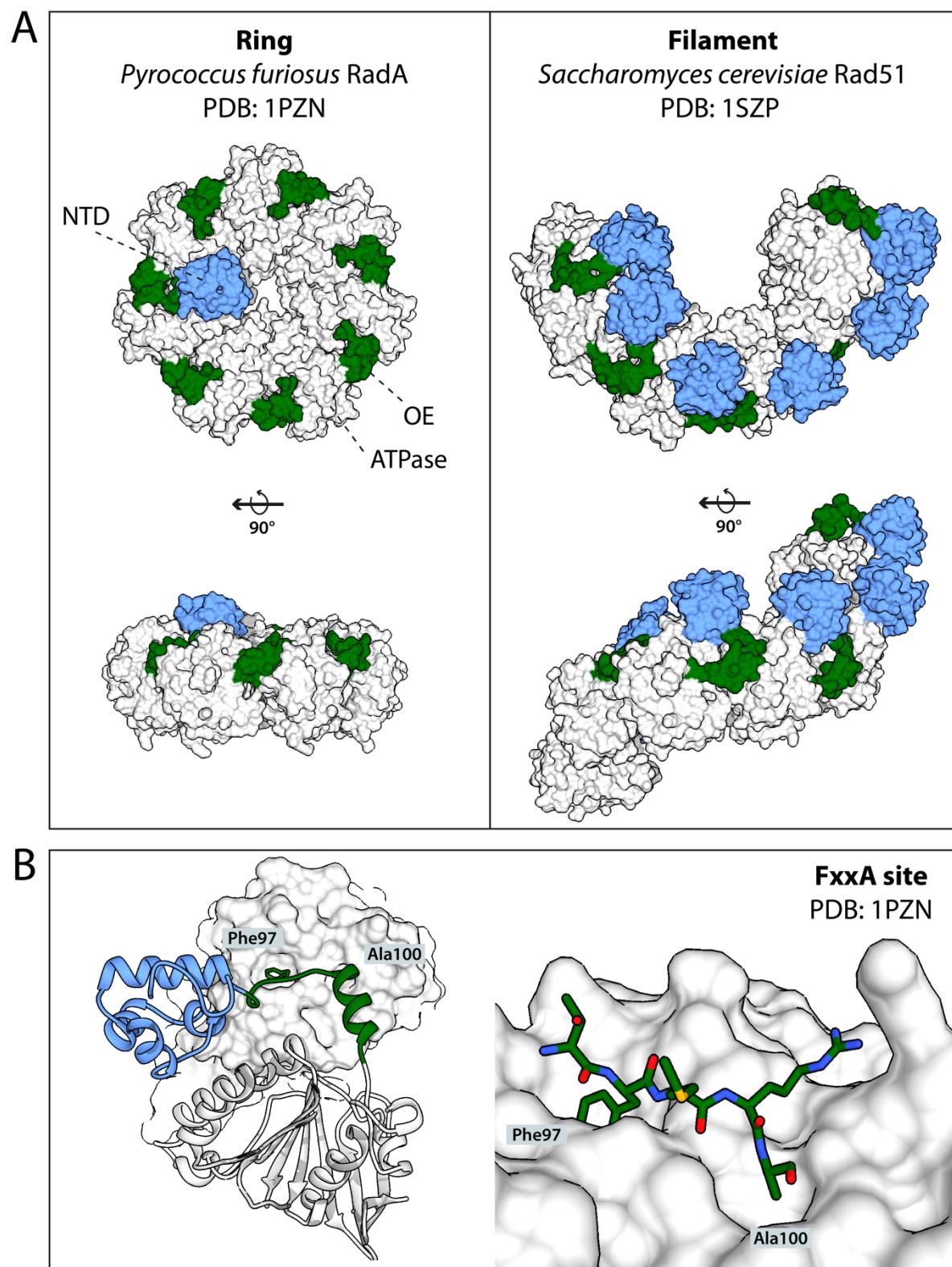
Figure 4. Rad51 oligomerisation happens through a conserved mechanisms and gives rise to different oligomeric species such as (**A**) heptameric rings formed by *Pyrococcus furiosus* RadA (left, PDB: 1PZN) or helical filaments in *Saccharomyces cerevisiae* Rad51 (right, PDB: 1SZP). (**B**) Rad51 oligomerisation is mediated by the FxxA motif located on the oligomerisation epitope. The FxxA motif binds through a combination of hydrophobic contacts via the Phe and Ala side-chains, and backbone-backbone hydrogen bonds (PDB: 1PZN).

The oligomerisation mechanism is highly conserved in the RAD51/RecA family in both ring and filament structures, and is likely to have arisen early in the evolution of its common ancestor. The bulk of the interface between adjacent RAD51 protomers is mediated by the oligomerisation epitope (OE), located between the NTD and the ATPase domains (**Figure 4B**). The OE contains a conserved FxxA motif, whose hydrophobic Phe and Ala side chains bind cognate pockets on the ATPase domain of a second RAD51 protomer. Importantly, the FxxA phenyl is buried within a deep hydrophobic cavity and this hot-spot residue is responsible for a large fraction of the free energy of binding.[80,81] Hydrogen-bonds are formed by FxxA motif backbone amides with the upper-most β-strand of the ATPase, resulting in an anti-parallel β-sheet augmentation. This interface is also maintained in the DNA-bound nucleofilament state (discussed below). The oligomerisation epitope is flexible, which allows a significant degree of conformational freedom between adjacent RAD51 subunits: while the FxxA contacts are relatively static, the orientation of ATPase and N-terminal domains can vary significantly. This flexibility of the OE explains the variety of oligomerisation morphologies observed in different orthologs and different assembly states. The closed ring arrangements seen in structures of *Pf*RadA (**Figure 4A**, left) and *Ec*RecA require the ATPase domains to remain co-planar, with a pure rotational symmetry axis crossing the ring centroid. Movement between consecutive ATPase domains results in a rise along this axis and transforms the oligomer into helical structure with a screw symmetry, like observed for the *Sc*Rad51 filament crystal structure (**Figure 4A**, right).

The transition from a ring conformation, having no pitch, to a filament with a non-zero pitch, involves the rotation of the globular ATPase domains relative to each other. For this to happen, backbone movement occurs in a flexible part of the OE immediately following the FxxA motif and preceding the ATPase domain, which corresponds to residues Glu98-Thr103 in *Hs*RAD51. This relatively small-scale local conformational transition enables large-scale movement at the DNA-binding region on the opposite side of the ATPase domain, as illustrated in **Figure 5**.
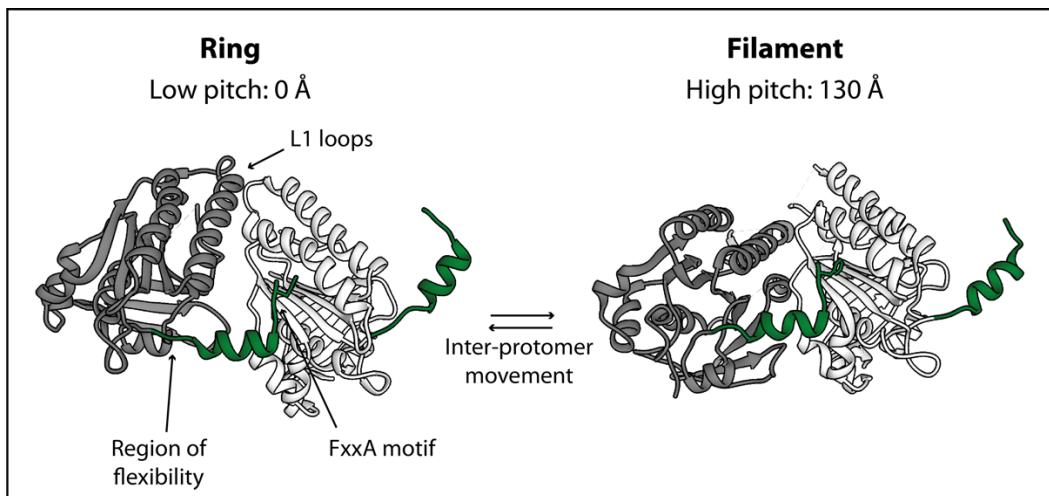
Figure 5. Movement between adjacent protomer ATP domain determines the assembly morphology. Conformational change at a region of flexibility in the oligomerisation epitope is required for the ATPase domains to rotate relative to each other. Image on the left shows a dimer extracted from a *Pf*RadA heptameric ring (PDB: 1PZN) while the dimer on the right is from a high-pitch *Sc*Rad51 filament X-ray structure (PDB: 1SZP). One ATPase domain is held in identical position in both images (light grey), while the other is allowed to move (dark grey).

1.5.4 RAD51 nucleofilament: structure and function

To exert its catalytic function, RAD51 first has to form an active, ATP-bound presynaptic nucleoprotein filament (NF) with the invading ssDNA strand. Human RAD51 can also bind DNA in the absence of nucleotide cofactor or if bound to ADP, however, these nucleoprotein filament states are unable to catalyse strand exchange. While ATP-bound state is essential for an efficient homology search[82], ATP hydrolysis is not necessary for this process[68,69] and hydrolysis-inhibited Ca^{2+}-bound state has enhanced strand-exchange activity, whereas the conversion ATP to ADP decreases it.[70]

Structures of RAD51 NFs have been determined at various resolutions using cryo-electron microscopy for the human protein and its orthologs. Early low-resolution studies with *Ec*RecA established that the prokaryotic NF has a helical morphology, where RecA oligomers wind around the ssDNA or dsDNA, forming a right-handed helix with the DNA strand lying close to the filament axis [83,84]. A similar morphology was later also observed for human RAD51 (**Figure 6A**). NFs are flexible with significant local polymorphism, contain ~6 protomers per turn and three bases or base-pairs per protomer, and maintain DNA in an under-wound, extended conformation compared to classical Watson-Crick B-DNA structure.

EM studies also revealed that the identity of the bound nucleotide co-factor significantly modulates a number of helical parameters in the NF. Importantly, the helical pitch,

that is, the height of a full turn of a helix along a helical axis, is significantly higher in the active ATP-bound RecA (~95 Å) versus the inactive apo/ADP-bound states (65-85 Å).[85]

Two relatively recent publications by Xu et al.[76] and Short et al.[86] report remarkably similar near-atomic resolution structures of the active pre-synaptic human RAD51 NF in the presence of a non-hydrolysable ATP mimetic AMP-PNP (**Figure 6B**). In these, the active filament has an average helical pitch of 103 Å and an inter-protomer displacement of 57°, closely resembling the structures of active yeast Rad51 and bacterial RecA filaments. 6.3 protomers or 19 nucleotides are positioned per turn of the right-handed NF helix, in line with previously reported RAD51:ssDNA binding stoichiometry of 1:3. The ssDNA in the pre-synaptic NF is extended by approximately a half compared to the length of duplex B-DNA.

In these structures, nucleobases are arranged as tightly stacked triplets resulting in a B-DNA-like stacking that favours Watson-Crick base pairing with the homologous template during strand invasion (**Figure 6B,C**). Each RAD51 protomer forms distinct contacts with three sets of adjacent nucleotide triplets in a regular, repetitive manner. The bulk of the RAD51-DNA interface is formed between the ssDNA phosphate backbone and the DNA-binding loops, thus orienting the nucleobases towards the helical axis (**Figure 6C**). Hydrogen bonding and charge-charge interactions make up majority of this interface. Val273 side chain from loop L2 of each RAD51 protomer inserts between every base triplet to enforce the tight stacking in an otherwise extended sugar-phosphate backbone conformation.

Three distinct interfaces contribute to the protein-protein interaction between adjacent RAD51 protomers in the active NF (**Figure 6D**). The largest interface is mediated by the FxxA motif from the oligomerisation epitope in a similar fashion to free RAD51 oligomers. A second, much smaller interface results from the stacking of Tyr54 aromatic side-chain, located on the NTD of one protomer, with the Phe195 on the ATPase domain of an adjacent protomer situated in the 5′ direction of the bound ssDNA. A third interface in the active NF is formed by a nucleotide binding site that is occupied by the AMP-PNP molecule and a metal ion co-factor, which are then covered by a different set of residues from the ATPase domain of the adjacent protomer, resulting in a tightly sandwiched nucleotide binding mode.
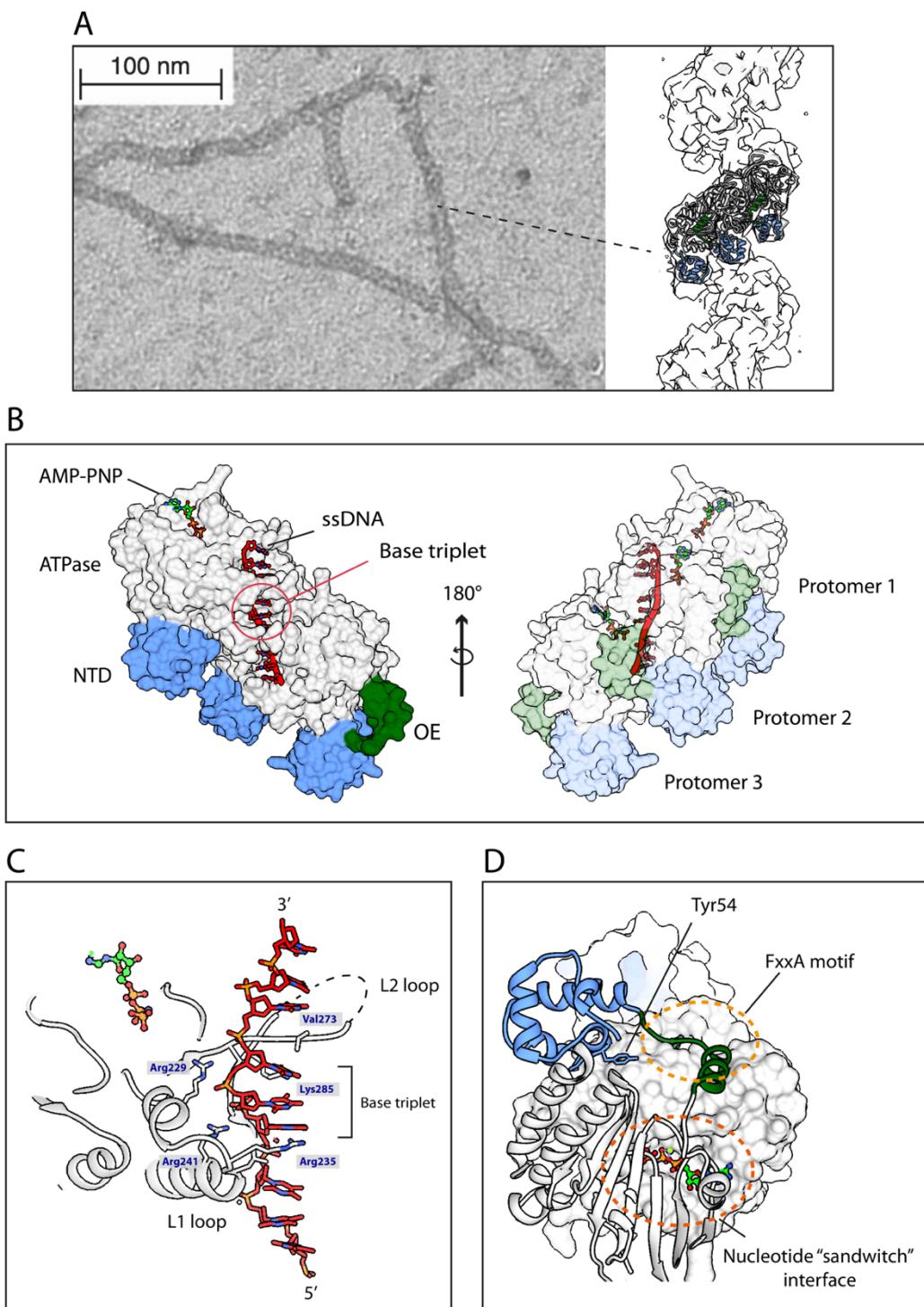
Figure 6. Pre-synaptic RAD51-ssDNA nucleofilament structure. (**A**) Electron micrograph of RAD51-ssDNA nucleofilament (reported by Davies and Pellegrini[87]) (**B-C**) High-resolution cryo-EM structure of the RAD51-ssDNA nucleofilament[76] (PDB: 5H1B). (**B**) A three-protomer segment of pre-synaptic NF. Three nucleotides in ssDNA bind per single protomer with nucleobases stacked in triplets and pointing outward for efficient base-pairing. Non-hydrolysable ATP mimetic AMP-PNP molecules (light green) are sandwiched between adjacent protomer ATPase domains (grey), stabilising the extended filament conformation. (**C**) RAD51-ssDNA interaction is mediated by the two DNA-binding loops L1 (Arg229-Ala240) and L2 (Val269-Asn290) protruding from the ATPase domain. (**D**) Three distinct interfaces form between adjacent RAD51 protomers in the active nucleofilament.

1.5.5 Allosteric regulation of RAD51 function

Several allosteric pathways have been proposed to control the different steps of RAD51-catalysed strand exchange. ATP hydrolysis rates increase when RAD51 binds DNA[65,88] and such up-regulation of ATPase activity likely stems from conformational change induced within the RAD51 filament by DNA binding and transmitted to the ATPase active site. A report by Chen and colleagues[89] proposed that this DNA-sensing pathway is transmitted through a cis-trans switching of a peptide bond between *Sc*Rad51 Asp280-Ser281 dyad (Asp222-Ser223 in *Hs*RAD51) located at the junction between the DNA-binding loop L1 and the catalytic Walker B residues (**Figure 7**). The peptide bond linking the two residues is observed in the energetically unfavourable *cis* conformation in all DNA-free structures of RAD51/RecA family members. In DNA-bound nucleofilament structures, on the other hand, the peptide bond switches to a *trans* conformation. The authors propose that this brings the *Sc*Rad51 catalytic Glu221 side-chain closer to the reaction centre and increases catalytic rates as a result of DNA binding. However, no further studies have confirmed the validity of the proposed mechanism of activation. An earlier study on *Pyrococcus furiosus* RadA by Shin and colleagues examined the effect of recombinase-disabling L1 loop mutations on ATP hydrolysis rates, and found only marginal effects on ATPase activity, indicating that other mechanisms than L1-Walker B translation could be responsible for DNA sensing.[78]

A second allosteric pathway regulates the dependence of the filament state on the identity of the bound nucleotide. Because ATP and ADP differ by the presence of a γ-phosphate, it is likely that this moiety is central to inducing the inter-protomer rotation necessary to achieve the active ATP-bound filament state. Several allosteric mechanisms have been proposed for how γ-phosphate sensing is transmitted to large-scale conformational change based on structural and mutagenesis data. In the reported active RAD51/RecA filament structures, non-hydrolysable ATP analogues bind between the two "half-sites" of two adjacent ATPase domains, where one half-site is composed of the phosphate binding P-loop and the surrounding nucleoside-binding residues. The other half-site is provided by a different set of residues from an adjacent protomer. Specifically, the nucleotide molecule is covered by a β-hairpin structure spanning residues Tyr315-Glu322 in *Hs*RAD51, also called the "ATP cap", (**Figure 7**). Pro321 from the ATP cap stacks tightly against the adenine base, while the preceding Cys319 backbone carbonyl further stabilises the interaction with the base by hydrogen-bonding with the exocyclic adenine NH2. A Glu322 residue that follows Pro321 does not directly interact with the nucleotide, but instead forms a salt bridge with a P-loop Arg130 from the other ATPase domain, contributing an additional inter-protomer link.
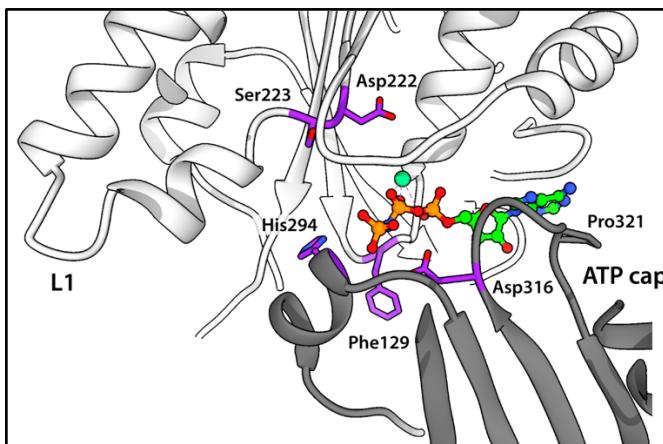
Figure 7. Two-way allosteric coupling of ATP hydrolysis and DNA binding. Residues considered to be important for allosteric signalling are highlighted in purple, AMP-PNP nucleotide is shown in green carbons, adjacent protomers are represent by light grey and dark grey ribbons. Residues Asp222-Ser223 have been proposed to act as a conformational sensors linking DNA binding to increased ATP hydrolysis rates through cis-trans peptide bond switching. Asp316 and Phe129 are putative γ-phosphate sensors that stabilise the high-pitch active NF.

It is likely that the ATP-cap contacts contribute to the stabilisation of the extended, active filament state, however, on their own, are insufficient to induce it, since identical contacts would be possible with an ADP cofactor. It is thus reasonable to suggest that there exists a γ-phosphate sensor that distinguishes between the cofactors. Asp316, located within the ATP cap β-hairpin, has been proposed to harbour such function by Amunugama and colleagues (**Figure 7**).[90] Its side chain is positioned adjacent to the γ and β-phosphates of the ATP-PNP molecules in NF structures and can interact with the gamma phosphate though a cation-mediated salt bridge, as evidenced by X-ray crystallography.[90] The Asp316 residue is conserved in eukaryotic and archaeal RAD51 orthologs, whereas in bacterial RecA and eukaryotic RAD51 paralogs it is replaced by a lysine that can directly bind the phosphate and does not require the intermediate cation.[90] The human RAD51 D316K mutant, which reverts this position to the bacterial lysine, has a markedly enhanced strand exchange activity and a decreased ATPase activity compared to the wild type protein, and exhibits a significantly higher NF stability as determined by SPR.[90] The authors propose an allosteric mechanism wherein this residue senses the ATP γ-phosphate, stabilising an ATP-capped conformation, with the concurrent inter-protomer rotation resulting in an increased filament pitch compared to ADP-bound or nucleotide-free filament. In a separate paper from the same group, the authors propose that the conserved lysine in RAD51 paralogs RAD51B-RAD51C confers increased stability to the nucleofilament.[44]

A conserved phenylalanine residue corresponding to Phe129 in human RAD51 (Phe140 in *Pf*RadA, Phe187 in *Sc*Rad51) has also been proposed to act in γ-phosphate sensing (**Figure**

7). It is located in the middle of the P-loop and the phenyl side chain forms a flexible cap over the P-loop phosphate site. In the *apo* form, lacking a nucleotide co-factor, and in ADP-bound states, all published atomic resolution structures reveal the side chain to be in a "closed" or "up" conformation, with the phenyl covering the γ-phosphate binding site. When bound to nucleotides with a γ-phosphate, i.e. ATP and AMP-PNP, the side chain has to undergo a rotamer flip to an "open" or "down" conformation that accommodates the additional phosphate (**Figure 7**). In the active filament structure, the open Phe129 interacts with Ala293, His294 and Arg299 from the adjacent protomer, possibly stabilising the inter-protomer orientation. The closed state of Phe129 would not be compatible with the active conformation of the NPF, as the phenyl side-chain would sterically clash with the Ala293 backbone carbonyl from the adjacent half-site. This would prevent the positioning near the catalytic site of His294, which has been shown by mutagenesis to be involved in the control of ATP hydrolysis.

Thus, it appears that a fine-tuned network of allosteric regulators exists within the RAD51 NF that maintain a sufficient level of ATP hydrolysis necessary for post-synaptic disassembly and recycling of RAD51 molecules, while ensuring that futile consumption of ATP is otherwise prevented, for example, in DNA-unbound RAD51. Hydrolysis-disabled mutants can catalyse strand invasion, but do not efficiently dissociate from the post-synaptic heteroduplex. It is likely that the hydrolytic dissociation of the γ-phosphate induces the conformational sensors of ATP, such as Asp316 and Phe129, to cease the stabilisation the extended inter-protomer orientation, and the filament reverts back to a lower-pitch, inactive state.

1.6 Regulation of RAD51 by BRCA2

In addition to the intrinsic allosteric pathways within the RAD51 nucleofilament, many regulatory proteins are involved in maintaining correct spatiotemporal control of RAD51 function.[91] BRCA2 is the most significant RAD51 protein mediator both mechanistically and from a clinical perspective, and will be discussed in the following section. Interestingly, there is no BRCA2 homolog in yeast, an important model organism in the study of DNA damage repair, and its functional homolog *Sc*Rad52 exerts similar mediator effects on *Sc*Rad51.

BRCA2 (breast cancer 2) is a human tumour suppressor gene that first gained significance as mutations were identified that are associated with a large increase in breast and ovarian cancer incidence. Women who inherit mutations in the BRCA2 gene have an 85% chance of developing breast cancer.[92] Homozygous BRCA2-/- homolog mutants are embryonic lethal in mice.[93] Its orthologs are found across the eukaryote domain and differ widely in size

and number of constituent functional domains, suggesting significant evolutionary flexibility in structure and function. Cells from a wide range of species that have been rendered deficient in BRCA2 are sensitised to DNA damaging agents and have impaired ability to repair DSBs.[94,95,96]

The 3418 amino acid protein product of BRCA2 is a regulator of homologous recombination and acts on molecules in this pathway, most crucial is its role as a mediator of RAD51, on which it exerts a number of functions.[97] In the absence of DNA damage, BRCA2 sequesters RAD51 in a monomeric form to prevent its oligomerisation and formation of RAD51 nucleofilament with off-target DNA substrates.[98] Once the DSB repair pathway is activated, BRCA2 orchestrates the assembly and disassembly of oligomeric RAD51 on DNA. Consistent with this, BRCA2 is required for the formation of nuclear RAD51 foci following exposure to ionising radiation.[99]

Purification and biochemical characterisation of full-length BRCA2 by Jensen *et al*.[100] elucidated many of the mechanistic steps this protein carries out to ensure efficient and tightly controlled homologous recombination and is to date the most comprehensive analysis of its *in vitro* function. BRCA2 contains several DNA binding domains – three oligonucleotide-binding (OB) folds and a helical domain, which interact with ssDNA, and a tower domain that binds dsDNA. The *in vitro* study showed that BRCA2 exhibits a strong preference for binding ssDNA and tailed ss/dsDNA substrate over dsDNA duplex, consistent with its role of targeting of recombination machinery to resected DSBs.[100] *In vitro* strand exchange is inhibited when an excess of RAD51 binds the dsDNA template in addition to the ssDNA substrate. Thus, the full-length BRCA2 protein was shown to stimulate RAD51-mediated strand exchange by enhancing active NF formation on ssDNA and decreasing inhibitory RAD51 binding to dsDNA.[100] Pre-incubation of ssDNA with RPA impairs *in vitro* strand exchange, because RAD51 cannot displace the high affinity ssDNA-RPA interaction. The authors showed that catalytic amounts of BRCA2 restore RAD51 binding to ssDNA by rendering it capable of RPA displacement, in the absence of a direct interaction between BRCA2 and RPA.[100] The same study reported that BRCA2 can stimulate strand exchange using both 3′ and 5′ tailed ssDNA/dsDNA substrates, as opposed to *E. coli* and *U. maydis* homologs, which act specifically on 3' overhangs.[100] Moreover, BRCA2 was found to inhibit ssDNA-dependent RAD51 ATPase activity.[100]

1.6.1 BRC repeats

Two distinct regions for binding RAD51 have been identified in BRCA2 (**Figure 8**). In the central part of the protein, encoded by the exon 11, are located a series of eight conserved ~30-40 residue long sequence fragments termed BRC repeats.[101] BRC repeats are variably spaced within the human protein – the distance between BRC1 and BRC2 is 210 residues, while that between BRC3 and BRC4 is 96. Corresponding repeats of same number in related organisms, e.g. BRC4 in mammals such as dog and mouse, are highly conserved, whereas sequence conservation between repeats of different number in the same ortholog is lower.[102] This suggests that the eight BRC repeats in mammals emerged as a result of sequence duplication before the radiation of the mammalian class.[102] The number of BRC repeats varies dramatically as one explores BRCA2 homologs of more distant eukaryotes. For example, the fungus *Ustilago maydis* has only a single BRC repeat in its BRCA2 ortholog, whereas the protozoan *Trypanosoma brucei* has 15 BRC repeats.[103,104] Step-wise removal of *T. brucei* BRC repeats has been shown by Trenaman *et al*. to perfectly correlate with the number of Rad51 nuclear foci observed after induced DNA damage.[105] It is therefore likely that BRC repeat number correlates with Rad51-mediator activity.

Isolated individual BRC repeats or short BRCA2 fragments containing one or more repeats have been shown to exert many of the effects on RAD51 that are ascribed for the full-length protein. A polypeptide construct encompassing a BRC3-BRC4 fragment fused to the BRCA2 DNA binding domain (DBD) was shown to be sufficient to stimulate RAD51-ssDNA nucleofilament formation and strand exchange, despite being only a fraction of the size of the full-length protein.[106]
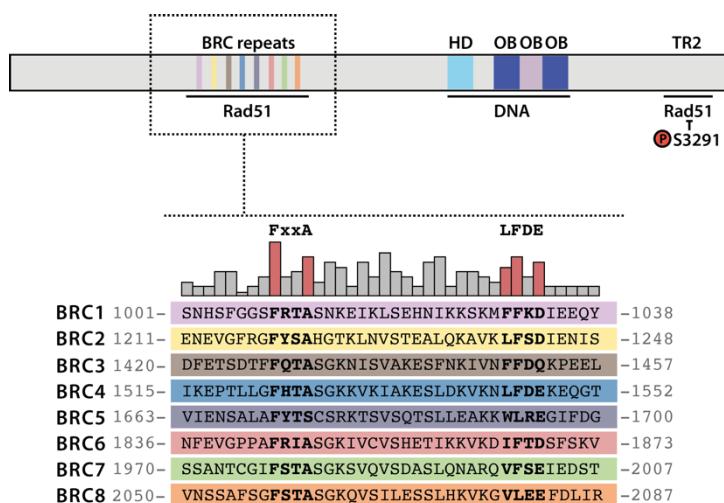
Figure 8. A schematic representation of the BRCA2 protein product with RAD51 and DNA-interacting domains highlighted (top). Alignment of the 8 human BRC repeats with conserved FxxA and LFDE motifs highlighted (bottom).

The exact mechanisms by which the BRC repeats modulate RAD51 function either as isolated peptides or in the context of the full-length BRCA2 protein are elusive. There appears to be functional differentiation between repeat groups 1-4 and 5-8, as shown in a study by Carreira and Kowalczykowski.[107] The study reports that individual BRC repeats 1-4 bind free RAD51 with higher affinity than repeats 5-8, stabilise the active form of RAD51-ssDNA nucleofilament by inhibiting its ATPase activity, and decrease RAD51 binding to dsDNA. On the other hand, repeats in second group, comprising BRC5-8, bind free Rad51 with low affinity but RAD51-ssDNA nucleofilament with high affinity, and do not affect ATPase activity or dsDNA binding. The authors propose that two groups BRC1-4 and BRC5-8 have evolved different functions to promote the efficient nucleation and growth of RAD51 ssDNA nucleofilament, respectively.

Remarkably, other studies have shown that, on the contrary, isolated BRC repeats prevent the assembly of RAD51 on ssDNA, rather than stabilise the filament, and are therefore inhibitors of homologous recombination when taken out of context of the full-length BRCA2 protein.[87] In fact, when overexpressed in cells, isolated BRC4 has been shown to disrupt RAD51 foci formation and sequester RAD51 in the nucleus.[108] BRC4 peptide also disrupts free RAD51 oligomers in vitro.[97] These effects are consistent with the structural observation that the BRC repeat binding would overlap with RAD51 oligomerisation epitope (**Section 1.6.2**, **Figure 9A,D**). To date, structural explanations of how these conflicting observations can be reconciled are lacking. These discrepancies may be addressed to some extent by the fact that some of the published *in vitro* experiments use GST-fused repeats, which are expected to form dimeric species in solution, and are therefore not strictly isolated repeats, whereas other studies use free linear peptides.

1.6.2 Structural basis of the RAD51-BRC repeat interaction

The first and so far only structural model of a BRC repeat in complex with RAD51 was determined by Pellegrini and colleagues using X-ray crystallography.[80] To obtain a crystallisable complex, the *Hs*RAD51 C-terminal ATPase domain, lacking the OE and the N-terminal domain, was recombinantly fused to BRC4, the highest-affinity BRC repeat, separated by a flexible linker, and the construct produced in a bacterial expression system. The resulting model elucidates the overall fold of the repeat when bound to the ATPase domain and suggests a number of critical structural features, or "hot-spots", that drive the interaction (**Figure 9A**). The most outstanding feature of the binding mode is the interface formed by the repeat residues 1524-FHTA-1527, interacting with the FxxA binding site on the RAD51 protomer (**Figure**

24

9B). These residues bind in an almost identical manner as the FxxA motif from the oligomerisation epitope (**Figure 9A,B**). Phe1524 constitutes the most extensive hydrophobic contact of the whole interface, while Ala1527 binds a much smaller nearby pocket. These FxxA motif residues are conserved in six out of eight human BRC repeats (**Figure 8**) and have been shown to be critical for binding RAD51.[109,81]

Besides the FxxA motif tetrad, the rest of the repeat binds differently compared to the OE (**Figure 9A**). The N-terminal residues Leu1521-Gly1523 do not make significant contacts with the ATPase domain and form a short, solvent-exposed α-helix, the significance of which is unclear. 1528-SGK-1530 residues that lie immediately C-terminal to the FxxA motif form a β-turn, stabilised by hydrogen bonds between backbone amides and side-chain hydroxyl groups of Thr1526 and Ser1528 (**Figure 9B**). This turn causes the repeat to bend back in the opposite direction, folding into a β-hairpin whose peptide backbone hydrogen bonds with the ATPase central β-sheet and extends it in an inter-molecular fashion. A conserved cationic residue corresponding to Lys1530 in BRC4 forms a salt bridge with RAD51 Asp187 (**Figure 9B**).

At its C-terminal half, spanning residues Lys1536 to Glu1548, the peptide folds into an α-helix that produces additional contacts with the ATPase domain through a combination of hydrophobic and polar interactions (**Figure 9C**). Residues Ile1534, Leu1539, Val1542, Leu1545 and Phe1546 form a continuous hydrophobic interface with RAD51 by projecting their side chains into the ATPase domain surface from a shared helical face. Of the five residues, Leu1545 and Phe1546 are significantly buried within cavities formed by RAD51. Further to the C-terminus, Glu1548 forms a salt-bridge with a nearby Arg250 and can potentially also interact with an adjacent Arg247. The combination of two hydrophobic residues at fixed positions within the repeats, represented by Leu1545 and Phe1546 in BRC4, with an acidic residue, represented by Glu1548, is a highly conserved feature found across all mammalian BRC repeats. Because BRC4 is the most studied and highest-affinity repeat, this feature has been termed the "LFDE" motif and has been shown to be critical for high-affinity binding in biochemical and cellular experiments,[81] and thus represents an additional interaction hot-spot.
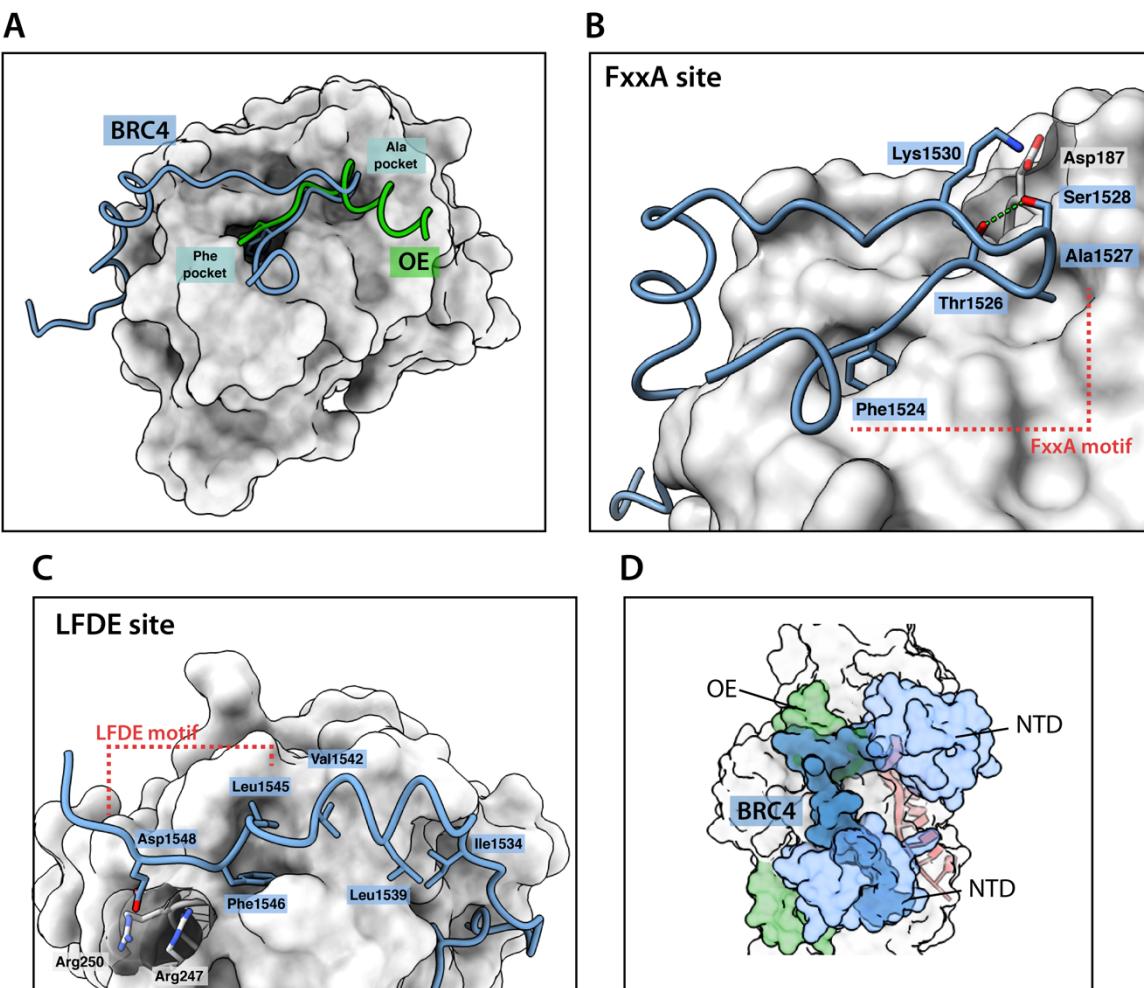
Figure 9. X-ray crystallographic structure of BRC4 repeat (blue) in complex with RAD51 ATPase domain (grey) (PDB: 1n0w). (**A**) Cartoon depiction of the BRC4 peptide superimposed with an oligomerisation epitope from *Pf*RadA (PDB: 1pzn). Near-identical backbone trace is seen at the FxxA site but not beyond. (**B**) Binding of the BRC4 repeat at the FxxA site of RAD51. Hot-spot residue Phe1524/Ala1527, and β-turn stabilising residue Thr1526/Ser1528 side-chains are rendered as sticks. (**C**) Binding of the BRC4 repeat at the LFDE site of RAD51. Hydrophobic interface-mediating residue side-chains and acidic Asp1548 are depicted as sticks. Salt bridge-mediating RAD51 arginines are shown behind a semi-transparent surface. (**D**) Superposition of the BRC4 repast structure with a cryo-EM model of a pre-synaptic nucleoprotein filament.

The BRC repeats have been reported to span a wide range of affinities for RAD51, with BRC4 being the tightest binder.[107,110] As evidenced by the BRC4 crystal structure, there is direct competition between the binding of BRC repeats and the RAD51 OE (**Figure 9A,D**). The OE has been shown to bind RAD51 with 800 nM affinity, which is more than a magnitude lower than BRC4.[111] It is therefore remarkable that BRCA2 serves a RAD51 mediator function and stimulates its assembly on ssDNA, whereas isolated BRC peptides disrupt it.

1.6.3 TR2 region

The C-terminal TR2 region (sometimes called CTRB) of BRCA2, encoded by exon 27, also binds RAD51 and is important for HR.[112] The binding preferences of TR2 were reported in two simultaneous publications by Davies and Pellegrini[87] and Esashi *et al*.[113] TR2 binds RAD51 oligomers both in the absence of DNA and in the context of a Rad51-ssDNA nucleofilament. It protects the RAD51-ssDNA nucleofilament from depolymerisation by the isolated BRC4 repeat in an ATP hydrolysis-independent manner and inhibits the formation of 1:1 RAD51-BRC4 complexes. Such effect is not observed on the depolymerisation of free RAD51 oligomers or RAD51-dsDNA nucleofilament. Phosphorylation of TR2 on Ser3291 by cyclin-dependent kinases (CDKs) disrupts this interaction in a cell cycle-dependent fashion. This post-translational modification ablates RAD51 binding at the TR2 and serves a molecular switch for turning off homologous recombination during the G2-M transition.[114]

1.7 Shuffling of BRC repeats reveals modular contributions to binding

Integration of previously published structural and affinity data indicates that binding of the conserved BRC repeat FxxA and LFDE motifs is contingent on the folding of two distinct secondary structure elements, a β-hairpin and an α-helix, respectively, that split the bound repeat into two structural modules (**Figure 9A**). For convenience, these modules are termed "FxxA" and "LFDE", reflecting the hotspot motifs that they contain. Previously, Rajendra *et al*.[81] demonstrated that an FxxA module from BRC4 and an LFDE module from BRC5 can be recombined into a new repeat, termed BRC4-5, in manner that preserves binding to RAD51, as evidenced by ELISA and streptavidin bead co-precipitation.

The modular nature of BRC repeats prompted a collaborative project led by the group of Prof Florian Hollfelder (Department of Biochemistry, University of Cambridge), and our group, with the aim of investigating BRC repeat modularity. The findings uncovered in the project directly informed the aims of the work presented in this thesis. A pre-print describing this initial work has been published,[115] and also includes data from **Chapter 2** *Characterisation of a chimeric BRC8-2 repeat.*

Understanding modular contribution to binding would provide insight into BRC repeat function. For example, it may be possible that the two repeat groups 1-4 and 5-8, that have been shown to manifest distinct biochemical effects on RAD51 nucleofilament,[81,116] arise as a result of modular differentiation. Quantitative binding measurements of individual modules are

challenging given their low intrinsic affinities.[81,117] Instead, FxxA and LFDE modules from each of the 8 parental human repeats were shuffled in every possible combination at a crossover point immediately C-terminal to the β-hairpin element, corresponding to the peptide bond between residues Lys1530 and Lys1531 in BRC4 (**Figure 10B**), leading to a total of 56 novel sequences that are not found in nature (**Figure 10C**).

A

	FxxA module	LFDE module
BRC1	SNHSFGGS**FRTA**SNK	EIKLSEHNIKKSKM**FFKD**IEEQY
BRC2	ENEVGFRG**FYSA**HGT	KLNVSTEALQKAVK**LFSD**IENIS
BRC3	DFETSDTF**FQTA**SGK	NISVAKESFNKIVN**FFDQ**KPEELH
BRC4	KIKEPTLLG**FHTA**SGK	KVKIAKESLDKVKN**LFDE**KEQGT
BRC5	SVIENSALA**FYTS**CSR	KTSVSQTSLLEAKK**WLRE**GIFDG
BRC6	NFEVGPPA**FRIA**SGK	IVCVSHETIKKVKD**IFTD**SFSKV
BRC7	SSANTCGI**FSTA**SGK	SVQVSDASLQNARQ**VFSE**IED
BRC8	VNSSAFSG**FSTA**SGK	QVSILESSLHKVKG**VLEE**FDLIRT

C 8^2 = 64 chimaeric repeats

E K_d (nM), FxxA module (columns 1–8), LFDE module (rows 1–8):

	1	2	3	4	5	6	7	8
1	151	445	105	32	4175	795	274	19
2	92	278	10	13	3081	175	35	6
3	2834	4051	245	899	8643	9506	2973	138
4	187	205	15	38	3354	1497	262	17
5	2210	3306	358	2080	1636	2735	3825	249
6	300	861	1826	1422	8334	9780	446	3176
7	1227	3097	561	209	1962	2028	1151	534
8	4176	7947	1109	861	2661	2407	11110	573

F $\Delta\Delta G_{parental}$ (kCal/mol), with $\Delta\Delta G_{LFDE1-8}$ (last column) and $\Delta\Delta G_{FxxA1-8}$ (bottom row):

	1	2	3	4	5	6	7	8	$\Delta\Delta G_{LFDE1-8}$
1	n/a	0.45	-0.35	-0.50	1.24	-0.25	-0.24	-1.60	-0.18
2	-0.47	n/a	-1.90	-1.20	0.88	-1.31	-1.62	-2.44	-1.15
3	1.57	1.60	n/a	1.30	1.52	1.06	1.00	-0.58	1.07
4	0.53	0.40	-1.08	n/a	1.51	0.52	0.13	-1.26	0.11
5	0.87	0.93	-0.33	1.24	n/a	-0.22	0.60	-0.79	0.33
6	-0.81	-0.38	0.10	0.49	0.43	n/a	-1.18	0.17	-0.17
7	0.63	0.99	0.28	0.00	0.21	-0.29	n/a	-0.24	0.22
8	1.55	1.74	0.88	1.03	0.59	0.01	1.52	n/a	1.05
$\Delta\Delta G_{FxxA1-8}$	0.55	0.82	-0.35	0.34	0.91	-0.07	0.03	-0.96	

Figure 10. Shuffling of the FxxA and LFDE modules from 8 parental BRC repeats leads to binders with improved affinity. (**A**) Alignment of the 8 parental BRC repeats that were shuffled. (**B**) BRC4:RAD51 complex showing the cross-over point at which the FxxA/LFDE modules were delineated for shuffling (**C**) Depiction of the tested repeats, both shuffled and parental. (**D**) Diagram showing the fluorescence polarisation assay reaction used for affinity determination. (**E**) Affinity matrix of the shuffled and parental repeats expressed as K_D. (**F**) Calculated $\Delta\Delta G_{parental}$ values for individual shuffled repeats and their per-module averages shown at the bottom of each column (FxxA) or on the left of each row (LFDE).

The resulting peptides and the eight parental repeats were cloned into a bacterial expression plasmid as fusions to a GB1 protein tag to aid expression and solubility. All repeat fusions were successfully expressed and purified in sufficient quantities for binding assays. Their affinities for a monomeric RAD51 ATPase domain surrogate protein, HumRadA22, were successfully determined using a microfluidic assay setup, described previously by Gielen *et al.*,[118] that couples nanolitre droplet formation with a fluorescence polarisation (FP) competition assay, allowing rapid collection of a very large number of data points across

concentration gradients of the competitor GB1-BRC fusion. The FP assay used fluorescein-tagged BRC4 peptide as a fluorescent probe that is progressively displaced by the competitor (**Figure 10D**).

The parental repeat affinities determined in these experiments were in line with previously published K_D values and broadly maintained a similar relative ranking (4>1>3>2>8>7>5>6) of repeats to what has been reported.[107,110] Importantly, repeats 1, 2, 3 and 4 displayed higher affinity (median K_D = 245 nM) than repeats 5, 6, 7 and 8 (median K_D = 1636 nM), confirming the previously reported differentiation between the two groups.

Systematic analysis of the chimeric repeat affinities revealed a number of trends that would not be discernible in the context of parental repeats (**Figure 10E**). All of the chimeras containing FxxA[BRC5] are relatively weak binders, with K_D values above 2 μM, consistent with observations by Rajendra *et al.*,[81] and likely resulting from a disruptive Ala to Ser substitution at the FxxA motif. Remarkably, the 56 new chimeras span an even wider range of affinities than the 8 parental repeats, from 6 nM to 11 μM, and a number of combinations bind more tightly than BRC4. Intriguingly, several of the shuffled repeats containing FxxA[BRC8], which originates in the "weak" 5-8 group of parental repeats, were amongst the tightest-binding sequences overall. BRC8-2 was the highest-affinity repeat in the whole matrix, with a K_D of 6 nM, exceeding both of its parental repeats BRC8 and BRC2 more than 40-fold.

To enable analysis of module-specific contributions, affinities were first derived in terms of Gibbs free energy ($\Delta G°$, **Eq. 1**) and then, for each chimeric repeat, the effect of shuffling was expressed as the average of the two differences of $\Delta G°$ between the chimeric repeat and its parents ($\Delta\Delta G_{parental}$, **Eq. 2**).

$$\Delta G° = R \cdot T \cdot \ln K_D \textbf{ (Eq. 1)}$$

$$\Delta\Delta G_{parental} = ((\Delta G°_{parent1} - \Delta G°) + (\Delta G°_{parent2} - \Delta G°))/2 \textbf{ (Eq. 2)}$$

Where: R = universal gas constant (8.314 J·K^{-1}·mol^{-1}), T = temperature (K)

A negative value for $\Delta\Delta G_{parental}$ means that the net effect of recombining the two constituent modules is beneficial for binding, whereas a positive value indicates that on average the chimeric repeat is worse than its parents. One can then calculate the mean $\Delta\Delta G_{parental}$ values for a given module across the series, that is, an average of all values in a column for an FxxA module or in a row for an LFDE module of the recombination matrix (**Figure 10F**). For example, $\Delta\Delta G_{FxxA-BRC5}$ equals 0.91 kCal/mol, indicating that the FxxA module from BRC5 is most detrimental to binding when recombined with other LFDE modules. $\Delta\Delta G_{FxxA-BRC8}$, on the

other hand, equals -0.96, meaning that $FxxA^{BRC8}$ is the most favourable FxxA module for binding. Surprisingly, the LFDE module from the same parental BRC8 is generally detrimental to binding ($\Delta\Delta G_{LFDE-BRC8}$ = 1.05 kCal/mol), which clearly demonstrates that parental repeat affinities are not indicative of both modules' contributions to binding. For LFDE modules, $LFDE^{BRC2}$ has the highest contribution to affinity with an $\Delta\Delta G_{LFDE-BRC2}$ of -1.15. Together the top-contributing modules, $FxxA^{BRC8}$ and $LFDE^{BRC2}$, recombine to form the highest-affinity chimeric repeat, BRC8-2. The K_D of this repeat is 6 nM, representing a more than 6 times tighter binding than BRC4.

This analysis shows that both LFDE and FxxA modules make meaningful contributions to the interaction and greatly expands our understanding of module-specific effects that are obscured in the context of the parental repeats. However, the structural basis of these contributions still remains poorly understood. Besides the importance of the two conserved tetrad motifs, it is hard to predict what other structural factors are at play based on the sequence data and the BRC4 complex structure. The high affinity BRC8-2 repeat binds more strongly than any of the other repeats. This property makes it interesting for the development of pharmacological probes targeting RAD51. Understanding the structural features that drive the high-affinity interaction can inform the design of pharmacological probes, and the BRC 8-2 sequence itself may provide an attractive template for the development of binders.

1.8 RAD51 in protozoan parasites

Protozoans are an informal phylogenetic category that describes a highly diverse group of unicellular eukaryotes characterised by having independent motility and heterotrophic feeding.[119] A number of protozoans are important human parasites and cause severe disease, particularly in tropical and subtropical geographic regions.[120] Most notable examples are *Plasmodium*, *Trypanosoma*, *Leishmania* and *Toxoplasma* species. Despite being evolutionarily distant from metazoans such as humans, many protozoan parasites have Rad51 and BRCA2 homolog genes identified. Moreover, in a number of species, the two proteins have been functionally characterised and shown to contribute to specific aspects of the parasite life cycle and pathogenicity. A number of these roles are discussed below.

Trypanosoma brucei, the causative agent of African sleeping sickness, is capable of host immune evasion through a process termed variable surface glycoprotein (VSG) switching, where the expression of alternate coat glycoproteins on the parasites' surface is achieved by recombination of silent VSG genes into an actively expressed locus. McCulloch and Barry have shown that the recombination mechanism that underlies VSG switching is mediated by

Rad51.[121] Interestingly, the *T. brucei* BRCA2 homolog has 15 BRC repeats of which 14 are identical in sequence, suggesting a relatively recent expansion in repeat number. A study by Hartley and McCulloch demonstrated that reducing the number of these repeat copies results in impaired VSG switching rates.[104]

In *T. cruzi*, Rad51 has a protective effect on host-induced oxidative stress and promotes parasitaemia.[122] Moreover, Rad51 is important for the process of genetic exchange in the *T. cruzi* parasite,[123] which lacks sexual mechanisms of recombining genetic information and therefore relies on the formation of fused cell hybrids.

Leishmania species and other kinetoplastids are unique among eukaryotic organisms in that they do not regulate gene expression at transcriptional level, that is, through promoters and transcription factors, and all genes are transcribed as long polycistrons of 10-100 genes with similar expression levels. One of the ways in which *Leishmania* responds to changing environment is by altering the copy number of specific genes. This can manifest as aneuploidy or amplification of smaller DNA segments. Genes can be multiplied within the chromosome or as extra-chromosomal amplicons. Laffitte *et al*. have shown that such amplicons underlie several antibiotic resistance mechanisms in *Leishmania* species.[124] Formation of these extra-chromosomal elements is facilitated by homologous recombination and therefore Rad51 is important for the resistance mechanisms.

The specific roles of Rad51 in protozoan parasites make it an attractive potential therapeutic target. Selective targeting of parasitic Rad51 orthologs may provide novel therapeutic avenues for treatment of neglected diseases. Understanding of the structural determinants of the Rad51:BRCA2 axis can guide the development of such interventions. BRCA2 orthologs in protozoan parasites vary significantly in size and domain composition, and contain variable numbers of BRC repeats, ranging from two in *L. infantum* to fifteen in *T. brucei*. To date, no structural studies of protozoan Rad51 and BRCA2 have been reported in the public domain, and only qualitative investigations of these interactions have been published.

1.9 RAD51 as a therapeutic target

Radiation therapy (RTx) is a key method in cancer treatment. It has been estimated that around 50% of oncology patients receive some form of radiation as part of their treatment.[125] RTx acts primarily by causing DNA damage in rapidly dividing cancer cells.[125] Moreover, RTx can prime the immune system to destroy metastatic secondary tumours via a variety of processes, such as increased presentation of tumour antigens.[126] Despite its wide application,

tumour relapse is extremely common due to a variety of RTx resistance mechanisms. RAD51 and other recombinogenic factors have been shown to be over-active in RTx-resistant tumours, and help such cells survive the genotoxic lesions.[127] RAD51 knock-down has been shown to sensitise resistant tumour cells to RTx.[128] Besides its application to RTx, RAD51 inhibition can synergise with other cancer-specific proteins, causing synthetic lethality in a variety of tumours. For example, activation-induced cytidine deaminase (AID) is an enzyme that is normally expressed in activated B-cells, causing mismatch mutations that are beneficial for antibody maturation in lymph nodes, and is tightly controlled to avoid unwanted DNA damage.[129] Overexpression of AID is observed in many lymphoid and solid cancers, leading to genomic instability that contributes to tumorigenesis. RAD51 inhibition with a small molecule compound has been shown to be effective against AID-positive tumours in mouse xenograft models.[130] Because several DSB repair pathways are available to healthy cells, specific RAD51 inhibition may achieve a sufficient therapeutic window in tumours where synthetic lethality is possible.

Name	Structure	Biochemical mechanism	Potency	Notes	Ref
DIDS		Strand exchange inhibition via DNA binding competition. DNA mimic.	SPR: K_D = 2 μM	Toxic to cells. Chloride channel inhibitor.	[131]
B02		Strand exchange inhibition via unknown mechanism.	Strand exchange: IC_{50} = 27 μM	No biophysical data.	[132]
CYT-0851	n.a.	n.a.	AID+ cell killing: EC_{50} = 2 μM	No biophysical data.	
RI-1		DNA binding inhibition. Covalent binding to Cys319 on protomer-protomer interface.	DNA binding FP: IC_{50} = 5-30 μM Cell sensitisation to MMS: EC_{50} =15-25 μM	Covalent inhibitor. Michael acceptor.	[133]
RI-2		DNA binding inhibition. Putative binding to protomer-protomer interface.	DNA binding FP: IC_{50} = 44 μM Cell sensitisation to MMS: positive	Derived from **RI-1** but not covalent inhibitor.	[134]
IBR2		RAD51:OE/BRCA2 interface binding (putative).	BRC4 competition: IC_{50} = 0.11 μM	Not reproducible in our lab.	[135]
CAM833		RAD51:OE/BRCA2 interface binding. BRCA2 binding competition.	ITC: K_D = 366 nM FP: K_D = 355 nM RAD51 foci inhibition in cells: IC_{50} = 5 μM	Crystal structure (PDB: 6TW9).	[117]

Table 1. Small-molecule pharmacological modulators of RAD51 function reported in literature. Compounds are depicted in a neutral ionisation state. AID+ denotes cells overexpressing activation-induced cytidine deaminase.

With potential applications for RTx and precision oncology, RAD51 holds great promise as a novel therapeutic target for cancer treatment. Despite this, to date there is only one clinical

trial of a RAD51 inhibitor reported, a Phase I/II study of a small-molecule candidate **CYT-0851** developed by Cyteir Therapeutics, with the chemical series originally described in patent US10336746B1. No mechanism of action (MOA) or biochemical data have been reported for this compound. The scarcity of clinical investigations and pharmacological probes for RAD51 is likely reflective of the challenges associated with targeting this protein, which presents challenging potential binding sites for inhibition. Targeting the ATP site may seem like a reasonable mechanism of action at first glance, yet its relatively flat and polar nature, as well as the need to compete with endogenous nucleotides, renders it sub-optimal. An overview of published pharmacological probes is provided in **Table 1**. One of the first RAD51 inhibitors reported, a DNA mimic **DIDS** is a chloride channel blocker that has been shown to competitively inhibit RAD51 DNA binding and downstream strand exchange.[131] However, the low potency, high negative charge and potential lack of specificity would severely limit its pharmacological application. The RAD51:OE/BRCA2 binding interface is a protein-protein interaction that has instead gained attention as a druggable site for inhibition. Recently, a potent small-molecule RAD51 inhibitor (**CAM833**) disrupting the RAD51:OE/BRCA2 interface has been reported, with a peptidomimetic structure that binds at the FxxA site with a K_D of 366 nM and inhibits IR-induced RAD51 foci formation at an IC_{50} of 5 μM in the A549 cell line.[117] **CAM833** is therefore the first example of a biophysically, structurally and cellularly validated probe inhibiting RAD51 function.

The DNA repair research community would benefit from novel, more potent RAD51 inhibitors with validated MOAs. Moreover, such molecules may lead to the development of effective drugs for cancer treatment. Targeting protein-protein interactions (PPIs), such as the RAD51:OE/BRCA2 interface, is a challenging task due to the physicochemical nature of these interaction. The following chapter describes the issues surrounding PPI targeting, as well as a number of potential solutions.

1.10 Pharmacology of protein-protein interactions

Protein-protein interactions (PPIs) are specific physical contacts between proteins that exert a functional role in the organism. This is opposed to random encounters resulting from macromolecular crowding and non-specific interactions. PPIs are essential to almost all of the physiological processes in the cell, and their analysis is crucial for understanding normal organismal function and elucidating disease aetiology. Despite their critical role in the functioning of the cellular biochemical machinery, PPIs are under-represented as a target class

in the development of pharmacological probes in target-based drug discovery.[136] Traditional medicinal chemistry as we know it today experienced dramatic advances in the second half of the 20th century, enabled by the myriad of novel organic synthetic reactions, recombinant DNA technology and an increase in assay throughput. This mode of drug discovery focused on modulating a limited number of privileged target classes, mainly enzymes, nuclear receptors, ion channels and GPCRs, that had native ligands, such as biochemical substrates, with molecular weights comparable to small-molecule compounds, as opposed to those targets that interact with large macromolecules.[137] For example, many enzymes have active sites that bind small-molecule-like substrates at a well-defined cavity, containing a mixture of hydrophobic and polar features, which provides the drug designer with a convenient "druggable" orthosteric site for a medicinal chemistry work. PPIs, on the other hand, tend to exhibit substantially different physicochemical profiles. Interaction interfaces between proteins tend to be mediated by more shallow binding cavities and encompass much larger areas of the protein's surface.[136] This renders the development of pharmacological probes challenging, as it becomes increasingly difficult to achieve a sufficient number of favourable contacts between the protein and the synthetic compound. This limited druggability problem is particularly acute for intracellular PPIs that are inaccessible for alternative modalities of larger size, such as antibodies and other biologicals.
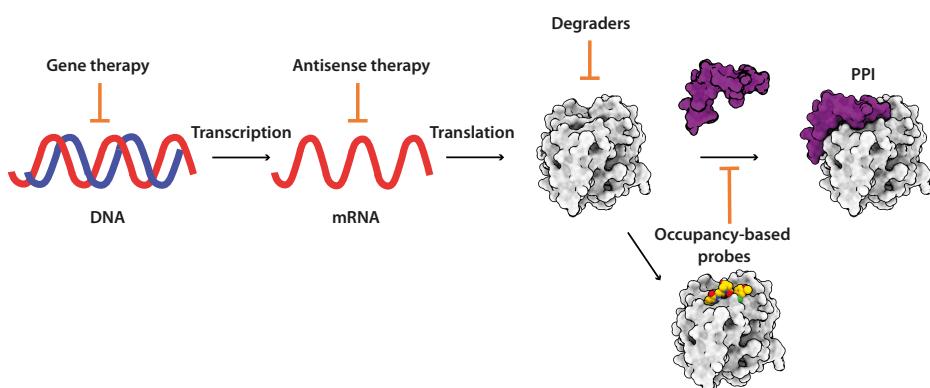
Figure 11. Diagram showing some of the different therapeutic modalities that can modulate protein-protein interactions at various stages of a protein's life cycle. Grey and purple surfaces represent two generic protein binding partners. Yellow spheres represent a small molecule inhibitor of the PPI.

The need to expand the druggable target space beyond the privileged target families has been recognised by the biomedical community, and has spurred several international collaborative initiatives, such as Illuminating the Druggable Genome and Target 2035, that aim to harness state of the art advances in biology and chemistry to find pharmacological modulators of hitherto undrugged proteins.[138] Therapeutic modalities that aim to modulate protein function can be mechanistically classified as either having occupancy-based

pharmacology or exert their activity by changing the protein level or identity (**Figure 11**). Modalities that alter protein level or its identity, such as antisense oligonucleotides, protein degraders and gene therapy, can potentially circumvent the druggability problem completely, because high-affinity binding to a druggable functional site on the protein is not necessary. Such drugs are gaining ever-increasing prominence and are now mature fields of their own, yet have specific limitations that hinder widespread application. Traditional occupancy-based modalities bind the target and alter its function by virtue of direct physical contact. This is the core paradigm of most medicinal chemistry, and suffers from the classical druggability pitfalls, as a suitable functional site is required for targeting.

Early-stage small-molecule drug development is in essence a multi-dimensional optimisation problem where the scientist iteratively improves on a number of pharmacological properties of a molecule in order for the drug to be efficacious. Among these, physicochemical properties of a compound are often maintained within what is termed the "drug-like space", with particular constraints on molecular weight, hydrophobicity and hydrogen bond donor/acceptor count.[139] A compound that is too large and hydrophobic is likely to have poor solubility and membrane permeability, leading to low bioavailability, as well as potential off-target toxicity. There is a common (but not universal) trade-off between these properties and compound potency, that becomes particularly acute when targeting PPIs.

A number of chemical and biological approaches that aim to overcome druggability limitations of PPIs have been applied in drug discovery.[136] For example, fragment-based drug discovery (FBDD) is a methodological framework in medicinal chemistry that aims to sample ligand chemical space more efficiently during early-stage screening, leading to small-molecule binders with higher affinity than compounds discovered using more traditional high-throughput methods.[140] In FBDD, libraries of fragment molecules containing low-complexity compounds of molecular weight of less than 300 Da are screened against a protein target using sensitive biophysical techniques such as NMR, SPR or X-ray crystallography, that allow the detection of weak affinities for molecules of such size. By virtue of covering a larger fraction of theoretical chemical space at a given molecular weight, as well as reduced conformational degrees of freedom, identified fragment hits are likely to *a priori* form more efficient contacts with the target than a drug-like hit from a high-throughput screen, thus providing more optimal starting points for their subsequent synthetic derivatisation.

An alternative modality to small molecules for targeting PPIs are conformationally constrained macrocyclic peptides.

1.11 Peptide macrocyclisation

Peptides have a long history of being used as therapeutics, with insulin being the best known and most widely administered peptide drug to this day, which was prepared and applied therapeutically as early as the 1920s.[141] Traditional peptide therapeutics were typically isolated from mammalian sources and mimic endogenous peptide ligands in humans, and act by supplementing dysregulated pools of the endogenous molecule. More recently, spurred by the advances in biology and chemistry, fully synthetic peptides became the focus of many drug discovery projects, and the target space of peptide therapeutics expanded beyond proteins with well-established native peptide ligands.

Peptide binders tend to have higher molecular weights and occupy larger binding interfaces than small-molecule drugs, and can therefore be more appropriate for the targeting of PPIs. Moreover, naturally occurring epitopes found in PPIs can be used as templates for the development of peptidic binders. Despite these positive attributes, peptide development has been hampered by a number of pharmacological pitfalls that are inherent to this modality and are primarily concerned with their ADME (Absorption, Distribution, Metabolism, Excretion) properties. A crucial disadvantage of peptides is their susceptibility to proteolysis *in vivo* by the plethora of proteases found in different physiological compartments both inside and outside the cell. Proteolytic cleavage makes peptides unsuitable for oral administration, as the various non-specific gastro-intestinal proteases completely ablate their oral bioavailability. IV-administered peptides typically have very short plasma-stabilities due to proteolysis in plasma and rapid renal clearance. Moreover, as peptides tend to be large compounds with significant conformational flexibility and contain many polar functional groups, their application to intracellular targets is limited by their poor plasma membrane permeability.

Peptide macrocyclisation is a drug discovery strategy where additional covalent linkages are introduced into a linear peptide with the aim of improving their pharmacological properties.[142] Conformationally constrained peptides are much less likely to form a favourable fit with a protease active site, therefore macrocyclisation can profoundly enhance a peptide's stability. Moreover, the pre-organisation of the peptide may result in higher binding affinity to the target, stemming from a reduced entropic penalty of binding. Crucially, peptide macrocycles can display an improved ability to cross the cell membrane and thus engage intracellular targets due to their smaller hydrodynamic radius and reduced conformational degrees of freedom.
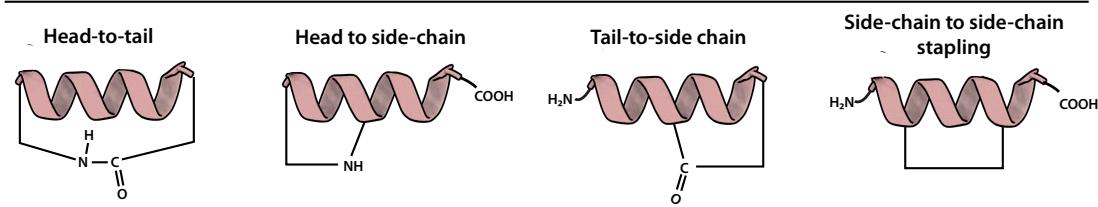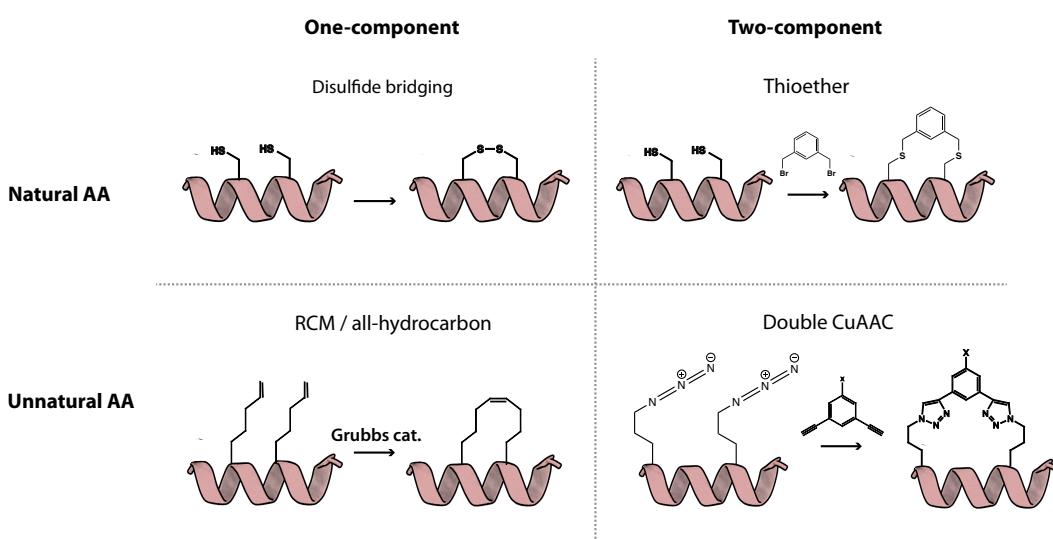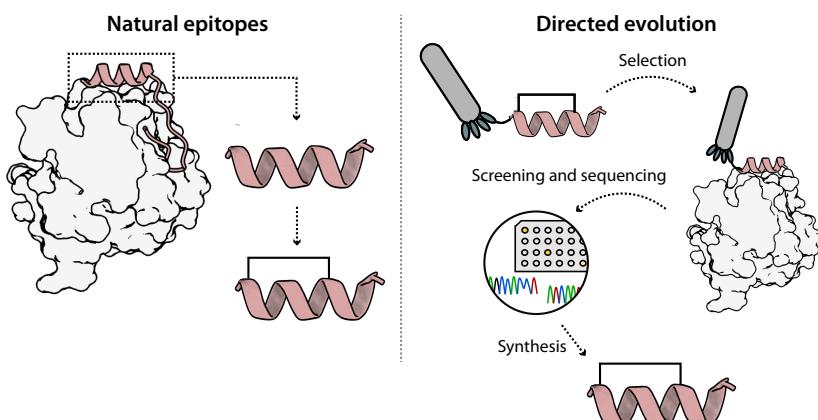
Figure 12. **Overview of peptide macrocyclisation methods.** Pink helix represents a generic peptide of interest. (**A**) Main geometries of covalent linkages. (**B**) Representative stapling reactions of various types classified according to the number of reaction components and the type of amino acid that is being stapled. (**C**) Binder sequences can be obtained from different sources, such as existing natural epitopes or directed evolution.

A variety of macrocyclisation strategies have been utilised in drug discovery and can be classified according to which parts of the peptide are linked (**Figure 12A**). Terminal amines and carboxylates may be covalently linked to form an amide bond in a "head-to-tail" approach.

Alternatively, amino acid side-chains can be linked with the termini or another side-chain. Moreover, macrocyclisation can be classified as either one-component or two-component (**Figure 12B**). One-component reactions contain the reactive groups within the amino-acid residues of the linear peptide, and the covalent bridge is introduced in an intra-molecular fashion. For example, ring-closing metathesis (RCM) reactions incorporate non-natural amino acids with olefin side-chains that react form the macrocycle via metal-catalysed metathesis. Because all the reactive parts constituting the covalent linkage are introduced in the peptide from the outset, RCM is a one-component stapling reaction, meaning that no external reactive linker is necessary.

In two-component reactions, a separate linker moiety is added that reacts with the two residues to join them (**Figure 12B**). For example, double copper-catalysed alkyne-azide cycloaddition (CuAAC) has been employed to cyclise peptidic inhibitors of the CK2α:CK2β kinase interaction.[143] The authors used a bis-alkyne linker to connect two azidoalanine residues introduced into a natural epitope binding the target protein. Two-component reactions can also be performed with trivalent linkers that join three residues simultaneously, yielding bicyclic peptides. For example, 1,3,5-Tris(bromomethyl)benzene (TBMB) and other similarly electrophilic linkers have been reacted with phage display libraries of bicyclic peptides containing genetically encoded cysteines that were then selected against targets of interest in a directed evolution approach.[144]

The nature of amino acids used for macrocyclisation has profound methodological implications for peptide development. Unnatural amino acids (UAAs) dramatically expand the chemical space and allow many different chemistries to be employed. The incorporation of UAAs into peptides is straightforward when done by solid-phase peptide synthesis (SPPS), as many Fmoc-protected amino acids are commercially available. However, for methods that require ribosomal synthesis, such as phage or mRNA display, UAAs present additional challenges, as native Aminoacyl-tRNA synthetases are not optimal for incorporating them into tRNA. Proteinogenic amino acids, on the other hand, permit only a handful of macrocyclisation reactions, but are compatible with a variety of recombinant expression systems. Recombinant methods can be advantageous in cases of long peptides, as synthesis becomes increasingly challenging, resulting in lower yields. Moreover, recombinant methods do not require the use of large volumes of hazardous solvents such as DMF. Cysteine in particular has found wide application for many bio-conjugation tasks, as its nucleophilic thiol side-chain can be reacted selectively with a variety of electrophilic groups. Many examples of cysteine-based macrocyclisation strategies have been reported.[145,144,146]

1.12 Peptide stapling

Peptide stapling is a macrocyclisation approach in which the peptide is covalently cross-linked through the side-chains of two separate and amino acid residues (**Figure 12B**).[147] Traditionally, the term "stapling" has been used to describe specifically the side-chain to side-chain macrocyclisation of α-helical epitopes. α-helices have been recognised to mediate a large fraction of PPIs, as shown by the analysis of complexes deposited in the PDB,[148] and their regular shape and structural stability render α-helices attractive motifs for stapling. Historically, RCM has been the reaction of choice for α-helical stapling, as the length of the two alkene side-chains can be optimised to join various inter-residue distances, here referred to as "stapling architectures".[149] Peptides were typically obtained by taking the α-helical motif out of the parental polypeptide and the reactive residues introduced on the face opposite of the interface to minimise clashes with the binding partner. For simplicity, in this dissertation, the term "stapling" is applied to encompass side-chain to side-chain cyclisation of all secondary structure elements, such as β-hairpins and loops, as these are becoming increasingly commonly utilised. For example, Wendt and colleagues recently reported a successful campaign to inhibit Wnt signalling by targeting β-catenin with a head-to-tail linked β-sheet mimetic that has an additional inter-strand cysteine staple introduced in the middle.[150]

1.13 Aims and scope

The high-affinity BRC repeats constitute an attractive template for developing peptidic RAD51 binders. Moreover, chimeric BRC repeat data suggests that improved binding can be achieved by domain shuffling or other sequence modifications. The key objective of this thesis is to expand our understanding of BRC repeat binding and integrate it with the latest advances in peptide stapling to obtain potent pharmacological probes that inhibit RAD51 function. To achieve this objective, the following specific aims were addressed:

I. To characterise of the interaction between BRC8-2 and RAD51

In order to ascertain the applicability of BRC8-2 for peptidic probe development, it is necessary to validate the preliminary data, which suggests that BRC8-2 is a more potent binder than its parental peptides, as well as BRC4. For this, I performed a biophysical evaluation of its binding to RAD51. To expand our understanding of the atomic determinants of the interaction, which is crucial in guiding subsequent design of molecular probes, a structural characterisation of the complex between BRC8-2 and RAD51 was done using X-ray crystallography. To further probe the finer atomic details of binding, I performed additional mutagenesis and biophysical experiments to support my structural analysis. The resulting data contribute to the present understanding of how BRCA2 exerts its RAD51-modulatory function. Moreover, the structure provided crucial information for the subsequent design of stapled peptides.

II. To develop a recombinant stapled peptide screening platform

Peptide stapling presents a promising avenue for inhibiting RAD51, and the high-affinity BRC8-2 repeat can be utilised as a template for this strategy. A robust approach for preparing and screening many stapled peptides with different architectures is necessary to obtain molecules with optimal characteristics. For this, I used a two-component cysteine-stapling methodology that employs heterobifunctional divinyl linkers. To rapidly evaluate the different stapling architectures that can be introduced into this peptide, I devised a recombinant platform for the small-scale preparation and screening of cysteine-stapled peptides.

III. To design, screen and characterise stapled BRC8-2 repeats

The crystal structure of the BRC8-2:RAD51 complex can inform the design of cysteine pairs for stapling. Bis-cysteine mutants of BRC8-2 were rationally designed in a structure-

guided manner in order to obtain RAD51 inhibitor peptides that are proteolytically stable and maintain high affinity binding. These were evaluated using the recombinant screening platform. Different stapling architectures were found to be compatible with high-affinity binding. Based on these findings, selected stapled peptides were prepared at higher scale and tested in a variety of biophysical and biochemical assays. Crystal structures of stapled BRC8-2 peptides were determined, revealing the atomic-level features of the staple moiety.

IV. Characterisation of the BRC:Rad51 interaction in *L. infantum* orthologs

The protozoan *Leishmania infantum* is the causative agent of leishmaniasis and selectively disrupting its RAD51:BRCA2 ortholog interaction may lead to novel pharmacological interventions to tackle this disease. Understanding the structural and biophysical basis of this interaction and its comparison with the human orthologs would be a useful starting point for drug design. Using biophysical and structural approaches, I probed the binding characteristics of *L. infantu*m BRC repeats. I determined novel structural features that define the binding of *Li*BRC1 to *Li*Rad51.

2 Characterisation of a chimeric BRC8-2 repeat

The microfluidic fluorescence polarisation measurements performed by Dr Laurens Lindenburg identified BRC8-2 as the highest affinity repeat, and its constituent modules as most potent contributors to binding. This high affinity binder can serve as a template for development of peptide-based pharmacological probes, as it provides an advantageous starting point for chemical and sequence optimisation. However, the high-throughput approach in which these peptides were screened is susceptible to systematic error, for example, if some of the BRC repeat sequences are consistently degraded by the proteases during their expression. In order to confirm these preliminary affinity values and to establish BRC8-2 as a functional Rad51 inhibitor, I prepared the peptide in highly pure, un-tagged form and evaluated it in a range of biophysical and biochemical assays.

2.1 BRC8-2 is a potent inhibitor of RAD51 function *in vitro*

Full-length human RAD51 forms polydisperse linear oligomers in solution, mediated by the inter-domain oligomerisation epitope (OE). Because both OE and BRC repeats directly compete for the same interface, biophysical measurements of BRC repeat binding to *Hs*RAD51 are complicated by the need to compete with RAD51 oligomerisation. A simplified, truncated construct, containing only the monomeric ATPase domain of RAD51, which is the *bona fide* binding partner of BRC repeats, would allow direct one-to-one affinity measurement. Previous attempts in our lab to isolate the ATPase domain of human RAD51, lacking the N-terminal domain and the oligomerisation epitope, failed due to the intrinsic instability of the truncated protein. Instead, I used a highly soluble thermostable surrogate protein HumRadA22, based on the archaeal RadA ortholog from *Pyrococcus furiosus*. This mutant was previously engineered in our lab and serves as a faithful mimetic of the human RAD51 ATPase domain, in which surface residues encompassing the FxxA and LFDE binding sites, as well as nearby regions have been mutated to the human sequence.[111]

I recombinantly prepared BRC8-2, its parental repeats BRC2, BRC8, and the natural repeat BRC4. Briefly, peptides were expressed in *E. coli* as fusions to a GB1 protein expression tag, purified by IMAC, after which the tag was proteolytically removed with TEV protease, yielding free peptides that were then purified to homogeneity by reversed-phase chromatography (RPC), ensuring removal of degradation products. All of the peptides were

prepared in milligram quantities and were >90% pure as confirmed by HPLC and MS (**Appendix 8.1** *RP-HPLC chromatograms and MS spectra of purified peptides*).

Isothermal titration calorimetry (ITC) was used to determine peptide binding affinities to HumRadA22 (**Figure 13A**). The ITC data is broadly consistent with the affinity ranking initially observed by microfluidic FP, but important differences were observed. BRC4, the most potent natural binder according to previous reports, bound with a K_D of 23 nM, which is consistent with the microfluidic FP value (38 nM) and other published affinities. The parental BRC8 repeat exhibited the lowest affinity in the set, with a K_D of 56 nM, which is a significantly smaller K_D value than one from microfluidic FP. Surprisingly, BRC2 appears to be a more potent binder than BRC4, with a K_D value of 14 nM, as opposed to 278 nM in the microfluidic FP screen. These significant discrepancies can be attributed to a number of factors. First, the presence of a GB1 expression tag, as used in the microfluidic FP, may interfere with binding. Despite including a long, flexible linker between the tag and the peptide, it is possible that the GB1 domain forms non-specific interactions with some of the peptides, imposing an energetic penalty on binding, or, alternatively, it can have a direct steric clash with HumRadA22. More importantly, no downstream purification steps were employed after IMAC in the microfluidic FP screen, meaning that full-length peptides had to be quantified using SDS-PAGE gel densitometry, possibly contributing to decreased accuracy in concentration. Gratifyingly, BRC8-2 was tightest binder as measured by ITC, with a K_D of 4 nM.
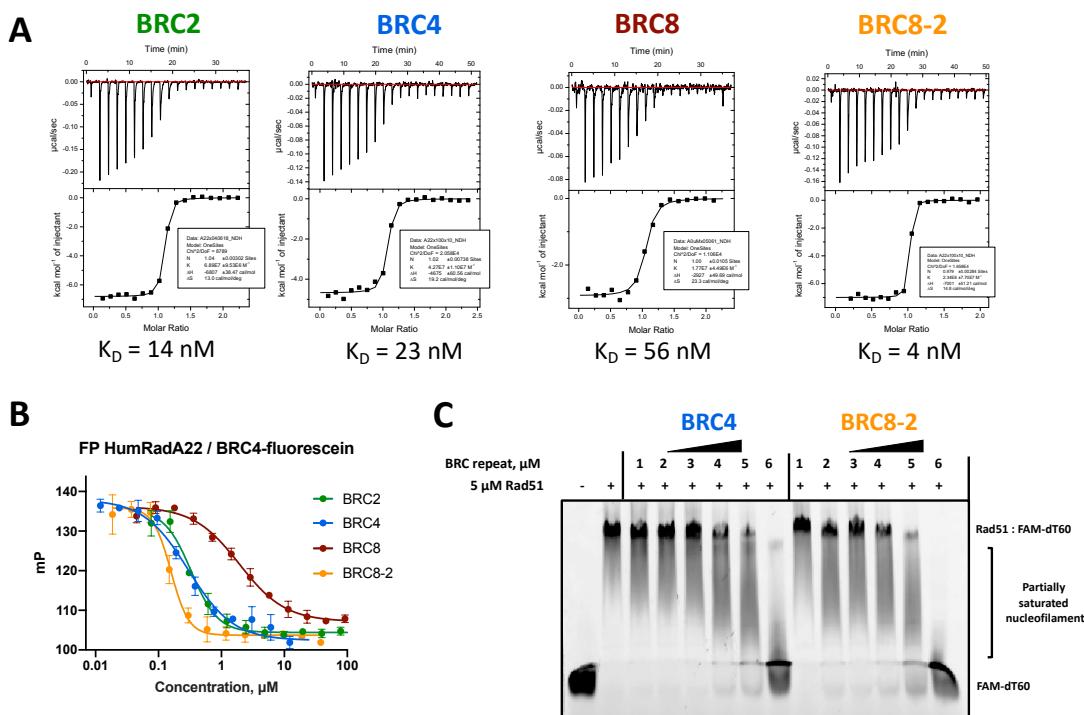
Figure 13. BRC8-2 is a potent RAD51 binder and a functional inhibitor. (**A**) Direct binding ITC titrations of BRC repeats with HumRadA22. In all titrations, 100 μM peptide was injected into cell containing 10 μM HumRadA22 over 19 injections. (**B**) Fluorescence polarisation (FP) competition assay titrations of BRC repeat peptides. Fluorescently labelled BRC4 (10 nM) probe was pre-incubated with 100 nM HumRadA22, to which peptide dilution series were added. Data shown are the means of triplicate measurements ± SD. (**C**) Electrophoretic mobility shift assay (EMSA) of *Hs*RAD51 (5 μM) with FAM-dT$_{60}$ (100 nM) and various concentrations of competitor peptides. BRC8-2 can efficiently disrupt *Hs*RAD51-ssDNA nucleofilament formation on the FAM-dT$_{60}$ oligonucleotide.

To further validate this affinity ranking, plate reader-based competition FP measurements were performed with the same peptides using a fluorescein-labelled BRC4 peptide as probe (**Figure 13B**). Here, a similar trend is observed, with BRC8-2 being the most potent inhibitor, and BRC8 the least potent, with BRC4 and BRC2 spanning intermediate values.

In a functional validation step, the same BRC repeat peptides were evaluated for their ability to inhibit full-length *Hs*RAD51 function. Electrophoretic mobility shift assays (EMSAs) were set up by adding fluorescently tagged dT$_{60}$ oligonucleotide to human RAD51 that has been pre-incubated with varying amounts of BRC repeat peptides. Reaction products were analysed on native PAGE. BRC8-2 was able to disrupt nucleofilament formation in a dose-dependent manner, with a complete decrease in the amount of shifted dT$_{60}$ at super-stoichiometric peptide concentrations relative to RAD51. At the saturating conditions necessary for the assay, it is impossible to quantitatively compare inhibition by different

peptides. However, the EMSA data clearly demonstrates that BRC8-2 potently inhibits *Hs*RAD51 function rather than just the archaeal surrogate HumRadA22.

Peptide	ITC K_D, nM	FP IC50 ± SE, nM	FP K_D, nM
BRC2	14	331 ± 31	27
BRC4	23	218 ± 32	14
BRC8	56	1931 ± 165	203
BRC8-2	4	158 ± 9	8

Table 2. Data summary for ITC and FP measurements of BRC8-2, its parental repeats BRC8 and BRC2, as well as BRC4. Affinity ranking broadly confirms the trends from chimeric peptide screening experiments.

2.2 BRC8-2 inhibits RAD51 foci in human cells

RAD51 localisation at the sites of DSB repair can be observed by immunofluorescence (IF) as nuclear foci that form following treatment with ionising radiation. Previously, overexpression of BRC4 in human cells has been shown to inhibit foci formation by inhibition of BRCA2 mediator activity and RAD51 self-oligomerisation.[108] Having validated the *in vitro* functional activity of BRC8-2, its inhibitory potential in cells was characterised. For this, I prepared a set of mammalian expression constructs containing the repeats BRC4 and BRC8-2 fused to N-terminal GFP and a nuclear localisation signal, expressed under the CMV promoter. The subsequent cell-based experiments were performed by Dr Pedro Zuazua-Villar at our collaborator Prof Jessica Downs' lab.

U2OS cells were transfected with either of the two BRC repeat constructs, or with an empty GFP-NLS control. RAD51 foci were counted by IF microscopy in GFP-positive cells after exposure to ionising radiation, or in the absence of radiation treatment (**Figure 14A**). A small number of foci were detected in GFP-NLS control cells that were not subjected to any radiation treatment, likely arising as a result of replicative stress (**Figure 14B,C**). Treatment of GFP-NLS control cells with a 3 Gy dose of IR resulted in a large increase in the number of RAD51 foci, consistent with the formation of many DSBs (**Figure 14B,C**). In contrast, cells transfected with either BRC8-2 or the positive control BRC4 had a greatly reduced number of RAD51 foci both prior and after radiation treatment (**Figure 14B,C**). Variability in protein expression levels and transfection efficiency make a quantitative comparison of BRC4 and BRC8-2 mediated inhibition challenging. Both peptides appear to cause more than a two-fold decrease in foci formation compared to GFP-NLS control in both treated and untreated cells, with no significant difference.

Because homologous recombination is limited to the S and G2 phases of the cell cycle, change in RAD51 foci count can also reflect a larger population of cells being in the G1 phase. To account for this, cell cycle profiles were examined by flow cytometric measurement of DNA content (**Figure 14D,E**). No significant change in G1 populations was observed for BRC4 and BRC8-2 transfected U2OS cells, meaning that the differences in foci counts are not cell-cycle related.

The IF experiments demonstrate that the chimeric repeat BRC8-2 is a potent inhibitor of Rad51 function in a cellular context, where it prevents translocation of RAD51 to sites of DNA damage. An increased pan-nuclear signal for RAD51 in cells transfected with BRC8-2 was observed compared to GFP-NLS control cells (**Figure 14F**), consistent with a previously proposed mechanism of action in which BRCA2 sequesters RAD51 in a diffuse nuclear state.[151] Acting as a competitor, BRC8-2 prevents the BRCA2-mediated recruitment of the recombinase by direct competition with its binding.
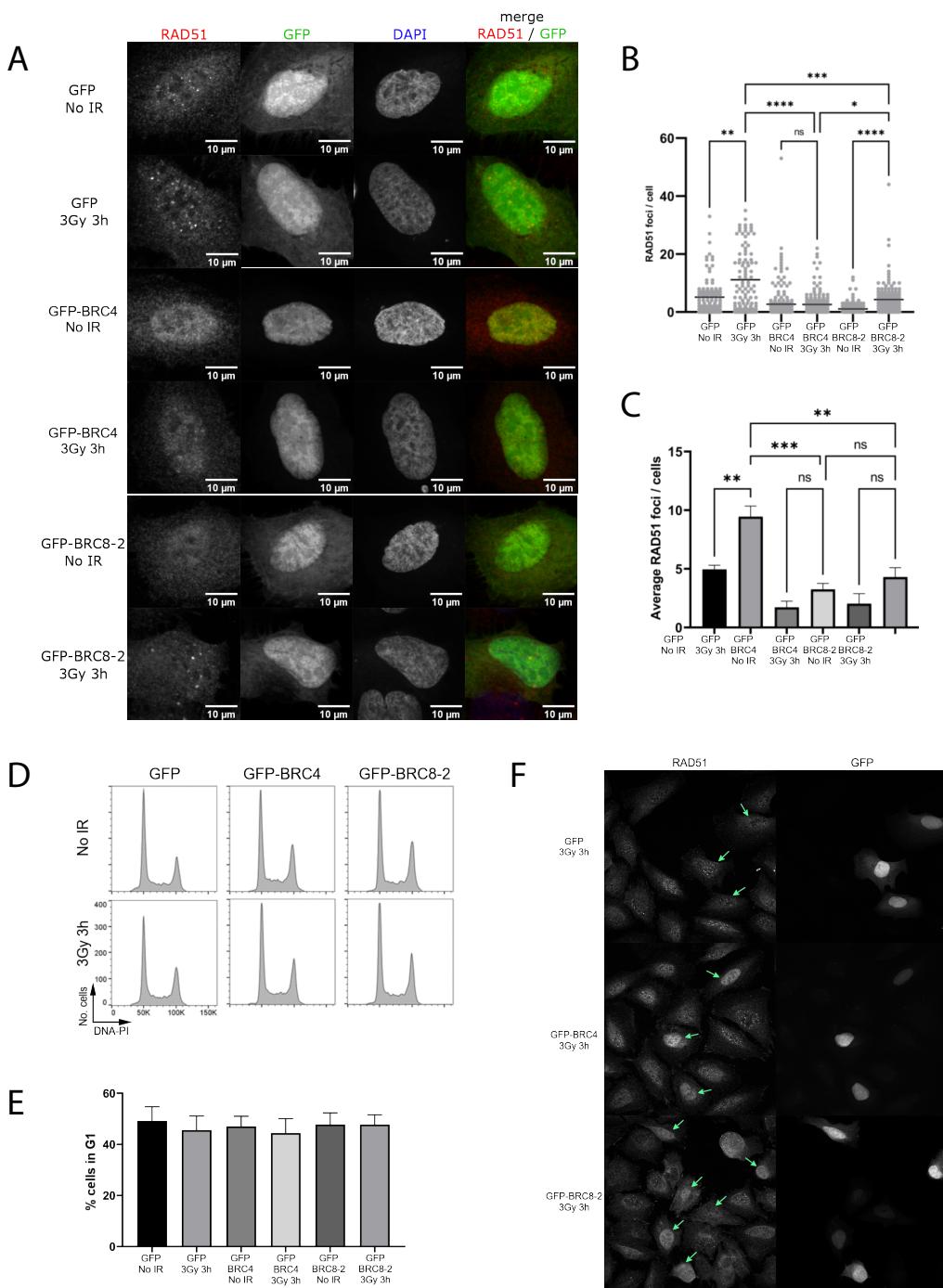
Figure 14. BRC8-2 impairs RAD51 foci formation in human U2OS cells. (**A**) Representative images of U2OS cells expressing GFP, GFP-NLS-BRC8-2, or GFP-NLS-BRC4. Cells were monitored 3 h after no treatment or irradiation with 3 Gy (3Gy 3h) for GFP fluorescence and stained with α-RAD51 and DAPI. (**B**) Dot plot graph from one biological replicate plotting the number of RAD51 foci per GFP positive cell. Median values are indicated with a bar. Statistical analysis was done using Kruskal Wallis rank sum test followed by Dunn's procedure for pairwise comparison (*p <0.005, **p<0.001, n.s.= not significant). (**C**) Bar graph showing the average median RAD51 foci per GFP positive cell from three independent biological experiments. Data are presented as mean ± SEM, n = 3 biological repeats. **p<0.001, n.s.= not significant. (**D**) Representative cell cycle profiles from GFP positive cells analysed by flow cytometry 3 hours after no treatment or irradiation with 3 Gy (3Gy 3h). (**E**) Bar graph showing the percentage of cells in G1 phase. Data are the mean values from three

independent experiments ± SEM. (**F**) The pan-nuclear signal of RAD51 in GFP-BRC8-2 and GFP-BRC4 expressing U2OS cells is greater than in the GFP control cells. Labels and conditions are the same as for (**A**). The arrows indicate GFP-positive cells. All experiments in this figure were conducted and the figure prepared by Dr Pedro Zuazua-Villar.

2.3 X-ray crystallographic study of BRC8-2

Having completed the functional survey of BRC8-2, I set out to investigate its binding mode with RAD51 and to elucidate the structural determinants of its high affinity interaction. For this, I decided to perform X-ray crystallographic studies of BRC8-2, as well as its parental repeats BRC2 and BRC8, in complex with RAD51.

For successful X-ray crystallographic studies of macromolecular complexes, it is of paramount importance to minimise assembly heterogeneity and reduce conformational degrees of freedom within the assembly. WT human RAD51 forms linear oligomeric assemblies of varying length, where a flexible oligomerisation epitope (OE) of one RAD51 protomer interacts with the ATPase domain of a successive molecule. The polydisperse nature of such system and competition between the OE and the BRC repeat would strongly disfavour crystallisation. The C-terminal ATPase domain contains the *bona fide* interaction interface between RAD51 and BRC repeats, as elucidated by Pellegrini *et al.*,[80] and should be on its own suitable for crystallographic studies. Unlike the approach taken by Pellegrini *et al.*, where a fusion between the *Hs*RAD51 ATPase domain and the repeat yielded a stable complex, I decided to express the binding partners as separate molecules. For this, the stable archaeal surrogate mimetic HumRadA22 was used again, as its previously determined apo structure closely matches the surface of WT *Hs*RAD51 ATPase domain (PDB: 5KDD).

Overexpression of small peptides in bacterial culture can be challenging, therefore BRC repeats were prepared as TEV-cleavable fusions to an N-terminal His_8-GB1 expression tag. The two partners of the complex were expressed separately in bacterial cells. The N-terminally His_8-GB1-tag BRC fusion-expressing cell lysate was loaded on a Ni-NTA column, followed by application of untagged HumRadA22 lysate, and subsequent elution (**Figure 15A**). The co-precipitation approach produced high yields of complex which were then subjected to TEV cleavage, tag removal by reverse IMAC and final purification by size-exclusion chromatography (**Figure 15A**). In total, HumRadA22 complexes were purified with all three peptides (BRC2, BRC8, BRC8-2).
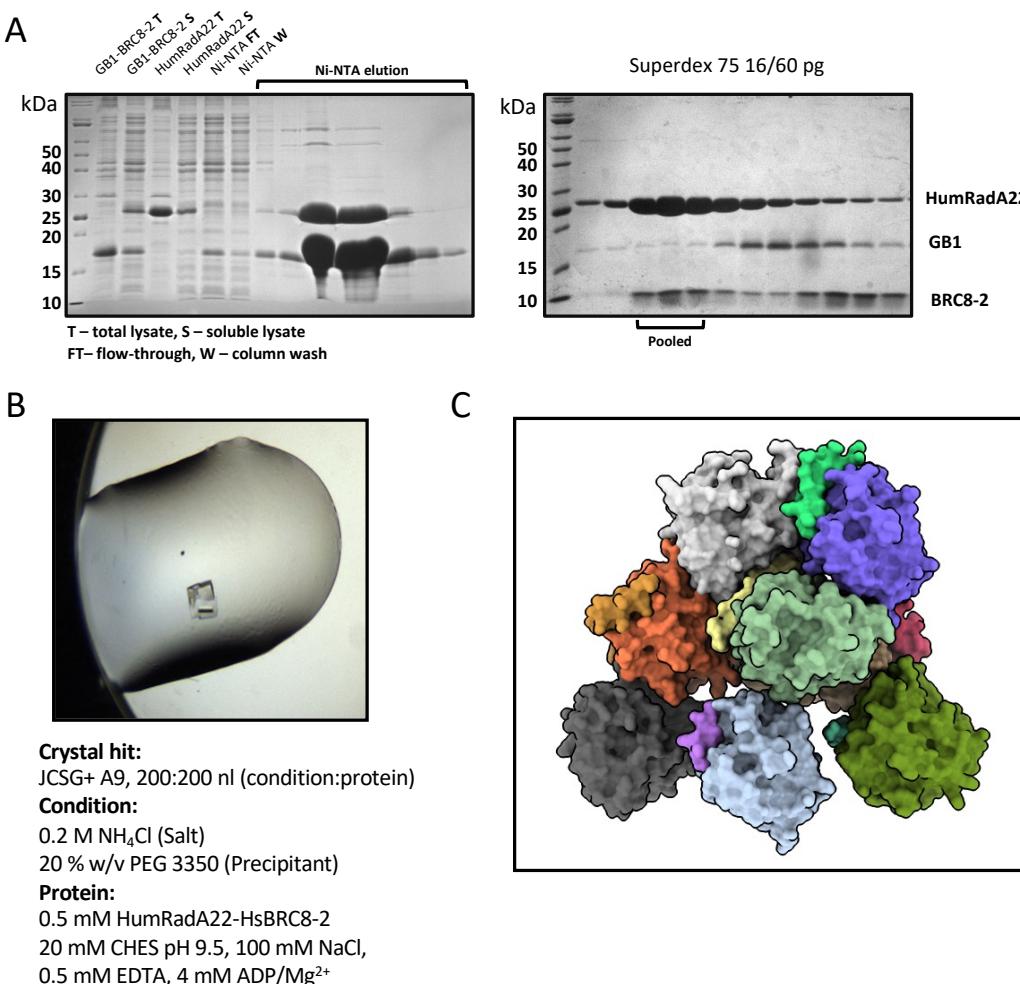
Figure 15. Preparation and crystallisation of BRC8-2:HumRadA22 complex. (**A**) Coomassie stained SDS-PAGE gel depicting the different steps of HsBRC8-2:HumRadA22 complex purification. Individual complex components were solubly expressed in *E. coli*. His-tagged GB1-BRC8-2 was loaded on Ni-NTA resin, followed by application of HumRadA22. The co-eluted complex shows little degradation of the GB1-fused BRC8-2 peptide and high purity. SEC of TEV-cleaved complex successfully separates remaining GB1 contaminants and results in pure complex (pooled fractions). T = total lysate, S = soluble lysate, FT = flow-through, W = wash. (**B**) Crystallisation drop and conditions for the BRC8-2:HumRadA22 crystal hit. (**C**) Asymmetric unit contents with each chain depicted in different colour.

Concentrated HumRadA22-BRC repeat complexes were crystal-screened in 96-well MRC plate format using a variety of pre-mixed commercially available crystallisation screens, with ADP and Mg^{2+} used as stabilising cofactors. Of the four complexes screened, only BRC8-2 yielded suitable diffracting crystals (**Figure 15B**). Diffraction data were successfully collected without further optimisation of the crystallisation condition. The crystal diffracted to 1.9 Å resolution and the complex structure was determined (PDB ID: 6HQU). There are eight complexes in the asymmetric unit (**Figure 15C**), all of which are very similar to each other with an average RMSD of 0.664 Å for 198 Cα atoms in monomeric RAD51. The BRC8-2

peptide was visible in seven out of eight complexes. Electron densities for bound peptides were observed prior their fitting and refinement (**Figure 16**). In all complexes in the asymmetric unit, the overall conformation and binding mode of the peptide was near identical, but with varying extent of visible electron density. The best-defined complex chains B and J were used in the subsequent analyses.
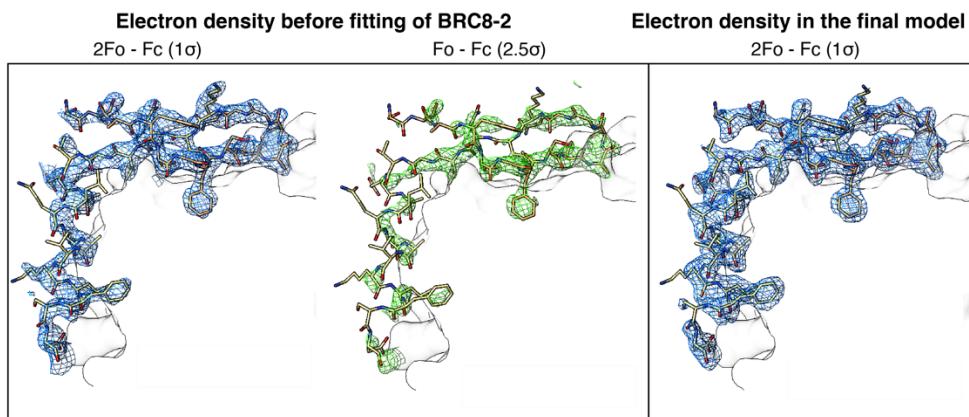
Figure 16. BRC8-2 peptide electron densities prior and after fitting of the peptide. Representative electron density maps are shown for the peptide chain J before the peptide has been modelled in (left) and after fitting with subsequent rounds of refinement (right).

BRC8-2 forms a more extensive β-hairpin and a more optimal α-helix

Comparison of the refined HumRadA22:BRC8-2 structure with that of the RAD51:BRC4 complex (PDB: 1n0w) reveals a similar overall topology (**Figure 17A**). As anticipated, FxxA residues Phe2058[BRC8] and Ala2061[BRC8] form extensive hydrophobic contacts with the FxxA binding site, identical to BRC4 (**Figure 17B**). Ala2061[BRC8] is followed by a β-turn that is stabilised by the Thr2060[BRC8] and Ser2060[BRC8] side-chain hydroxyls, after which the peptide folds back in the opposite direction, resulting in a β-hairpin structure that extends the central β-sheet of monomeric RAD51 in an inter-molecular fashion. At the C-terminal half, the BRC8-2 LFDE module acquires an α-helical secondary structure, and, similarly to BRC4, forms a second interface with HumRadA22. The LFDE motif of BRC4 is replaced by LFSD in the BRC2 module, conserving hydrophobic and acidic hot-spot residues. Leu1240[BRC2] and Phe1241[BRC2] bind the same hydrophobic groove that the BRC4 LFDE motif interacts with (***Figure 17*C**). Asp1243[BRC2] forms a salt bridge with Arg270 (corresponding to Arg254 in *Hs*RAD51), after which electron density becomes poorly defined at the C-terminus.
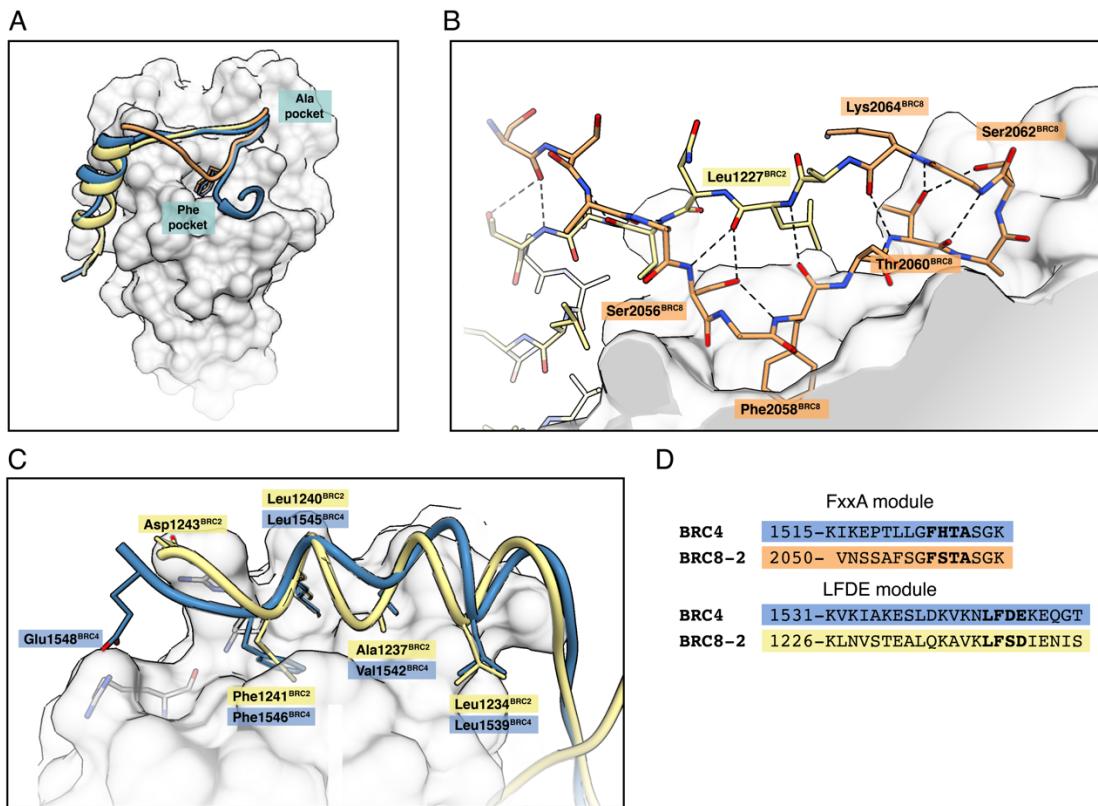
Figure 17. Crystal structure of monomeric RAD51:BRC8-2 complex superposed with RAD51:BRC4. BRC8-2 is depicted in orange and yellow, corresponding to BRC8 and BRC2 sequences, respectively. BRC4 is shown in blue. Peptides were superimposed by aligning the structures of their respective protein binding partners. Monomeric RAD51 is represented by a gray surface. Selected residues of the monomeric RAD51 are depicted in grey. (**A**) Overall topologies of the two peptides. Phe and Ala residues of the FxxA motifs are shown. (**B**) Hydrogen-bonding network of the BRC8-2 β-hairpin. (**C**) LFDE interface with side-chains of crucial residues depicted. (**D**) Sequence alignment of BRC4 and BRC8-2 FxxA and LFDE modules.

The most striking difference between the binding modes of the BRC4 and BRC8-2 peptides is in the extent of the intra-molecular hydrogen-bonding network that forms the β-hairpin in the FxxA module. In BRC8-2, the β-hairpin is significantly extended, with its N-terminal end, before $Phe2058^{BRC8}$, folded back towards the rest of the peptide and interacting with the anti-parallel strand downstream in sequence (**Figure 17B**). As a result, the hairpin feature spans residues $Ser2053^{BRC8}$ to $Thr1231^{BRC2}$, a total of 19 amino acids compared to just 9 in BRC4. A likely structural determinant for the formation of the extended hairpin is $Ser2056^{BRC8}$, whose side-chain fits tightly between the two anti-parallel strands of the peptide and the surface of monomeric RAD51 (**Figure 17B**). The $Ser2056^{BRC8}$ side-chain hydroxyl forms a hydrogen bond with the carbonyl of $Leu1227^{BRC2}$ and the NH of $Phe2058^{BRC8}$, allowing the peptide to fold back on itself in a favourable manner. BRC4 has a bulky and hydrophobic

51

Leu1522 at the equivalent position of Ser2056[BRC8], which is unlikely to satisfy the steric and electrostatic requirements of the topology observed in BRC8-2, forcing the N-terminus of the peptide to point away from the β-hairpin and the rest of the peptide. The extended β-hairpin is observed in all of the 7 peptides modelled in the asymmetric unit, independent of the different crystal packing interactions that surround them, which argues against the possibility of the conformation being a crystal packing artefact. The crystallographic data collection, processing and refinement statistics are provided in **Appendix 8.7**.

Differences in binding mode are also evident at the LFDE modules of BRC4 and BRC8-2. In BRC4, the interface-forming residue Val1542[BRC4] has a bulky hydrophobic side-chain projected into the ATPase domain (**Figure 17C**). To accommodate this, BRC4 has to form an outward facing bulge, resulting in a disrupted secondary structure geometry that deviates significantly from an optimal α-helix. This residue is changed to Ala1237[BRC2] in BRC8-2, causing a shift of the α-helix backbone towards the surface of the protein and resulting in a more helical geometry and tighter fit with HumRadA22 (**Figure 17C**).

2.4 BRC8-2 residue-specific contributions to affinity

The observed structural differences present several potential explanations for the high affinity binding seen for BRC8-2 and for the favourable interactions of its constituent modules, as expressed in terms of ΔΔG. To examine the contributions of these structural features to binding, a number of BRC8-2 mutants were prepared that introduce critical residues from BRC4 that were anticipated to reduce affinity based on the comparative analysis of the crystal structures (**Figure 18A**). BRC8-2[S2056V] has the hairpin-mediating Ser2056[BRC8] mutated to a valine, like in BRC4. Similarly, BRC8-2[A1237V] introduces a valine in place of Ala1237[BRC2], which was expected to disrupt helicity at the LFDE module, similarly to what is seen in BRC4 structure. Moreover, a double mutant BRC8-2[A1237V, S2056V], combining the two substitutions, was prepared to evaluate the cumulative effect of these changes. The peptides were evaluated in plate reader-based FP competition experiments (**Figure 18B, Table 3**). For the single mutants, a modest decrease in potency was observed, with K_D values increasing approximately two-fold. A more pronounced drop in affinity was determined for the double mutant, approaching that of BRC4, suggesting an additive effect from the disruptive mutations, reverting the peptide to a more BRC4-like binding mode, and confirming the structural interpretation. Attempts to test these peptides in ITC were not successful, as significant heats of aggregation at the high concentrations required in the syringe prevented accurate measurements, likely stemming from the increased hydrophobic character of the mutants. A

BRC8-2^{S2056A} mutant was prepared instead, disrupting the hydrogen bonding pattern of Ser2056, but not introducing a BRC4-like valine. BRC8-2^{S2056A} binding to HumRadA22 was measured by ITC and gave a K$_D$ of 6 nM, versus 4 nM for BRC8-2 (**Figure 18C, Table 3**). Accurate determination of low-nanomolar K$_D$ values by ITC is challenging due to uncertainty in isotherm fitting parameters, i.e., a very steep transition in heats at 1.0 molar ratio, therefore these differences cannot be deemed significant. However, there is an unambiguous difference in the enthalpy of binding (ΔH^{BRC8-2} = -7.00 kCal / mol vs $\Delta H^{BRC8-2S2056A}$ = -5.29 kCal / mol), suggesting a different binding mode for the BRC8-2^{S2056A} mutant.
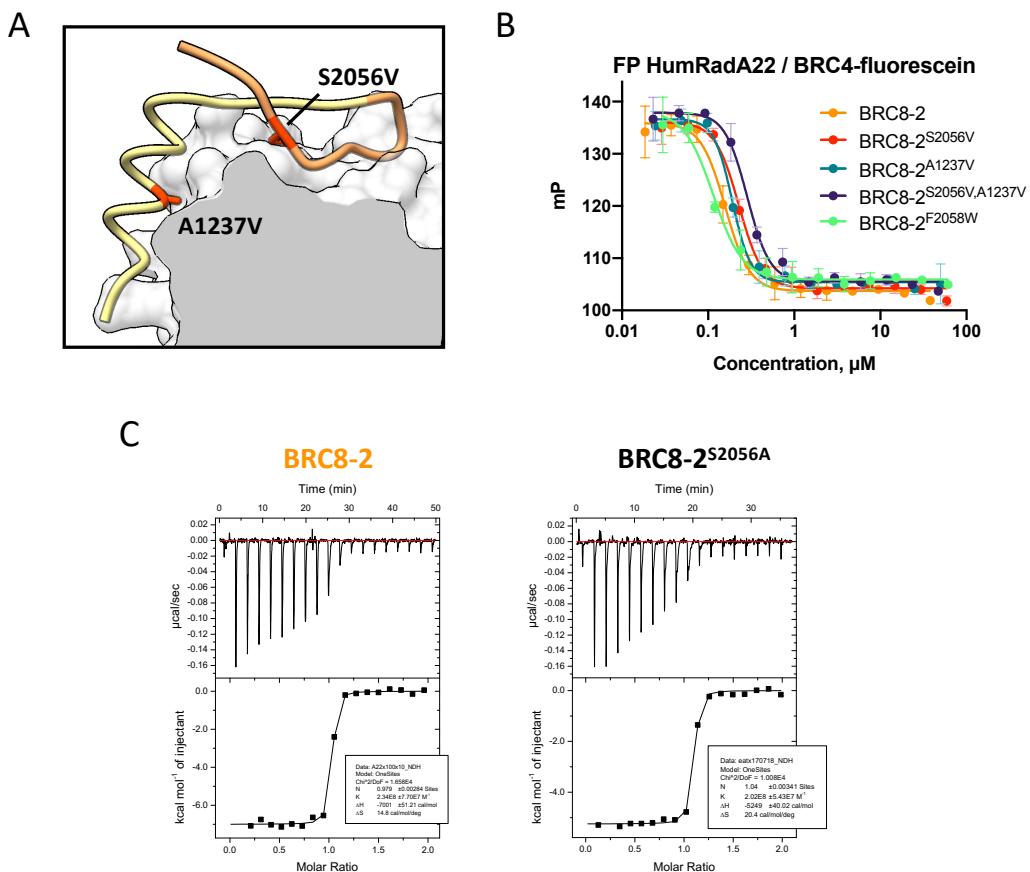
Figure 18. Specific residues in BRC8-2 contribute in a cumulative manner to high affinity binding. (**A**) Critical residues in BRC8-2 that were mutated to BRC4-like amino acids are highlighted in red. Ser2056 mediates the extended β-hairpin at the FxxA module, Ala1237 confers a more optimal helix on the LFDE module. (**B**) FP competition assay titrations of BRC8-2 mutants. Fluorescently-labelled BRC4 probe (10 nM) was pre-incubated with 100 nM HumRadA22, to which peptide dilution series were added. Data shown are the means of triplicate measurements ± SD. (**C**) ITC titrations of BRC8-2 and the BRC8-2^{S2056A} mutant into HumRadA22. All ITC experiments were performed with 100 μM peptide in the syringe and 10 μM HumRadA22 in the cell.

Interestingly, in the peptide shuffling matrix, the FxxA module from BRC3 has the second most favourable $\Delta\Delta G^{FxxA}$ value after BRC8 ($\Delta\Delta G^{FxxA3}$ = -0.35 kCal/mol, **Figure 10F**) and this module contains a threonine at the position equivalent to Ser2056^{BRC8}, which may

similarly enable an extended hairpin conformation by hydrogen-bonding to backbone amides. To further investigate this hypothesis, BRC3:HumRadA22 complex was also purified and crystallisation screening carried out, which yielded a number of hits, however, none of these diffracted.

Peptide	ITC K_D, nM	FP IC50 ± SE, nM	FP K_D, nM
BRC8-2	4	158 ± 9	7
BRC8-2^{S2056V}	-	228 ± 10	16
BRC8-2^{A1237V}	-	191 ± 8	12
BRC8-2^{S2056V,A1237V}	-	285 ± 13	21
BRC8-2^{S2056A}	6	-	-
BRC8-2^{F2058W}	-	115 ± 10	3

Table 3. Summary table of BRC8-2 mutant binding data.

Previous work on FxxA motif tetrapeptides and full-length BRC repeats has shown that replacing the phenylalanine with a tryptophan causes an increase in binding affinity.[152,81] To probe whether this modification can be utilised in the context of a BRC repeat in order to improve the binding affinity of BRC8-2, a mutant BRC8-2^{F2058W} peptide was also prepared and tested in the FP assay (**Figure 18B**). The peptide displays a marginally higher affinity over BRC8-2, an observation that needs further validation given the small effect size (BRC8-2 IC$_{50}$ = 158 ± 9 nM, BRC8-2^{F2058W} IC$_{50}$ = 115 ± 10 nM). The improved binding of a BRC8-2^{F2058W} mutant and other similar sequence-activity relationships can be utilised in the development of BR8-2 based pharmacological probes as a means of increasing target affinity.

2.5 Conclusions

The microfluidic FP screening of chimeric BRC repeats derived through modular recombination had produced novel sequences that exhibited improved binding over the eight wild-type repeats. Intrigued by the increased affinity of the chimeric repeat BRC8-2, I undertook an extensive examination of this peptide. The sum of biophysical and biochemical data obtained strongly supports BRC8-2 as a novel high affinity peptide. Using both ITC and FP experiments, I confirmed that BRC8-2 binds more strongly than any of the parental repeats or BRC4. Moreover, the crystal structure of BRC8-2:HumRadA22 complex shed new light on our understanding of BRC repeat-RAD51 interactions by highlighting novel structural features that modulate binding affinity.

High affinity epitopes can be effectively utilised as templates for the development of peptide-based pharmacological probes. For example, stapled peptides based on the p53 tumor suppressor protein are being investigated in the clinic as inhibitors of the MDM2 proto-oncogene.[153] The biophysical, biochemical and cellular data presented in this chapter provide a clear rationale for investigating BRC8-2 as a template for the design of peptidic RAD51 inhibitors. Moreover, the crystal structure allows for the rational, structure-based introduction of chemical modifications, for example, it can guide the placement of covalent linkages during peptide macrocyclisation. With this information at hand, I undertook the task of developing BRC8-2 into a potent inhibitor of human RAD51 using a peptide stapling approach.

3 Development of RAD51-targeting stapled peptides

3.1 Introduction

RAD51 presents several potential approaches for its inhibition. A classical small-molecule targeting of the enzymatic ATPase active site may be used to inhibit active NF formation, diminishing the strand exchange activity of RAD51. However, an enzymatic mechanism of action is likely to be challenging due to the highly conserved nature of the nucleotide binding site and the need to compete with the abundance of cytosolic nucleotide co-factors. For example, ATP is present in the cell in millimolar amounts, and binds RAD51 with micromolar affinity, meaning that a successful enzymatic small-molecule inhibitor would need to be of low nanomolar potency to achieve high occupancy, a formidable task for the relatively shallow and dynamic nucleotide interface.

Alternatively, the BRC repeat FxxA and LFDE module interfaces could be targeted to inhibit BRCA2 binding and/or self-oligomerisation. The high-affinity BRC8-2 repeat provides an attractive starting point for the development of macrocyclic peptide probes, as it encompasses a functional binding interface and contains a large number of residues with solvent-exposed side chains that can used for covalent linking in a peptide stapling approach. The peptide's low-nanomolar affinity can support high target occupancy, provided that cellular uptake is sufficiently high.

Evaluating all of the possible combinations of linked residues on the peptide would be a daunting task. For example, a peptide with *n* residues theoretically permits $\frac{n^2 - n}{2}$ different combinations of a single cysteine pair. For a 12 aa peptide, this means 66 possible pairs, whereas for a 32 aa BRC repeat this number increases to 496. Each type of chemical linker favours a range of inter-residue distances and macrocycle conformations and stapling residues can be introduced with this in mind. Linkers are small molecules with only a few intramolecular conformational degrees of freedom and thus impose a constraint on inter-cysteine distances, which significantly narrows down the search space if one seeks to preserve peptide binding. The above-described crystal structure of the shuffled BRC8-2 repeat can be used to inform the rational design of stapling positions to be introduced in the sequence.

In this work, I decided to use a cysteine-reactive stapling methodology to obtain macrocyclic peptide RAD51 inhibitors, utilising BRC8-2 as a high-affinity template. The potential advantages of different stapling methods have been considered in **Section 1.12** *Peptide stapling*. Moreover, my decision to develop a recombinant methodology rather than

use standard SPPS protocols was motivated by the fact that bacterial culture is more widely applicable to academic biology labs lacking the infrastructure for organic synthesis and procedures for disposal of large volumes of organic solvents, such as DMF. Successful application of recombinant peptide stapling could encourage the adoption of this method in the wider scientific community.

A set of divinyl linkers developed at the lab of Prof David Spring (Department of Chemistry, University of Cambridge) were used in this project. Composed of a heterocycle core and two electrophilic vinyl arms, these linkers act as Michael acceptors in reaction with cysteine thiols (**Figure 19**). Initially developed for irreversible re-bridging of reduced antibody disulfides, these linkers comprise a new generation of bio-conjugation reagents that demonstrate many advantages over traditional labelling and stapling chemistries. Importantly, these linkers:

1. Show exceptional selectivity over other nucleophilic side-chains and N-terminal amines.
2. Are applicable in a wide range of aqueous and organic conditions.
3. Result in irreversible and stable thioether products.
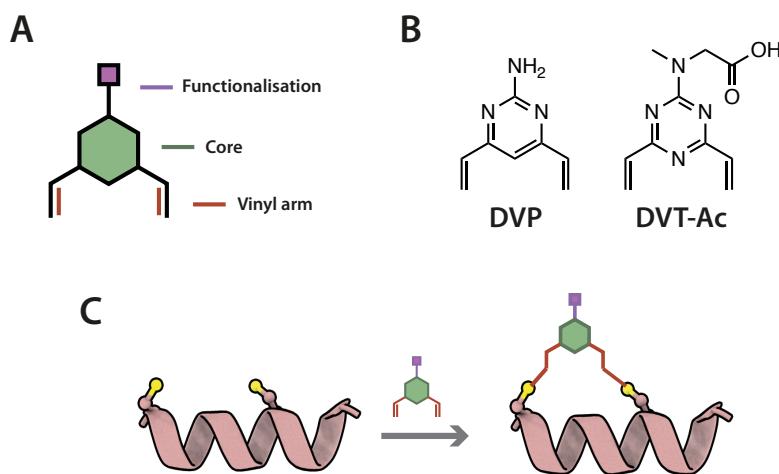
Figure 19. Divinyl linkers for cysteine stapling of peptides. (**A**) A diagram depicting the components of a divinyl linker. Reactive electrophilic arms are 1,3 to each other and each arm reacts with a cysteine thiol. The core heterocycle acts as scaffold and fine-tunes the reactivity of the vinyl groups. (**B**) Divinyl linkers used in this study. Divinylpyrimidine (**DVP**) contains an aminopyrimidine moiety and is less reactive. Acetyl-divinyltriazine (**DVT-Ac**) is a more soluble and more reactive triazine core linker. (**C**) A diagram depicting a stapling reaction of bis-cysteine helical peptide at *i,i*+7 residue positions encompassing approximately two helical turns.

It has been previously shown that these linkers can efficiently constrain α-helical motifs in small peptides, and are most optimal for joining cysteines at positions *i,i*+7, resulting in a

macrocyclic structure that encompasses two helical turns.[146] It is also anticipated that, with appropriate geometry considerations, these linkers can be applied to non-helical peptide folds.

3.2 A recombinant approach for peptide stapling

Introduction of a covalent linkage between different side-chains of a peptide can have a range of both beneficial and detrimental effects on peptide binding. Conformational changes imposed by the linkage can prevent the peptide from achieving an optimal binding mode and cause a drop in affinity. For example, a linker can completely disrupt α-helical secondary structure if the stapled residues are not optimally spaced. Moreover, the linker itself can clash sterically with the protein if introduced on a helical face that is not sufficiently solvent-exposed. On the other hand, stapling at an optimal position can have a positive contribution to affinity through pre-organisation of peptide conformation resulting in a reduction of entropic penalty, as well as through favourable linker-protein contacts. Rapid screening of different stapled peptides would allow for early identification of residue positions that are most optimal for linkage. To evaluate the effect of introducing cysteine linkages on BRC repeat binding, a small-scale recombinant production methodology was trialled. In this approach, bis-cysteine mutants were expressed in bacterial culture, purified and stapled at small scale in as few steps as possible, while obtaining sufficient purity and yields for biochemical evaluation. The two main obstacles to achieving successful production and isolation of linear peptides from bacterial cultures are low yield and proteolytic degradation, with the latter often contributing to the former. To address these methodological issues, an expression plasmid containing optimal expression tags for this purpose was designed (pPEPT1, **Figure 20A**). Using this vector, the peptide of interest (POI) was produced as a fusion to an N-terminal GB1 protein tag, derived from the B1 domain of streptococcal G protein. The GB1 tag is a small, monomeric and highly soluble globular protein domain that is very efficiently expressed in *E. coli*. On the C-terminus of the expressed construct a His$_8$-tag was introduced for the purification of the peptide fusion using immobilized metal affinity chromatography (IMAC) on Ni-NTA resin. The N-terminus of the construct also contains a Strep-II tag for a possible affinity purification on Strep-Tactin resin. The peptide DNA insert can be rapidly prepared using oligonucleotide annealing or one-pot PCR assembly and introduced seamlessly into the digested vector using sequence and ligation independent cloning (SLIC).

A bis-cysteine BRC8-2 model peptide (**SP2**) was designed for the testing and optimisation of this stapling methodology. The GB1-peptide fusion was expressed in 10 ml *E. coli* culture and purified on a 100 µl Ni-NTA spin column. As analysed by SDS-PAGE, the

fusion is overexpressed and efficiently purified from the bacterial culture (**Figure 20B**). The peptide runs at a higher molecular weight than calculated, a common phenomenon observed with GB1. 200 µl of ~100 µM GB1-peptide could be prepared in such way.
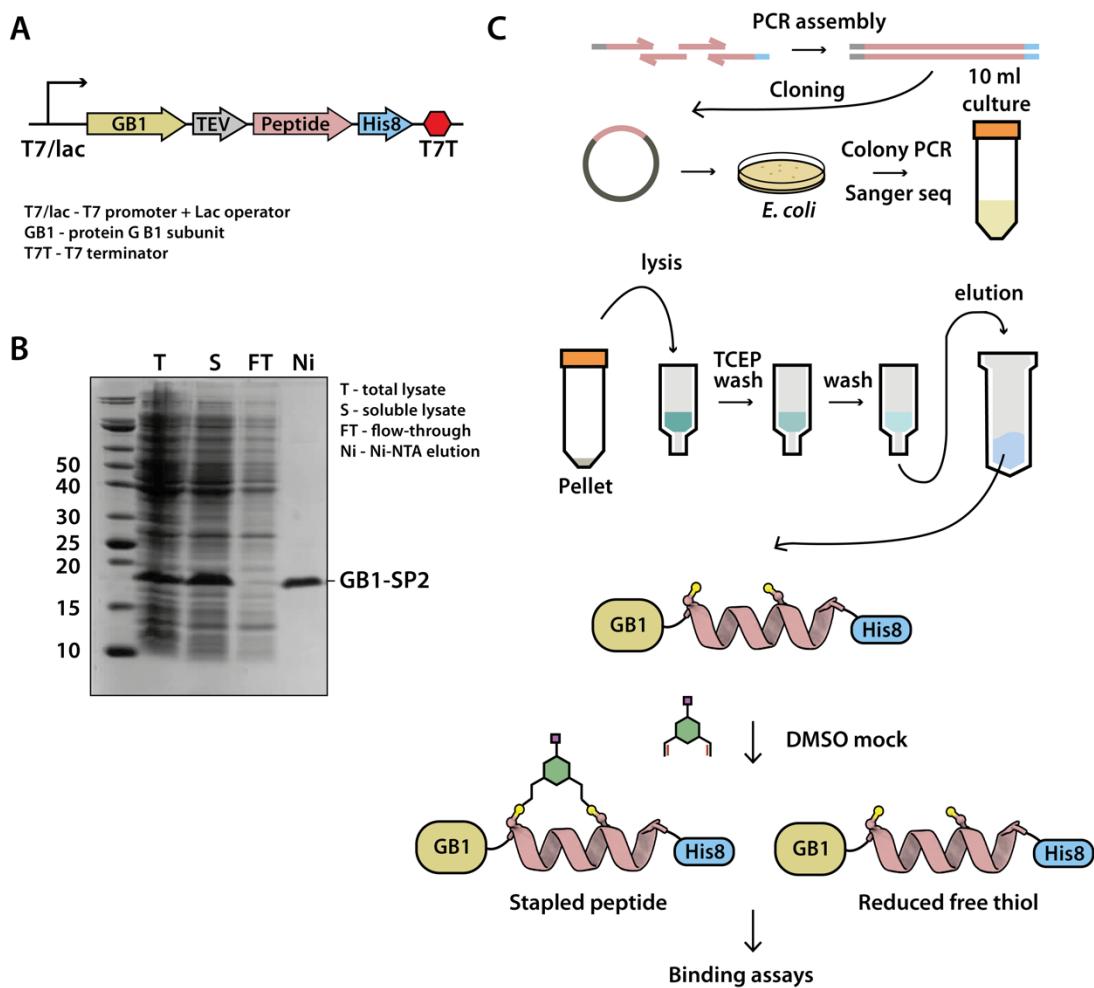
Figure 20. A recombinant stapled peptide preparation platform. (**A**) Illustration of the expression cassette used in the pPEPT1 vector for the preparation of GB1-fused stapled peptides (**B**) Purification of a GB1-fused model peptide **SP2** from a small-scale (10 ml) *E. coli* culture. T = total lysate, S = soluble lysate, FT = flow-though, Ni = Ni-NTA elution (**C**) An illustration of the small-scale stapled peptide preparation work-flow.

Having established that high levels of protein expression are achievable with this system, a general procedure for small-scale stapling of recombinant peptide fusions was devised (**Figure 20C**). Using this approach, the linker was added directly to the Ni-NTA output, after which the reaction was quenched and the product used in biochemical evaluation without any further purification steps, as the reagents are not expected to interfere with in vitro assays. The divinyl linkers used in this study have been reported to react rapidly and selectively with reduced cysteines. A successful divinyl stapling reaction has a two-step mechanism: the first step is a bimolecular Michael addition reaction between an electrophilic vinyl arm of the linker molecule and a deprotonated thiol nucleophile from one of the cysteines within the

peptide (**Figure 21A**). This is followed by a unimolecular ring closing reaction of the reaction intermediate **2**, where the second vinyl arm reacts with the second cysteine sulfhydryl, again through a Michael addition, giving a stapled peptide **3**. The intramolecular macrocycle closing step is in competition with an undesired bimolecular side-reaction, where a second linker molecule reacts with the remaining free cysteine (**Figure 21B**). This results in a double-linker product **4** that does not form the correct covalent linkage between cysteine side-chains and is therefore of no pharmacological utility compared to the linear peptide.
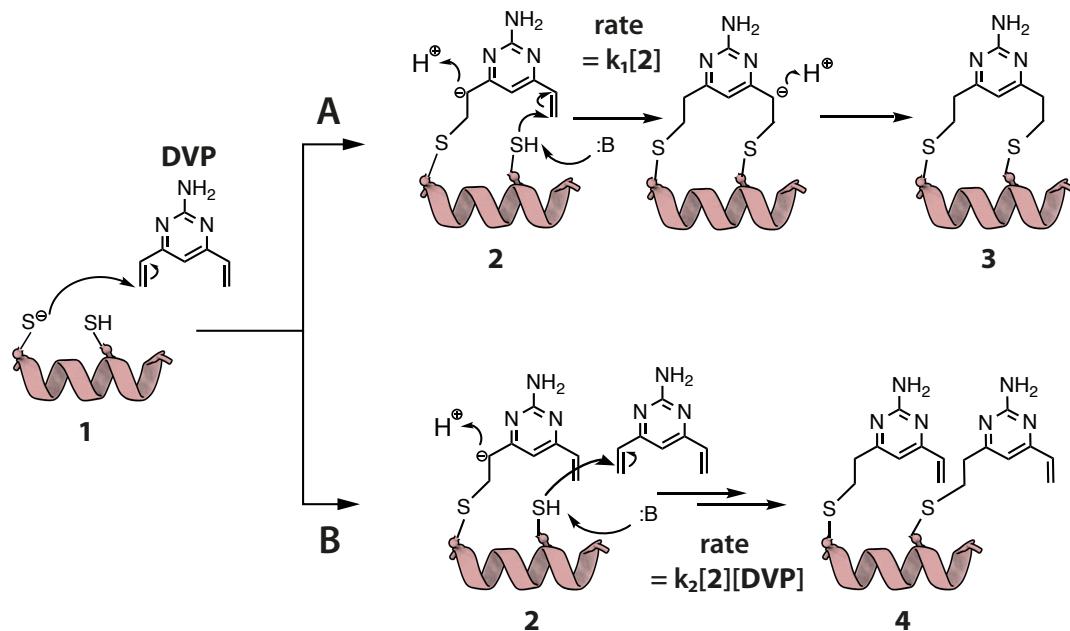
Figure 21. Curly arrow mechanisms for a bis-cysteine peptide reacting with a divinyl linker. (**A**) Reaction pathway for the formation of the correct stapled peptide product **3** via an intramolecular ring closing reaction of the intermediate **2**. (**B**) Reaction pathway for the formation of an undesired double-linker product **4** via a second **DVP** linker addition.

A number of kinetic factors influence the rates of these competing steps and favour the formation of either product. The first-order reaction rate for the macrocycle closing can be expressed as the product of the rate constant k_1 and concentration of the intermediate **2** (**Figure 21A**), whereas the attachment of the second linker follows second order kinetics and depends on concentrations of both intermediate **2** and the concentration of the divinyl linker (**Figure 21B**). If the concentration of the linker is decreased, the rate of the initial bimolecular addition is slowed down proportionally to this change in concentration. Thus, the unimolecular macrocycle closing reaction that follows is also slowed down to the extent that there is less intermediate **2** available. On the other hand, the bimolecular second linker addition is slowed down both due to lower intermediate **2** concentrations, *and* as a result of less free **DVP** linker available. Thus, lowering the reactive linker concentration in the reaction should slow down

this undesired reaction more and favour the formation of the correctly stapled product. A convenient way to maintain the linker concentration low, while achieving quantitative stapling of the peptide, is to use pseudo-dilution, an approach where one of the reactants is titrated into the reaction vessel in incremental amounts and is allowed to react before further addition of the same reactant.

The suitability of pseudo-dilution was tested on small-scale preparations of the bis-cysteine mutant GB1-**SP2** in combination with the **DVT-Ac** linker. Several titration conditions were trialled with varying volumes and time-points of linker addition (**Figure 22A**). While this is not a comprehensive scan of a large number of possible reaction conditions, a sufficient optimum could be established for use in subsequent screens of BRC8-2 mutants. Products were analysed by intact MS and relative sizes of mass peaks were used as a metric for purity. Chromatographic estimation of yields was not possible because the GB1-fused products are physico-chemically too similar to be separated on any matrix.

Representative mass spectra of the different conditions are provided in **Figure 22B**. In all of the conditions tested, the most abundant peak corresponds to the correctly stapled peptide **6** with an N-terminal methionine cleavage (m = 14773.97 Da). A smaller peak is also present corresponding to N-terminal Met-Ser cleavage of the same species (peptide **5**, m = 14686.89 Da, Δm = 87.08). Double-linker by-products (peptide **7**, m = 14994.2 Da) were also observed, and were favoured by faster linker titration, e.g. when comparing condition **a** and condition **b**. In addition to this, impurities of dimeric nature were detected at masses approximately twice the GB1-**SP2** construct (condition **i**). These peaks correspond to a range of disulfide- and divinyl-linked adducts (**Figure 23A**). The pKa of the free cysteine thiol is around 8 and cysteines within the context of a peptide or protein chain can exhibit a wide range of pKa values. The propensity of the cysteine side-chain to form disulfides is therefore susceptible to modulation by pH at the near-physiological conditions used here. Decreasing the pH from 8 to 7 caused a significant decrease in the abundance of the dimeric species (condition **h**), while the drop from pH 7 to pH 6 did not have any further observable effect.
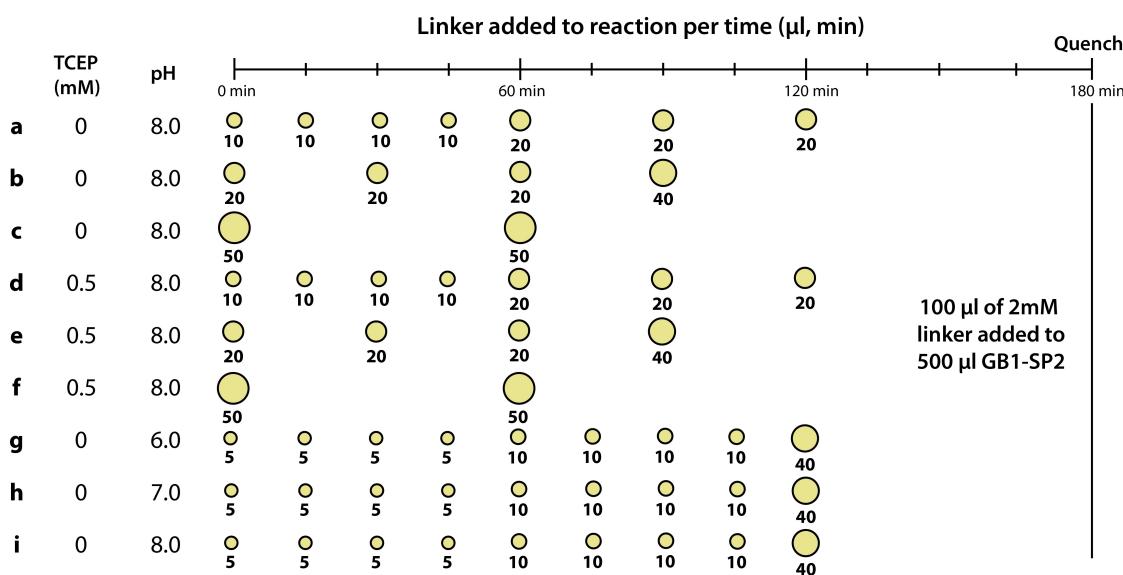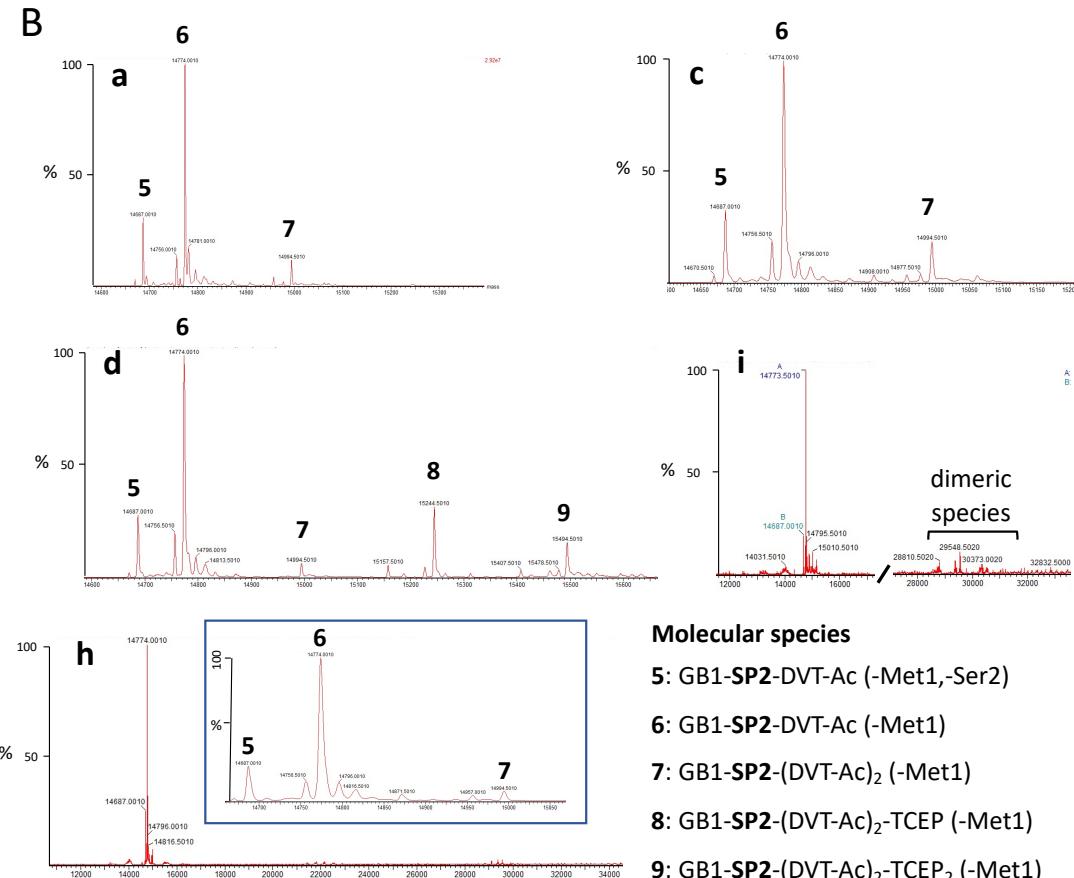
Figure 22. GB1-fused bis-cysteine peptide stapling optimisation. (**A**) A detailed overview of the stapling reactions conditions tested. Samples **a-i** represent combinations of titration timing, pH, and reducing agent parameters. Each reaction was done in 0.5 ml (final) at room temperature on a rotating mixer. (**B**) Mass spectra of representative stapling conditions from (**A**). **5** to **9** denote the different molecular species observed.

A critical consideration in cysteine bio-conjugation reactions is the application of a reducing agent. While it is important to prevent the formation of disulfide bridges, reducing

agents, particularly thiol-based ones, are susceptible to reactivity with the linker electrophile. TCEP has been commonly applied as a selective, non-thiol reducing agent, for example, in maleimide-coupling reactions.[154] Here, inclusion of even a relatively low concentration (0.5 mM) of TCEP in the divinyl stapling reaction resulted in the formation of undesired Cys-linker-TCEP adducts. The phosphine group of TCEP is nucleophilic and can react with the vinyl arms via Michael addition. In fact, alkyl phosphine addition to acrylamide has been suggested as a novel click reaction for bio-conjugation.[155] The putative structures of the observed TCEP adducts are shown in figure **Figure 23B**. To overcome this limitation, GB1-**SP2** was reduced on-resin with a TCEP-containing wash, after which the construct was washed with buffer lacking the reducing agent, and eluted shortly thereafter. Stapling was performed immediately on the eluted protein. At pH 7.0, disulfide formation rate is expected to be negligible relative to the Michael addition reaction, and the optimised condition **h** yielded the desired single-stapled product with very minor (<5%) formation of undesired by-products as analysed by MS (**Figure 22B**). The reaction is robust, resulting in highly pure product when performed with a range of different bis-cysteine mutants (**Section 3.3** and **Appendix 8.2** *Representative protein LC-MS spectra of GB1-fused stapled peptides*).
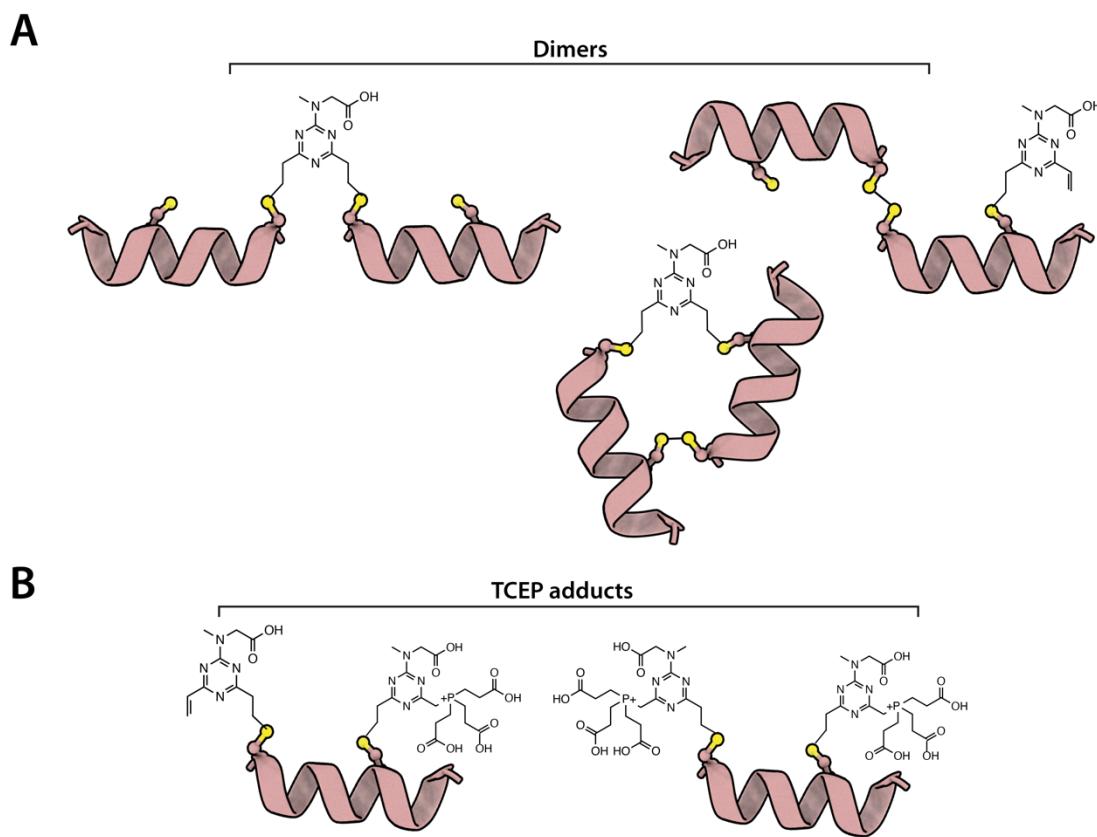
Figure 23. By-products of cysteine stapling. (**A**) Some of the dimeric species that may arise as a result of linking of two peptides though disulfide bridges or through the linkers, or both. (**B**) Schematic illustration of the putative TCEP adducts observed by MS.

3.3 Design and screening of stapled RAD51 binders based on the chimeric BRC8-2 repeat

With an optimised stapling procedure at hand, I proceeded with the design of BRC8-2 bis-cysteine mutants to be evaluated, with the following objectives in mind:

1. To introduce covalent linkages that would encompass a significant fraction of the BRC8-2 template, including the hot-spot motifs that are critical for binding.
2. To evaluate a variety of stapling architectures, including non-helical motifs in the macrocycle.
3. To decrease the size of the stapled peptides by truncating parts that are not essential for binding.

Peptides were designed in a structure-guided manner, informed by the crystal structure of the BRC8-2 complex (**Figure 24** & **Figure 25A**). Two design strategies were employed. In a traditional helical stapling approach, different *i,i*+7 Cys pairs were introduced at the solvent-exposed face of the α-helical part of the LFDE module (peptides **SP1**, **SP2**, **SP8**, and **SP9**). For example, in **SP2**, a Thr1231 residue at the start of the α-helix is replaced by a cysteine, which is then linked to another cysteine in place of Val1238, both of which are solvent-exposed, are found on the same face of the α-helix and have a continuous unoccupied space available between the two side-chains, allowing the linker to be accommodated by the α-helical surface. It can be inferred from the BRC8-2 complex X-ray structure that a different stapling arrangement at the helix, e.g. *i,i*+8, where Thr1231 is linked to Lys1239 instead, would likely result in a clash between the linker and the side-chain of Gln1235.

In a different approach, stapled BRC8-2 mutants were designed with at least one cysteine at the non-helical β-hairpin substructure of the peptide (**SP10**, **SP12**, **SP13**, **SP14**, **SP16**). The advantage of such stapling architecture is that a larger portion of the template can be cyclised, leading to potentially higher proteolytic protection, as well as increased pre-organisation of the peptide. Placement of cysteines in these mutants was guided by the X-ray model and did not conform to sequence constraints, spatial distances permitted by the linker molecule were considered instead. For example, **SP12** links a solvent-exposed Phe2055 near the N-terminus of β-hairpin with Leu1234 at the middle of α-helix of the LFDE module.
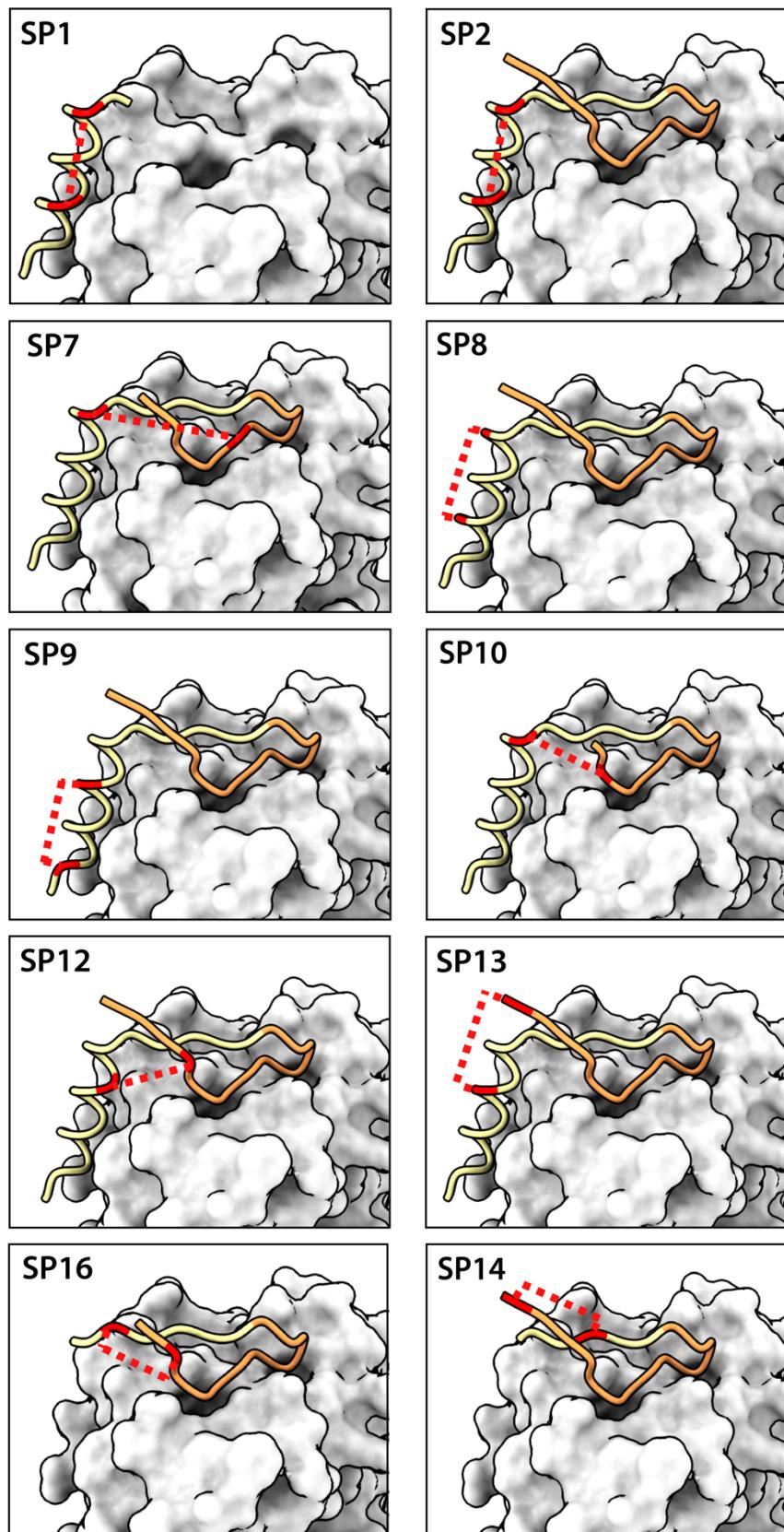
Figure 24. Designs of stapled peptides based on the BRC8-2:HumRadA22 complex structure. BRC8-2 backbone is depicted in orange and yellow, corresponding to BRC8 and BRC2 sequences, respectively. Residues mutated to cysteines are rendered in red. Dotted lines represent the approximate direction of the intended staple moiety.

Rather than using the full-length BRC8-2 sequence template, **SP10**, **SP12** and **SP13** were designed with N-terminal truncations in order to sterically and conformationally accommodate the backbone rearrangements that were anticipated to result from stapling. In a more extreme example of truncation, peptides **SP1**, **SP14** and **SP16** were designed to significantly decrease peptide size while maintaining either the FxxA or LFDE module. Previously, a BRC4 half-peptide encompassing residues Pro1519-Val1532 was shown to bind RAD51 with a K_D = 36 μM.[117] It is possible that the superior BRC8-2 template may improve on this significantly, allowing the development of probes that are reduced in size and therefore more likely be cell-permeable.

A negative control peptide **SP7** was designed with one cysteine replacing the Ser2059 from the FxxA motif (FSTA in BRC8-2) and another placed at Thr1231 of the LFDE module. Such a mutant is expected to have a significant drop in affinity upon stapling, as a result of the bringing together of two amino acids that are not in close proximity in the BRC8-2 binding mode, causing a disruption of the peptide fold.

All of the designed GB1-fused mutants were successfully purified from bacterial cultures and subsequently modified with the DVT-Ac linker. To examine the effect of introducing a covalent staple on affinity, a mock reaction was done in parallel for each peptide by splitting the Ni-NTA purification output, and adding TCEP + DMSO to the mock control reaction instead of the divinyl linker. TCEP is added to the mock reaction to prevent disulfide formation over the reaction time-course. The mutant cysteine thiols are designed to be solvent exposed based on the X-ray structure and the peptide binding mode is expected to accommodate these mutations without interference with affinity. Comparison with the reduced cysteine control, rather than just the cysteine-less BRC8-2, provides a more detailed examination of the modulation of binding by the stapling procedure. For example, a significant increase in affinity upon stapling would suggest that either the peptide becomes pre-organised and therefore has a lower entropic penalty of binding, or, alternatively, that the linker itself is forming constructive interactions with the target. Such observations would strongly support the choice of a particular stapling architecture over alternative designs.

Correct products were confirmed by MS for a number of the stapled peptide fusions with minimal impurities, and show very similar compositions to the initially optimised test conditions (**Appendix 8.2** *Representative protein LC-MS spectra of GB1-fused stapled peptides*). Peptides were evaluated for binding to HumRadA22 using the same fluorescence polarisation (FP) competition assay described in **Section 2.1**.

Representative FP assay measurements are shown in **Figure 25B,** with stapled peptide binding in green and mock reactions in red. The negative control repeat **SP7** ha a K_D of 39 nM in the free thiol state, whereas the stapled product has a K_D of 149 nM, a more than three-fold drop in affinity, confirming the modulation of binding by stapling. Gratifyingly, repeat mutants **SP2**, **SP8**, **SP9**, **SP10**, **SP12** and **SP13** all bind with high affinity following stapling (**Figure 25A**). For these peptides, differences in K_D were minimal between the corresponding stapled and mock reaction products, and all of these peptides bound with K_D values of 15-60 nM in either form. **SP12** bound more tightly after stapling and has the highest affinity observed amongst all of the mutants.

Figure 25. Screening of stapled BRC8-2 peptides. (**A**) Bis-cysteine mutants designed based on the BRC8-2 sequence and prepared in small-scale. FP K_D values are shown on the left. (**B**) FP competition assay graphs for representative peptides. Fluorescently-labelled BRC4 probe (10 nM) was pre-incubated with 100 nM HumRadA22, to which GB1-peptide dilution series were added. Data shown are the means of triplicate measurements ± SD.

    None of the truncated peptides **SP1**, **SP14**, **SP16** bound to saturation at the tested concentrations, either in stapled or unstapled form, reiterating the previously published

observations that the simultaneous binding of both FxxA and LFDE motifs is critical for high affinity interaction of BRC repeats with RAD51 (**Figure 25A**). While its binding could not be fitted, **SP14** showed the most substantial displacement of the fluorescent probe of the three short peptides, with an IC_{50} around 10 μM. Interestingly, **SP14** appears to bind more strongly after stapling (**Figure 25B**), indicating the stapling architecture is beneficial, despite an overall low affinity.

The small-scale preparation and screening of cysteine-stapled peptides established that rational, structure-based macrocyclisation can be applied to the chimeric BRC8-2 repeat. However, none of the repeats, either stapled or mock, bound with a higher affinity than the equivalent GB1-BRC8-2 template construct (K_D = 11 nM), suggesting that the introduction of the cysteine staple may be slightly detrimental for binding.

3.4 Optimisation of stapled BRC8-2 peptides

The small-scale screening data provided a detailed description of the effect of different stapling geometries on binding, however, the resulting products are not applicable in cell assays or for crystallographic studies. Scale-up and removal of tags is necessary to obtain the peptide in sufficient yields and a pharmacologically relevant form. The screening data were used to inform the design and preparation of BRC8-2 bis-cysteine mutants as free peptides, lacking any expression and purification tags. The preparation work-flow is illustrated in **Figure 26A**. Peptides were cloned into vectors having an N-terminal His_8-tag, followed by a GB1 tag and a TEV site. This allows the separation of the peptide from all of the fusion tags using a single TEV cleavage step. After expression in litre-scale *E. coli* cultures and purification on Ni-NTA resin, the linear peptides were cleaved with TEV protease, reverse-purified in a second IMAC step to remove the released N-terminal tags, followed by reversed-phase chromatography (RPC), which allowed the separation of degradation products from the full-length peptide. After this, the linear bis-cysteine peptides were stapled with a divinyl linker and purified by RPC again. Pseudo-dilution conditions required for the correct product formation were applied similarly to the screening protocol but adjusted for a larger reaction volume.
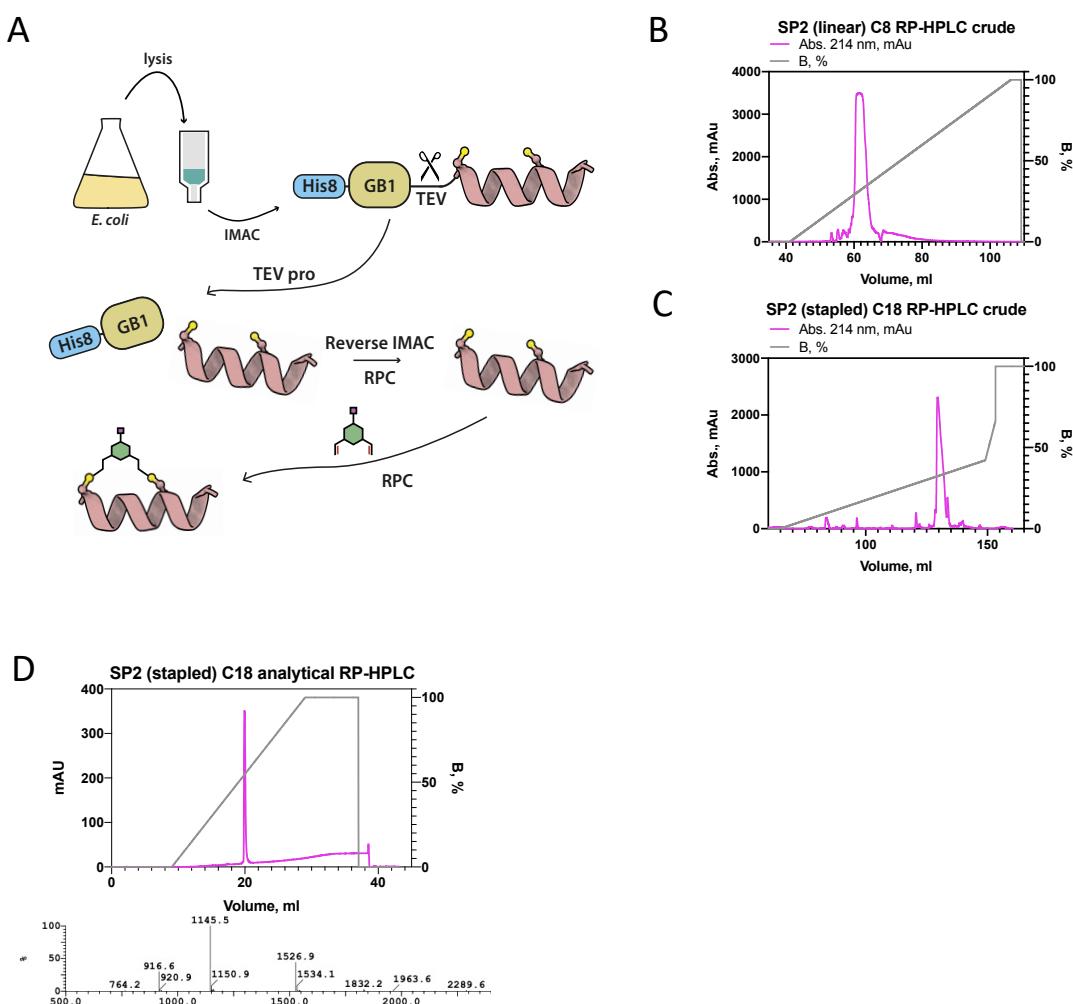
Figure 26. Preparation of stapled BRC8-2 peptides. (**A**) A diagram outlining the procedure for recombinant expression and purification of linear bis-cysteine peptides, and their subsequent stapling. Peptide is expressed as a fusion to an N-terminal His$_8$-GB1 tag and purified by an initial IMAC capture step. Cleavage with TEV protease and a reverse IMAC step then ensures removal of the tag. The linear peptide is then purified by RPC (typically C18 column), followed by stapling and a final C18 RPC step. (**B**) Prep-grade RPC (C8) trace of the TEV-cleaved linear **SP2** after a reverse IMAC step (**C**) Crude reaction product the same peptide after stapling with DVT-Ac under pseudo-dilution conditions. (**D**) Analytical HPLC trace (top) and MS spectrum (bottom) of the **SP2** final product.

SP2 was selected initially to evaluate the scale-up approach. RPC purification of the cleaved linear intermediate gave high yield (>10 mg) of the full-length product **Figure 26B**. Because the two cysteines had been reduced with TCEP prior the application onto the RPC column, and peptide was kept in acidic conditions during and after the first RPC step, no subsequent addition of reducing agent was required. The DVT-Ac Linker was titrated into phosphate-diluted pooled RPC fractions with magnetic stirring, forming a highly pure stapled product with no detectable double-linker contaminant (**Figure 26C,D**). Notably, a simplified stapling protocol lacking pseudo-dilution was also attempted, adding all of the linker added at

once. Here, the reaction products elute as two peaks with the second peak corresponding to the double linker product (data not shown). This again emphasises the need for careful consideration of reaction kinetics to obtain the correct stapled peptide.

Figure 27. Bi...
peptides prepa...
(green), nucle...
HumRadA22.

Bind...
humanised ...
filament for ...

		FxxA module	LFDE module	Linker	
	BRC8-2	VNSSAFSG**FSTA**SGKKLNVSTEALQKAVK**LFSD**IENIS			
Pantele	SP2	NH₃-**GS**VNSSAFSG**FSTA**SGKKLNVS**C**EALQKA**C**K**LFSD**IENIS-COOH		DVT-Ac	*peptides*
	SP24	NH₃-**GS**G**FSTA**SGKKLNVS**C**QALQKA**C**K**LFSG**-COOH		DVP	
	SP30	NH₃-**GC**SG**FSTA**SGKKLNVST**Q**A**C**QKAVK**LFSG**-COOH		DVP	
SP2 cy	SP31	NH₃-**GR₉GGC**SG**FSTA**SGKKLNVST**Q**A**C**QKAVK**LFSG**-COOH		DVP	
	SP32	NH₃-**GR₉GGC**SG**FSTA**SGKKLNVST**Q**A**C**QKAVK**LFSG**-COOH		DVP-Fluor	

the aim of imp d on the module.

SP30 is based on **SP12** and contains a more distant $i,i+18$ linkage in sequence across the two modules. The optimised peptides were significantly truncated at both termini compared to their parental sequences in order to decrease size with the aim of achieving improved membrane permeability. The peptides were also optimised to decrease negative charge, which has been likewise shown to be important for cellular uptake. First, the negatively charged DVT-Ac linker was exchanged for a neutral DVP (**Figure 19B**). Secondly, Asp1243 from the LFDE motif (LFSD in BRC8-2) was now at the C-terminus, meaning that this residue would carry two negative charges. To prevent this, Asp1243 was mutated to a glycine in both **SP24** and **SP30**. In the context of BRCA2, this residue interacts with RAD51 surface arginines through its acidic side-chain. A flexible glycine was deemed sufficient for binding, as the terminal carboxylate is only one carbon closer to Cα and can mimic the acidic side-chain of Asp1243. Moreover, a solvent-exposed Glu1232 was mutated to a glutamine to further decrease negative charge.

Gratifyingly, the resulting peptides were found to bind Hu...22 with affinities that are comparable to **SP2**, despite losing more than a quarter of residues (**Figure 27B**). Moreover, placement of a terminal glycine instead of Asp1243 at the LFDE motif (LFSD in BRC8-2) did not severely impair binding. **SP24** and **SP30** also inhibited nucleofilament formation in EMSA assays utilising full-length human RAD51 at comparable levels to BRC8-2 (**Figure 28**).

Peptide, µM

BRC8-2   SP2   SP24   SP30

0 0.1 0.3 0.9 2.7 8.3 25′   0.1 0.3 0.9 2.7 8.3 25   0.1 0.3 0.9 2.7 8.3 25   0.1 0.3 0.9 2.7 8.3 25

Rad51 : FAM-dT60

FAM-dT60

+ 5 µM Rad51
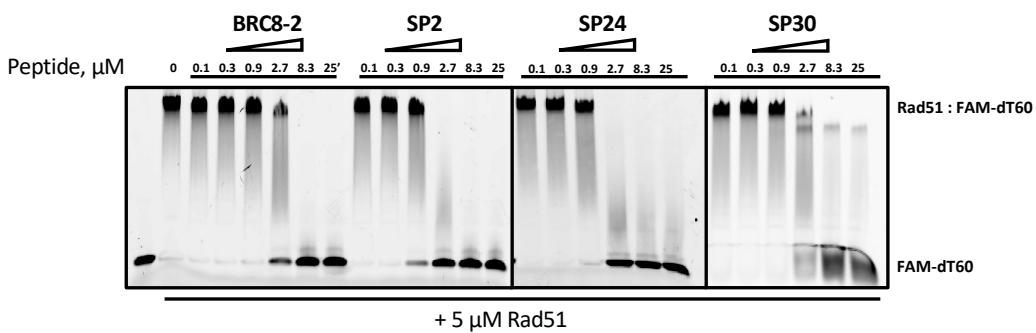
Figure 28. Competition electrophoretic mobility shift assays (EMSA) demonstrating the effect of stapled peptides on RAD51 nucleoprotein filament formation.

3.5 Proteolytic stability of stapled BRC8-2 peptides

A key objective of peptide stapling is to protect the linear sequence against proteolytic cleavage by the different proteases that are ubiquitous in the cell and physiological fluids. To evaluate this in the context of BRC8-2 peptides, I set out to investigate the effect of their

incubation with human serum. Because BRC8-2 peptides are large relative to many therapeutic peptides, traditional HPLC-based approaches may not be sensitive enough to detect cleavage of small fragments. I devised an activity-based assay instead. Briefly, 200 μM peptide was incubated with 20% human serum at 37°C and aliquots taken at various time-points and quenched with DMSO:ethanol to remove serum proteins. The quenched products were then tested by FP for their ability to inhibit HumRadA22:BRC4-fluor interaction.
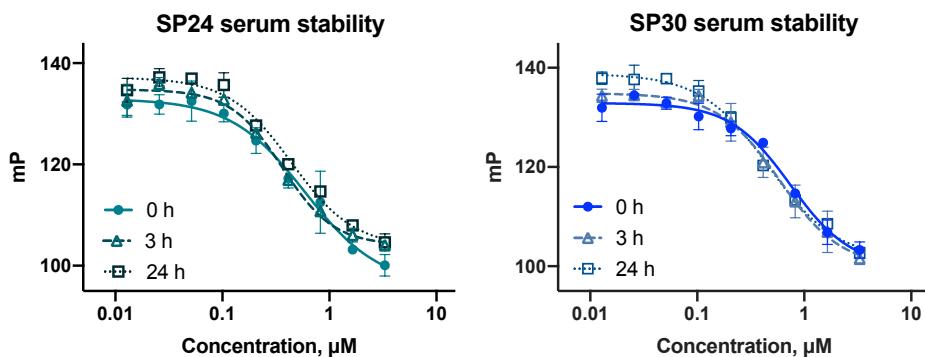
Figure 29. Activity-based serum stability assay of stapled peptides **SP24** and **SP30**. Peptides were incubated with 20% serum at 37°C, after which serum proteins were removed by DMSO:ethanol precipitation. Fluorescently-labelled BRC4 probe (10 nM) was pre-incubated with 100 nM HumRadA22, to which serum-treated peptide dilution series were added. Data shown are the means of triplicate measurements ± SD.

Remarkably, both peptides tested (**SP24** and **SP30)** had no noticeable decrease in activity even after 24 h of incubation with human serum. This could indicate either that the peptides lack labile proteolytic recognition sequences, or, if such sequences exist, that they are indeed protected by the covalent staple. Unfortunately, I was not able to evaluate the stability of unstapled control peptides with reduced cysteines, as these aggregated during the quenching step, therefore it was not possible to directly compare the effect of a covalent linkage between the cysteines. Furthermore, the cytoplasm is likely to present a more challenging proteolytic profile, and serum data cannot be simply extrapolated. Nevertheless, the fact that there was no loss of function observed after 24 h is very encouraging.

3.6 Modular functionalisation of stapled peptides

Both solid-phase and ribosomal peptide synthesis approaches have the advantage of using repetitive cycles of identical reactions to incorporate diverse building blocks with different side-chain chemistries. This makes peptidic probes suitable for modular design, where independently-developed functional elements are combined using robust and predictable orthogonal reactivities. Such modular design paradigms are finding increasing application in

biomedical sciences. For example, peptides that bind a particular target have been linked to anti-cancer cytotoxins, radioligands, proteolysis-targeting E3 ligase ligands, cell-penetrating motifs, among other functional elements. For two-component macrocyclisation reactions, the linker can provide an additional vector for such derivatisation. The linker can be developed and prepared separately from the linear peptide, allowing many functionalities to be combined with different linear peptide sequences. The divinyl linkers, for example, can be derivatised at the 5-position that is symmetrical to the two vinyl arms. Previous work using these linkers as antibody-labelling agents has shown that a cytotoxic payload can be attached via the linker to an antibody that then efficiently kills cancer cells.[156]

Despite efforts to decrease size, the RAD51-targeting stapled peptides reported above are large molecules that may not have sufficient membrane permeability. Cell-penetrating tags have been applied to a variety of cargoes, including therapeutic peptides, as a means to improve uptake.[157] These tags usually incorporate basic, cationic amino acid residues or synthetic mimics thereof. In order to improve cellular uptake, I designed **SP31**, a modified peptide based on the stapling architecture of **SP30**, with a recombinantly encoded polyarginine (Arg_9) cell-penetrating peptide (CPP) at the N-terminus (**Figure 30**). The highly charged tag was placed at the N-terminus rather than the C-terminus in order to avoid perturbing the electrostatic contacts between Gly1243 and RAD51 surface arginines.

Figure 30. Modular design of CPP-tagged and fluorescently-labelled peptides. (**A**) Sequences of peptides **SP31** and **SP32**. Both contain identical cysteine mutations and Arg9 tags but use different linkers. (**B**) Cation exchange step for the purification of linear **SP31** after TEV-cleavage and removal of the GB1 tag. (**C**) Click-compatible DVP alkyne linker used for the preparation of **SP32**. (**D**) Modification of DVP-alkyne linker after stapling to yield the fluorescent peptide **SP32**.

      SP31 was purified using a modified version of the recombinant approach described in **3.4** *Optimisation of stapled BRC8-2 peptides*. Initial attempts to use the same protocol failed, as the TEV-cleaved peptide strongly adsorbed through the polyarginine tag to the silica matrix of the C8 column in the first RPC step and could only be removed by a high-salt / GndCl wash. Instead, a cation exchange step was implemented using SP Sepharose resin that efficiently purified the cleaved linear peptide (**Figure 30**). After stapling with a DVP linker, the resulting product can be efficiently purified on a C18 RPC column.

      To aid a further derivatisation step, a click-compatible linker **DVP-Alkyne** (**Figure 30C**) was kindly prepared by Andrew Counsell from Prof David Spring's lab (Department of Chemistry). The alkyne handle can be used for addition of a variety of azide-tagged molecules.

This new linker was used to prepare **SP32**, which is identical to **SP31**, but with a fluorophore attached using copper(I)-catalysed azide-alkyne cycloaddition (CuAAC). A rhodamine-110-based green fluorophore was used instead of the more popular Alexa dyes, as it has a net neutral charge, unlike Alexa-488, which carries two negatively charged sulfonate moieties that can reduce cell uptake. The modification in **SP32** was intended for the tracking of cellular uptake and distribution of the peptide in cells by fluorescence microscopy.

3.7 Cell-based activity

At the time of submission of this dissertation, purified stapled peptides **SP24**, **SP30**, **SP31** and **SP32** have been provided to our collaborator Jessica Downs' lab for evaluation in RAD51 foci inhibition assays, analogous to what was reported with transfected BRC8-2. The resulting data will be included in future publications describing this work.

3.8 Double-stapled BRC8-2 peptides

The BRC repeats are large peptides that require the binding of both FxxA and LFDE modules for high-affinity interaction. This presents a particular challenge for intracellular targeting, as proteolytic liability is expected to increase with chain length. Moreover, larger peptide size is expected to be detrimental for membrane permeability. While stapling may mitigate these issues, it is conceivable that macrocyclisation of increasingly distant amino acids might not be sufficient. Several examples of double-stapled peptides have been reported in the literature that aim to introduce more stringent conformational constraints on the peptide through the introduction of two staple moieties, connecting two pairs of amino acid side-chains.[158]

Motivated by the success of divinyl-cysteine stapling of RAD51 binding peptides, I attempted to prepare double-stapled BRC8-2 peptides that would extend the potentially beneficial effect of macrocyclization to as large fraction of the peptide chain as possible. It was anticipated that simple addition of a reactive divinyl linker to the four-cysteine peptide would result in a complex mixture of products, as there are 3 possible linkage combinations in case of a complete reaction of such a peptide with two linkers. This makes recombinant preparation of such peptides challenging due to the need to separate these products. To this end, I tried two different methodologies. The first one involved template-assisted stapling of peptides pre-bound to the target protein. The second approach used solid-phase peptide synthesis (SPPS) of BRC8-2 four-cysteine mutants with orthogonal cysteine protecting groups. Unfortunately,

none of the approaches yielded the desirable products at sufficient purity. Nevertheless, I anticipate that with further optimisation of reagents and conditions, double-stapled BRC8-2 peptides can be prepared, and I discuss future steps to achieve this afterwards.

	FxxA module	LFDE module	Linker
BRC8-2	VNSSAFSG FSTA SGKKLNVSTEALQKAVKLFSDIENIS		
SP19	A C GFSTASGKKLNVS C EAL C KAVK LF C D		DVP
SP21	A C SGFSTASGKKLNVS C EAL C KAVK LF C D		DVP

Figure 31. Double-stapled BRC8-2 peptides that were designed in an attempt to increase the extent of conformational restraint.

3.8.1 Template-assisted double stapling

I hypothesised that pre-binding to a protein target may direct the reactive linker arms to form the desired linkage geometry, as opposed to a mixture of products. In the case of BRC8-2, I envisaged that introducing one pair of cysteines at the LFDE helix and another at the FxxA hairpin, while constraining the peptide conformation by binding to HumRadA22, which does not contain cysteines, might guide the reaction steps via physical separation of cysteine pairs. With this in mind, peptide **SP19** was designed (**Figure 31**), containing an N-terminal cysteine pair from **SP10** and a C-terminal cysteine pair derived from **SP9**, both of which maintained binding in the screening experiments.

SP19 was expressed as fusion to N-terminal GB1 expression tag and a C-terminal His-tag and purified by IMAC. Stapling reactions were performed with the **DVP** linker under pseudo-dilution conditions in an identical manner to the screening experiments in **Section 3.3**. Two reactions were done in parallel, one containing just the GB-fused peptide, and another a stoichiometric excess of HumRadA22. Despite applying a gradual titration of **DVP**, undesired three-linker adducts made up a significant fraction of the product, as detected by MS. Remarkably, adding the template appears to favour the formation of an undesired three-linker adduct. It is possible that the presence of template inhibits the intramolecular ring closing step by restricting the conformational freedom of the peptide, and thus increases the probability that the second cysteine will react with a free linker molecule. These experiments were also repeated with **SP21**, in which the first cysteine has been moved back by one position, intending to make it more accessible for ring closing. However, the same effect on the number of linkers is observed. Due to the unfavourable outcome of these reactions, the linkage species for the DVP$_2$ product were not investigated further.
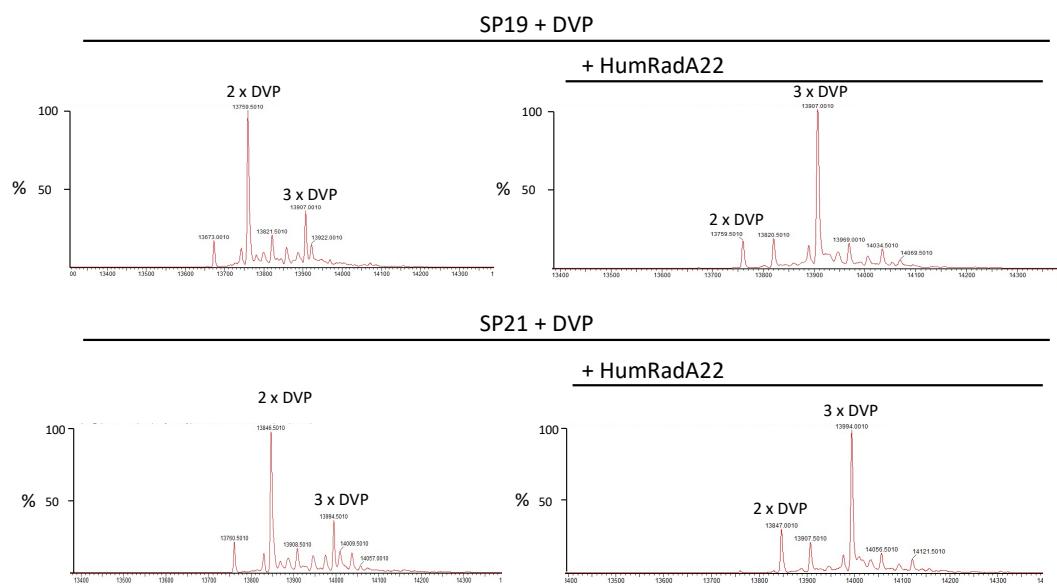
Figure 32. Template-assisted double-stapling. Mass spectra of GB1-SP19 and GB1-SP21 stapled with DVP linker alone or in the presence of a twofold stoichiometric excess of HumRadA22.

3.8.2 SPPS with orthogonal cysteine protecting groups

SPPS allows for the incorporation of a diverse set of synthetic amino acids that are not typically accessible through ribosomal synthesis. For the purpose of double-stapling, cysteine residues with orthogonal protecting groups at the thiol side-chains can be introduced into a peptide to control the macrocyclisation geometry through sequential deprotection-stapling steps. Orthogonal cysteine protecting groups have been utilised for the specific sequential conjugation of two different cytotoxic drugs to an antibody.[159] For such directed modification of BRC8-2 stapled peptides, I implemented trityl and S-tButyl, a pair of orthogonal protecting groups for thiol moieties that require different conditions for removal (**Figure 33B**). Trityl is acid-labile and eliminates during the global deprotection step with TFA. S-tButyl is stable in acids and bases but can be removed using disulfide reducing agents to yield the free cysteine thiol.
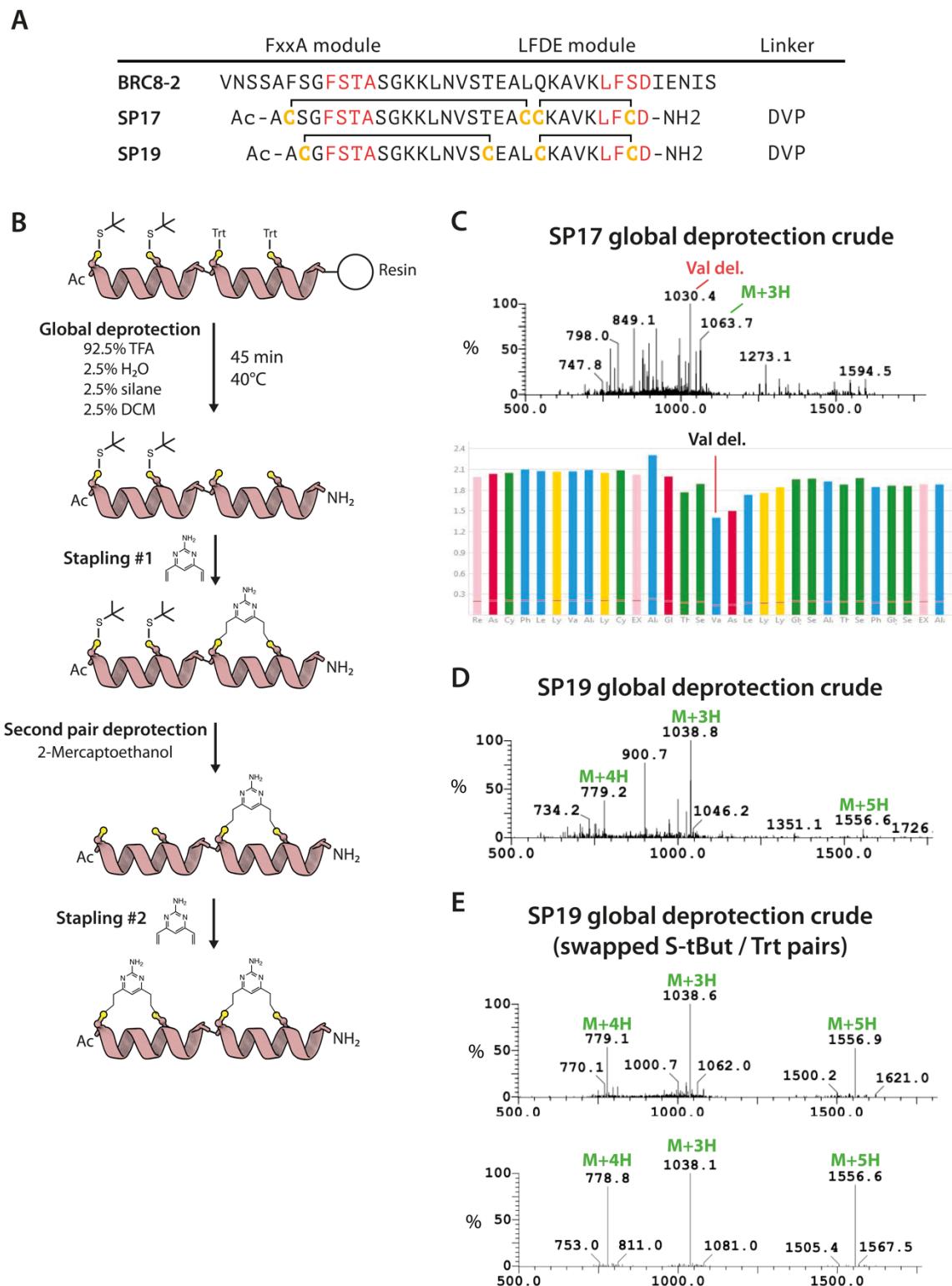
Figure 33. Solid-phase peptide synthesis of double-stapled peptides using orthogonal cysteine protecting groups. (**A**) Double-stapled peptides designed for SPPS preparation. (**B**) Reaction scheme for double-stapling of peptides using orthogonal cysteine protecting groups. (**C**) UV spectrometric quantification of coupling efficiency for each amino acid and LCMS analysis of the crude product from the global deprotection of SP17. (**D**) LCMS analysis of the crude product from the global deprotection of SP19. (**E**) LCMS analysis of the crude product (top) and RPC peak (bottom) from the global deprotection of SP19 after swapping Cys protecting groups.

SP17 was designed in an attempt to combine the stapling architectures from **SP9** and **SP12** (**Figure 33A**). Synthesis of **SP17** was attempted on an automated peptide synthesiser using standard Fmoc chemistry on RINK amide resin with amidation and acylation of C and N termini, respectively. HPLC and MS analysis of a small-scale preparation of **SP17** after TFA deprotection shows a product of extremely low yield, despite using double couplings and microwave assisted synthesis (**Figure 33C**). Peptide yield can be affected by a multitude of factors, such as on-resin aggregation, deletions due to inefficient coupling or side-reactions with the reactive cleaved protecting group moieties.[160] Synthesis failure tends to be sequence context-dependent, and hard to predict beforehand, and increases in severity with peptide length. In the case of BRC8-**SP17**, the highest-intensity peak in the crude deprotected product appears to be a Val deletion. This is further corroborated by the absorbance measurements during Fmoc-piperidine release, which shows a marked drop for the Val amino acid (**Figure 33C**).

An SPPS preparation of **SP19** was also attempted using the same conditions as for **SP17**. Small-scale global deprotection products were analysed by HPLC and MS, and reveal higher yield of the correct species, having N-terminal S-tBut protected and C-terminal deprotected Cys pairs (**Figure 33D**). In an attempt to improve on this, I decided to modify the **SP19** preparation by swapping the two pairs of orthogonal cysteine protecting groups, with the two Fmoc-Cys(S-tBut)-OH amino acids now being incorporated into the C-terminal Cys positions corresponding to the LFDE α-helix. The resulting TFA test cleavage product for the alternative preparation is more pure compared to previous syntheses, with expected m/z peaks being clearly dominant (**Figure 33E**, top). This shows that a simple modification to the SPPS protocol design can have profound effects on product quality. The RP-HPLC elution profile of **SP19** global deprotection crude product contains a number of overlapping peaks, indicating a chemically heterogeneous mixture. The fractions at the absorbance maximum, i.e. the major product, contain the correct peptide at higher purity, as evidenced by LCMS (**Figure 33E**, bottom). Encouraged by this, I attempted to scale up the purification to amounts suitable for downstream stapling reactions. Unfortunately, increasing the amount of peptide in the global deprotection step led to a massive drop in purity even when the total reaction volume was adjusted proportionally to the amount of resin used. Due to the limited ability to attend the Department of Chemistry facilities during the COVID-19 pandemic, SPPS-based preparation of **SP19** and other double-stapled peptides was not pursued further.

3.9 Conclusions

In this chapter, I have presented the successful application of a novel work-flow for the recombinant preparation and *in vitro* screening of cysteine-stapled peptides. The method can be applied for rapid and cost-effective evaluation of cysteine-stapling architectures to select lead designs for scale up and further optimisation. For this, I have combined a set of streamlined molecular biology and chemical biology methods. Sequence and ligation-independent cloning (SLIC) of peptide-coding DNA inserts circumvents the need for digesting and re-purifying the insert with endonucleases, which is typically done during restriction cloning. The use of a single *E. coli* strain (T7Express) reduces the time from cloning to protein purification, as there is no need to re-transform the vector into a separate expression strain. A single IMAC purification step allows many bis-cysteine peptides to be purified in parallel without the need for RP-HPLC steps, whereas stapling under pseudo-dilution conditions ensures quantitative conversion of the linear precursor into the correct macrocycle product. If appropriately planned, the production of these GB1-fused stapled peptides can be executed in three days from assembly PCR to the final stapling reaction (**Figure 34**). This methodology can be generalised to a variety of targets that have a suitable binding epitope, in particular, when structural data is available to guide staple design. I anticipate that it can also be expanded to different linker chemistries, as well as trivalent linkers that form bicyclic peptides. Moreover, this methodology can also be applied to evaluating sequences derived from the directed evolution of stapled peptides. For example, peptides selected by phage display may be evaluated in a similar way.
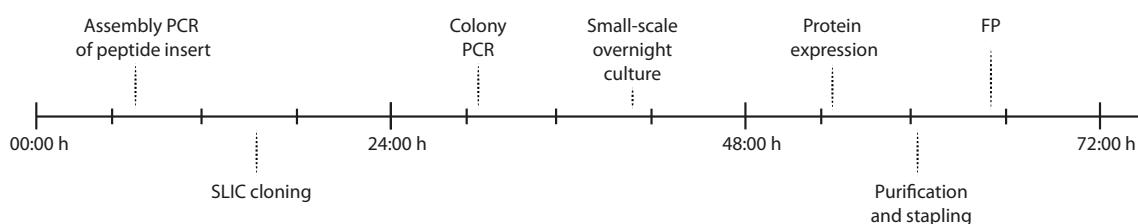
Figure 34. A best-case timeline of stapled peptide production using the small-scale recombinant work-flow.

The screening of different stapling architectures on BRC8-2 has shown that a variety of positions in the sequence can be covalently linked without disrupting binding. Crucially, both traditional helical staples and alternative architectures with large sequence distances are tolerated. For example, in the case of the **SP13**, an *i, i + 22* architecture was obtained while maintaining nanomolar affinity. It is likely that the N-terminal fragment of BRC8-2, preceding the fixed hot-spot residue Phe2058, can tolerate conformational rearrangements to accommodate the staple moiety linking the FxxA and LFDE modules.

It was not possible to significantly reduce the size of BRC8-2, as residues on both modules appear to be critical for binding, and macrocyclisation does not compensate for the loss of affinity when these critical contacts are removed. Several approaches may be used in future to improve the binding of significantly truncated BRC8-2 peptides, thus reducing the overall size of the peptidic probe. For example, **SP14**, containing the FxxA but not the LFDE motif, showed weak binding in the FP assay that appears to improve upon stapling (**Figure 25B**). This peptide may be affinity matured using a directed evolution method, such as phage or yeast display in combination with a soft randomisation mutagenesis[161] in which residues are varied but biased towards the original sequence. Alternatively, one could use error-prone PCR to introduce mutations at random positions of the **SP14** DNA insert and then select for improved binders using the aforementioned directed evolution approaches. Furthermore, rational design may be applied to introduce sequence modifications that increase binding energy. For example, in **Section 2.4** *BRC8-2 residue-specific contributions to affinity* I have shown that mutating Phe2056 to a tryptophan has a beneficial effect on binding. Synthetic fragment hits that bind RAD51 may be conjugated to the peptides, similarly to what was done in during the development of peptidomimetic CAM833.[117] Such modifications can contribute to an additive effect on affinity.

Using recombinant methods, I was able to purify stapled peptides **SP2**, **SP24** and **SP30** as free peptides in >10 mg quantities, which provided sufficient material for biophysical and biochemical studies. ITC measurements show that all three peptides bind with nanomolar affinities that are slightly inferior to the BRC8-2 template (**Figure 27B**). **SP2** appears to be the tightest binder, which is consistent with it having the least truncated template. **SP24** and **SP30** were significantly truncated and contained a number of charge-modifying mutations, yet experienced only a marginal drop in affinity compared to **SP2**.

Because the initial screening was done with a monomeric HumRadA22 archaeal surrogate protein, there was a concern that binding to this model molecule would not translate to full-length human RAD51. Gratifyingly, **SP2**, **SP24** and **SP30** peptides are able to disrupt the RAD51-ssDNA nucleoprotein filament *in vitro*, which is the desired functional consequence of these pharmacological probes.

Regrettably, I was not able to obtain double-stapled peptides with two cysteine pairs covering an increased fraction of the BRC8-2 template. Preventing the formation of mixed linkages during double stapling requires spatial or temporal separation of the two addition steps in order to obtain the desired product. In this chapter, I had optimised the initial steps necessary for an SPPS-based preparation of **SP19**, however, the synthesis was not brought to conclusion.

I anticipate that the strategy that I outline can be successfully finalised, but will require significant optimisation of purification steps, for example, through focused RP-HPLC gradients. Moreover, finding an optimal reducing agent and reaction conditions for the deprotection of S-tBut cysteines may be challenging.

Alternatively, I propose a biological approach for double-stapled peptide preparation using split inteins. Inteins are a class of proteins that catalyse their own excision out of a polypeptide sequence and re-ligate their terminal flanking regions (exteins) in a process called "protein splicing".[162] A subtype of this reaction called *trans*-splicing involves inteins that are split, usually into one significantly larger, globular fragment, and a smaller peptidic fragment, that reconstitute the full-length intein upon mixing and undergo splicing.[162] *Trans*-splicing has been applied to various biotechnological applications, such as fluorescent or isotope labelling of proteins.[162] For the purpose of preparing double stapled peptides, one can envisage that trans-splicing may be applied as a means to spatiotemporally separate the two stapling steps. By preparing the two stapled halves of a peptide separately, each containing a pair of cysteines, and then ligating them via *trans*-splicing, one could ensure tight control over the two steps. For this to be successfully applied, several caveats have to be considered, such as the promiscuity of the intein at the terminal flanking sequences (-2, -1, +1, +2 etc), which will be different for each peptide, as well as the presence of catalytic cysteines, which can also react with the electrophilic linker.

Intrigued by how the staple moiety behaves at the atomic level, I set out to determine the crystal structures of these peptides bound to HumRadA22, which are presented in the next chapter.

4 Structural studies of stapled BRC8-2 peptides

4.1 Introduction

Our ability to predict the effect of macrocyclisation on the pharmacological properties of a peptide remains poor. Structural data can elucidate some of the underlying phenomena at the atomic level and guide the design of new molecules. Only a few structures of Cys-stapled macrocycle peptide complexes have been reported, including bicyclic peptides (for example, PDB: 6Y13, 5EEQ, 5VK0, 4OS5). While it has been shown that the divinyl linker **DVT-Ac** can staple *i, i+7* architecture on α-helical peptides,[146] the conformation of the linker and its effect on the helix geometry are unclear. Screening and optimisation of stapled BRC8-2 peptides, as described in **Chapter 3**, demonstrated that a wide-range of cysteine linkage architectures were tolerated, and high-affinity binding was maintained, albeit with a drop in K_D compared to the template peptide. Importantly, stapled peptides with large inter-cysteine sequence distances and loop-like architectures were successfully prepared and applied as RAD51 inhibitors *in vitro*. Aiming to understand how the introduction of a staple moiety affects the binding mode and geometry of these peptides, I set out to determine the crystal structures of the peptides **SP2**, **SP24** and **SP30**.

4.2 Crystal structures of SP2, SP24 and SP30

Peptide:HumRadA22 complexes were screened in in 96-well MRC plate format using a variety of commercially available screens. Crystals were obtained for all three complexes, and diffraction data were collected without further optimisation of conditions. The best-diffracting crystals produced data at resolutions 3.0, 1.6 and 1.2 Å for HumRadA22 complexes of **SP2**, **SP24** and **SP30**, respectively. It is possible that the lower resolution observed for **SP2** stems from the extended flexible termini of this peptide having higher conformational heterogeneity and producing less optimal packing.

In all three complexes, electron densities are defined for peptide residues Ser2056-Asp/Gly1243, including the stapled cysteines and their linkers (**Figure 35A-C**). In the low-resolution **SP2** complex structure, the DVT-Ac linker is seen as an electron density blob located between the two cysteine sulphurs, and the conformation of this moiety is not clear at the atomic level. By contrast, in the **SP24** complex structure, which connects analogous cysteines on the C-terminal α-helix, the DVP linker can be observed in atomic detail and its conformation can be modelled accurately (**Figure 35B**). The two arms of the DVP linker in **SP24** have

acquired different orientations relative to the heterocycle core. The first arm, connected to Cys1231, has the C1-C2 bond perpendicular to the pyrimidine ring, whereas in the arm linking Cys1238, the bond is nearly co-planar. In both arms, the C1-C2 bond and the C-S bonds are in a *trans* conformation. Thus, the overall geometry of the staple minimises strain while accommodating the Cys-Cys Cα distances imposed by the helix on the two cysteines. The linker **SP24** does not interact with HumRadA22, but its core heterocycle stacks on top of endocyclic Leu1234, which is at +3 relative to the first cysteine of the peptide, creating a small hydrophobic core.

The helical geometry of the C-terminal LFDE α-helices in **SP2** and **SP24** structures is not perturbed by the introduction of the staple, with near-identical distances between cysteine α-carbons compared to BRC8-2. Moreover, alignment of isolated LFDE helices from these peptides with BRC8-2 has a backbone Cα RMSD values 0.235 and 0.310 Å, indicating minimal effect of *i,i*+7 stapling on secondary structure (**Figure 35D**). The low-strain linker conformation and unperturbed helicity observed in **SP2** and **SP24** crystal structures support the application of divinyl linkers as a general strategy for stapling α helical motifs using an *i,i*+7 architecture.

In the **SP30** structure, the DVP linker can be modelled into electron density connecting distant residues Cys2055 on the FxxA module and Cys1234 on the LFDE α-helix (**Figure 35C**). Linker arm density is clearly defined on the Cys1234 side (LFDE module), whereas Cys2055 appears more disordered (FxxA module), and fitting the residue to the electron density is more ambiguous. The linker core is situated between the side-chains of HumRadA22 Gln213 and Gln217, making surface contacts with these residues. Linker arms of **SP30** acquire a different arrangement compared to the helical architectures. Both C1-C2 bonds are perpendicular to the pyrimidine core. The C1-C2 bond at Cys1234 and C-S bond at Cys2055 are in a *gauche* conformation. This linker conformation in **SP30** appears to be necessary to accommodate the different cysteine pair orientation and shorter S-S distance, compared to the *i,i*+7 peptides.
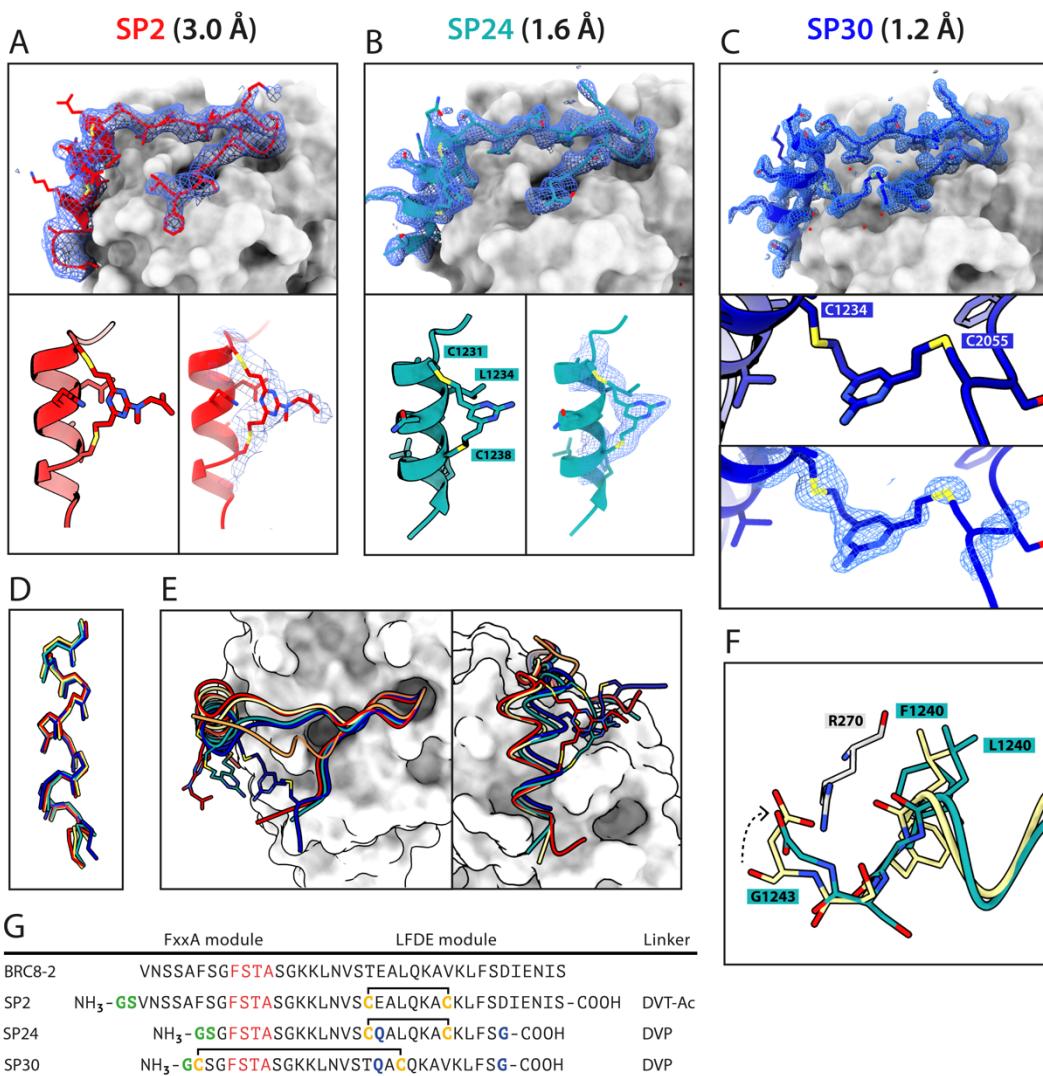
Figure 35. Crystal structures of stapled peptides **SP2** (red), **SP24** (green) and **SP30** (blue) in complex with HumRadA22 (grey surface). BRC8-2 is depicted for comparison in orange/yellow (FxxA/LFDE module). (**A**, **B** and **C**) Weighted 2mFo-DFc electron density maps after modelling of the peptide at 1σ level overlaid with the modelled peptide chains (top panels) and close-up images of staple moieties (bottom panels). (**D**) Alignment of the LFDE α-helices from the three stapled peptides and BRC8-2 based on backbone α-carbons of the helix. (E) A shift in the LFDE modules of **SP24** and **SP30** compared to **SP2** and BRC8-2. (**F**) Rotation of the LFDE α-helix in **SP24** results in a shift of Leu1240 and Phe1241 side-chains relative to their binding mode in BRC8-2. C-terminal Gly1243 forms a salt bridge with HumRadA22 Arg270 (*Hs*RAD51 Arg254), replacing an acidic side-chain. (**G**) Sequences of the crystallised stapled peptides aligned to BRC8-2.

All three stapled peptides maintain the critical FxxA and LFDE motif contacts that are conserved in BRC4 and BRC8-2. Remarkably, the LFDE modules of **SP24** and **SP30** undergo substantial movement relative to HumRadA22 and the rest of the peptide (**Figure 35E**). The observed motion affects more than a half of the peptide, encompassing residues Lys1226 to Glu1243. The movement is concomitant with a rotation of the LFDE α-helix around its helical axis, leading to the re-organisation of Leu1240 and Phe1241 side-chains in their cognate

hydrophobic clefts (**Figure 35F**). It is reasonable to hypothesise that the shift of the LFDE module is a consequence of the C-terminal Asp1243Gly mutation, which was introduced to minimise the overall negative charge of the peptide. The carboxylate terminus of the flexible glycine residue is one carbon shorter compared to an aspartate side-chain, and rotation of the LFDE helix brings it closer to Arg270 (*Hs*RAD51 Arg254) in the **SP24** structure, in order to maintain the terminal salt bridge (**Figure 35F**).

Despite containing the hairpin-mediating Ser2056, all three peptides **SP2**, **SP24** and **SP30** have lost the extended β-hairpin fold that was seen for BRC8-2 (**Figure 35E**). Instead, Ser2056 and Gly2057 residues, preceding the FxxA phenyl, point away from the peptide. This re-arrangement was anticipated in **SP30**, where the stapling architecture was designed to link the N-terminus of the peptide with the LFDE helix, whereas in **SP2** and **SP24** the cause for this movement is not immediately obvious. The cysteine mutations and the linker in **SP2** and **SP24** do not sterically preclude any of the interfaces necessary for hairpin formation. As with all comparative analyses involving backbone movements, one cannot exclude crystal-contact induced effects. In the **SP2** crystal structure, the two complexes in the asymmetric unit have non-crystallographic symmetry and their FxxA modules of the two peptides form reciprocal symmetric contacts, which may explain the different conformation. However, in the case of **SP24**, other complex molecules are not in close contact with N-terminal part of the peptide and there is a large solvent channel above it instead. It is therefore not clear why the extended β-hairpin fold has been lost for the two peptides.

4.3 Pre-organisation of stapled peptides

Previous reports of stapled peptide development have demonstrated that peptide stapling can pre-organise free peptide in solution by restricting the number of degrees of freedom towards the desired conformation. In particular, rationally introduced staples have been shown to enforce the secondary structures of α-helices, thus stabilising the bound state of a helical epitope.[163] This effect has also been hypothesised to be responsible for the improved membrane permeability of stapled peptides.

To evaluated if Cys-divinyl stapling can similarly pre-organise peptide conformation, I analysed the **SP24** and **SP30** using circular dichroism (CD) spectroscopy. For this, non-stapled versions of all three bis-cysteine peptides were prepared with reduced thiols to prevent disulfide bridge formation. Stapled peptides were also compared with the linear full-length BRC8-2 repeat.
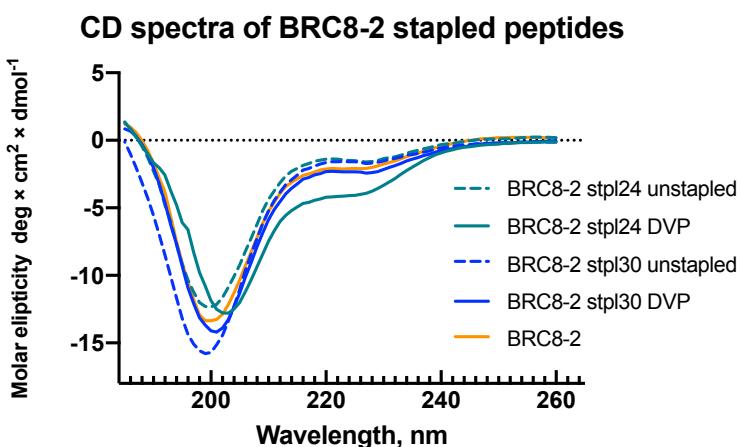
Figure 36. Circular dichroism (CD) spectra of BRC8-2 template and stapled peptides. Dashed lines indicate reduced free thiol versions of the bis-cysteine mutant. Data shown are smoothened values from three scans.

A mixed secondary structure composition is apparent from the CD spectra of all the peptides tested, with global minima around 200 nm indicative of a random coil, and a smaller local minima at 225 nm suggesting α-helical character (**Figure 36**). Covalent stapling of **SP30** has slight effect on the relative magnitudes of these minima, suggesting a small decrease in random coil, with a concurrent increase in α-helix or β-sheet (**Figure 36**, blue curves). **SP30** was designed with the intention of stabilising a large loop-like structure, part of which forms a β-hairpin. It is impossible to infer from the CD spectra of **SP30** whether the bound conformation is pre-organised.

SP24, on the other hand, shows significant increase in α-helical character upon stapling, as evidenced by an increase in negative molar ellipticity in the 210-230 nm range (**Figure 36**, green curves). **SP24** was designed with the intention to stabilise the helical secondary structure via the introduction of *i,i+7* cysteines, and the CD data confirm this effect.

4.4 Conclusions

In this chapter, I have presented the complex structures of the three BRC8-2 stapled peptides **SP2**, **SP24** and **SP30**, which have illuminated important characteristics of the staple moiety. The linker was shown to adopt a variety of conformations to accommodate different stapling architectures, while minimising steric and torsional strain. Introduction of *i,i+7* staple at the LFDE helix in **SP2** and **SP24** did not perturb helical geometry, endorsing the general applicability of this stapling chemistry towards helical epitopes. The linker orientation observed in **SP2** and **SP24** is likely to translate to other helices and can inform the rational design of binders for other targets. Based on the observed conformations, one could fine-tune the positions of cysteines to prevent steric clash between linker and target protein and

encourage favourable contacts. Moreover, in the case of **SP24**, the linker was shown to stack on top of an endocyclic residue side-chain, which adds an additional dimension for sequence optimisation. For example, introduction of an aromatic side chain at +3 position relative to the first cysteine may further stabilise the helical fold via π-stacking with the core heterocycle.

The three crystal structures validate the design strategies that were implemented in the initial design of these peptides. In the case of **SP2** and **SP24**, the cysteines were introduced in place of solvent exposed side-chains at the LFDE helix to avoid steric clash with the protein. Truncations at the C-terminus were introduced in **SP24** and **SP30**, as residues after Asp1243 appear to contribute little to binding. Confirming this, the truncated residues do not have defined electron density in the **SP2** structure, suggesting that they are disordered and not bound. Mutating the terminal Asp1243 to glycine removed an unnecessary and potentially detrimental negative charge, while preserving a salt bridge with Arg270, as shown by the complex structure. The structure of **SP30** confirms the successful linking of two distal sites on the peptide, connecting the LFDE helix with the N-terminus of the peptide without disrupting crucial contacts or secondary structure.

The structural data suggest new modifications for future optimisation of these stapled peptides. For example, one could attempt expanding the stapling architecture of **SP2** and **SP24** by introducing an additional third cysteine at the N-terminus and using a symmetric trivalent linker to create a bicyclic peptide. In fact, our collaborators at Prof David Spring's group attempted the preparation of trivinyl linkers akin to the divinyls used here, yet they were found to be chemically unstable for this application. Alternatively, one could use alkyne or azide-derivatised divinyl linkers as a means of connecting a third side-chain of a click-compatible unnatural amino acid. By introducing a second macrocycle one may further improve the proteolytic stability and other properties of the RAD51-targeting stapled peptide.

5 BRC:Rad51 interaction in *Leishmania infantum*

5.1 Introduction

Protozoan parasites are causative agents of a number of neglected tropical diseases.[120] Rad51 has been shown to have specific roles in the life cycles of a number of these organisms, making it an attractive potential therapeutic target. Selectively inhibiting Rad51 in protozoans, such as *T. brucei*, *P. falciparum* or *L. infantum*, may provide new avenues for fighting these devastating infectious diseases. This process can be accelerated by improved structural and biophysical understanding of the interactions that this protein makes. In this work, I focus on the Rad51:BRCA2 interaction axis in *L. infantum*, the causative agent of leishmaniasis, a tropical disease which in its visceral form can be fatal if untreated.

A detailed study conducted by Genois *et al*. provides the first examination of the Rad51:BRCA2 interaction in this organism.[164] *L. infantum* BRCA2 ortholog (*Li*BRCA2) was shown to be critical for maintaining genomic integrity in the protozoan, as *Li*BRCA2-null mutants displayed a decreased growth phenotype, impaired homologous recombination and chromosomal instability.[164] Moreover, in the absence of *Li*BRCA2, *L. infantum* Rad51 (*Li*Rad51) was not able to localise to the nucleus. In vitro, *Li*BRCA2 was shown to promote *Li*Rad51:ssDNA binding and D-loop formation, and inhibit *Li*Rad51:dsDNA binding. Moreover, an N-terminal fragment of *Li*BRCA2, containing the only two BRC repeats (*Li*BRC1 and *Li*BRC2), likewise stimulated ssDNA NF formation.[164]

In this chapter, I set out to investigate the *Li*BRC:*Li*Rad51 interaction in *L. infantum* to build upon the findings described above. I conduct a detailed examination of the structural determinants of *Li*BRC repeat binding and show that novel structural features define this interaction.

5.2 Sequence analysis of orthologous BRC repeats highlights potential differences in binding

Sequence alignment and comparison of conserved amino acids is a simple yet powerful approach to detect critical sequence features with functional roles. In order to evaluate the possible ways in which protozoan parasite BRC repeat binding to Rad51 might differ from those in human and other orthologs, repeat sequences from a set of representative model organisms were aligned with the BRC repeats from several important protozoan parasites.

Sequence conservation for the alignment was then quantitively evaluated using the BLOSOM62 scoring matrix (**Figure 37**). As anticipated, the Phe and Ala residues of the FxxA motif are almost universally conserved and are amongst the highest-scoring residues. Interestingly, a number of repeats contain a glycine instead of an alanine at the FxxA domain in several evolutionarily distant species, such as *Ustilago maydis*, where it is found on the single BRC repeat, and *L. infantum* BRC2. Another highly conserved feature is the SGK/R motif that forms the β-hairpin of the FxxA module, as observed in the structures of HsBRC4 and HsBRC8-2 (positions 13-15 in **Figure 37**). In particular, the high conservation of glycine is indicative that the β-turn is maintained across highly divergent organisms as a universal structural feature. C-terminal from the β-turn, a ΦxΦ motif can be identified, consisting of two conserved hydrophobic residues (Φ) separated by a non-conserved, solvent-exposed residue (positions 17-19 in **Figure 37**). Although the two hydrophobic residues are not buried within pockets like in the FxxA and LFDE motifs, their high conservation suggests that the continuous interface formed by these aliphatic side chains is crucial for the folding and binding of repeat. Unexpectedly, at the LFDE motif, the first residue (position 30 in **Figure 37**) appears to be rather poorly conserved. Instead, closer examination of sequences reveals that the protozoan organisms included in this analysis have charged or polar residues at this position, whereas in the so-called higher eukaryotes it is always occupied by a hydrophobic amino acid.
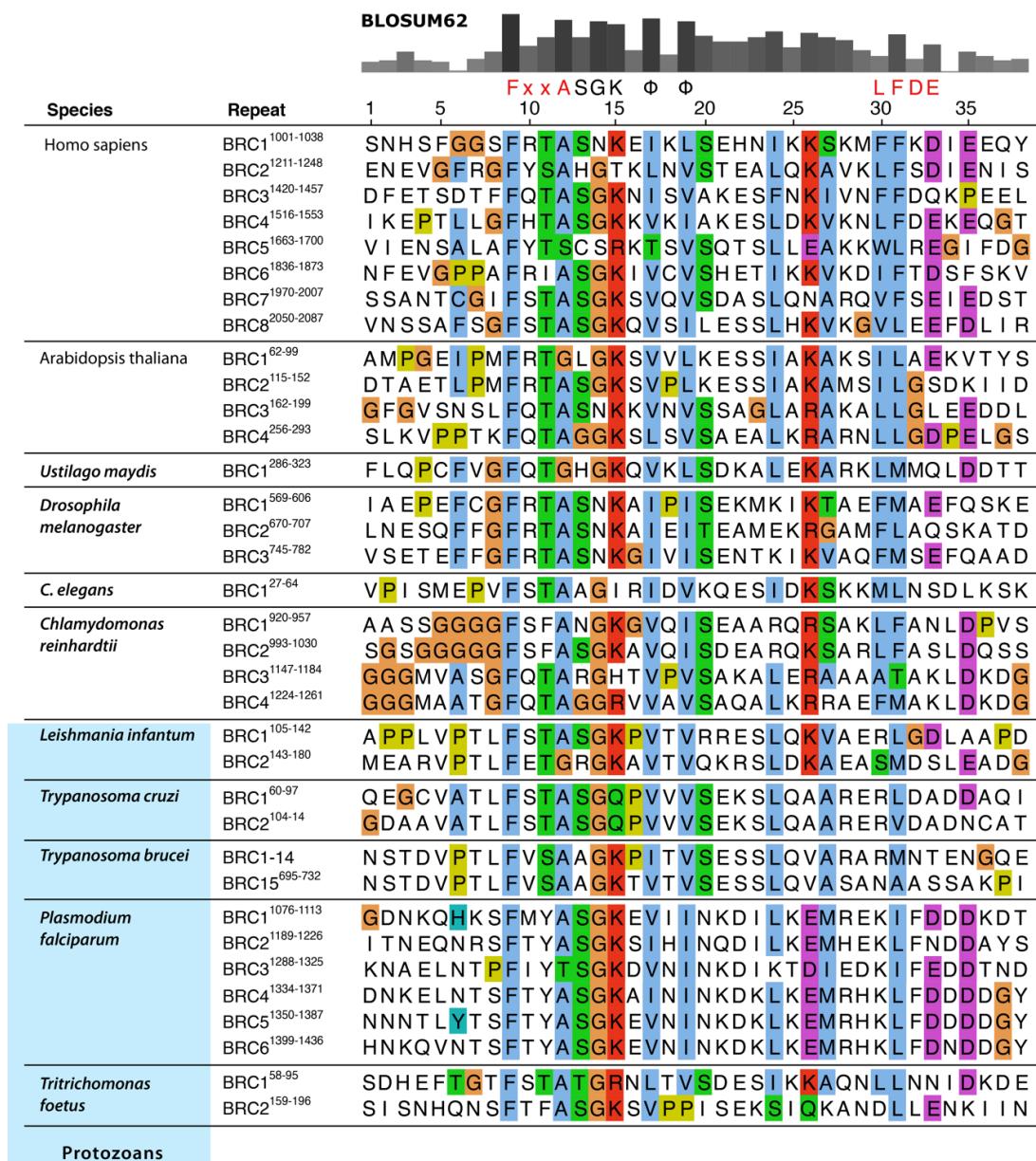
Figure 37. Sequence alignment of BRC repeats from a set of representative organisms and protozoan parasites. The grey bars on top represent BLOSUM62 alignment scores. Conserved residues are coloured using the default ClustalX colour scheme.

5.3 Purification of the *L. infantum* Rad51 ATPase domain

To investigate the interactions between the two *L. infantum* BRC repeats and *Li*Rad51, a monomeric version of *Li*Rad51 was necessary, as the oligomeric full-length protein is not suitable for biophysical and crystallographic studies of BRC repeat binding due to competition with self-oligomerisation and conformational heterogeneity. Purification of the ATPase domain of *Li*Rad51 lacking the OE and the NTD was therefore attempted. The sequence corresponding to *Li*Rad51 residues 134-376 (*Li*Rad51[ATPase], Uniprot: A4I3C9_LEIIN, **Figure 38A**) was codon-optimised for expression in *E.coli,* synthesised and cloned into pHAT2

bacterial expression vector containing an N-terminal His$_8$-tag. Small-scale expression tests showed that the protein can be produced in soluble form when expression is induced overnight at 15 °C. Scaling up to a larger culture volume and IMAC with Ni-NTA resin resulted in high amounts of purified protein, however, the product formed a viscous solution immediately after elution, followed by precipitation of insoluble aggregates.
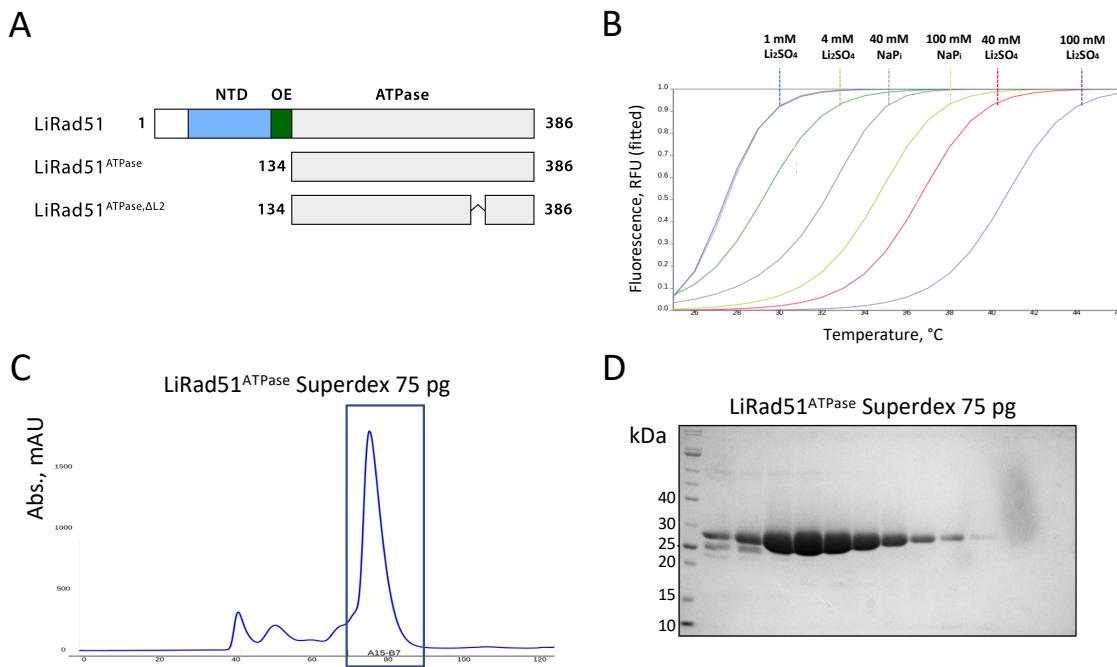
Figure 38. Purification of monomeric *Li*Rad51 ATPase domain. (**A**) Schematic drawing of domain organisation of full-length *Li*Rad51 and *Li*Rad51^{ATPase}. An N-terminal putative disordered region is depicted in white. (**B**) Differential scanning fluorimetry of *Li*Rad51^{ATPase} with sulphate and phosphate additives. All reactions were done with IMAC-eluted *Li*Rad51^{ATPase} in 50 mM Tris pH8.0, 150 mM NaCl. (**C**) Size-exclusion chromatography trace with *Li*Rad51^{ATPase} monomeric peak eluting at around 75 ml. (**D**) Coomassie-stained SDS-PAGE gel of fractions from the run shown in (C).

To improve solubility, I carried out a screen of stabilising buffer components using differential scanning fluorimetry (DSF). A 96-well format standardised screen available in our lab was used that covers a range of pH values, buffers, salts and additives that can improve the stability of the protein in solution. The fitted ΔT_m values suggested that the monomeric *Li*Rad51^{ATPase} is stabilised by phosphate and sulphate anions. In a subsequent buffer optimisation, lithium sulphate exerted a very large thermal shift in a concentration-dependent manner (**Figure 38B**). Crystallographic studies have shown that free inorganic phosphate can bind the Rad51 P-loop in place of nucleotide phosphates. It is possible that both phosphate and the similarly tetrahedral sulphate anions exert their stabilising effect through this interaction. Alternatively, it is also possible that the anions stabilise the protein surface elsewhere. Informed

by this result, sulphate was included in the purification, which, gratifyingly, enabled the production of stable LiRad51$^{\text{ATPase}}$ in >10 mg quantities (**Figure 38C,D**).

5.4 LiBRC1 binds LiRad51 more strongly than LiBRC2

To qualitatively evaluate the interaction between the two *L. infantum* BRC repeats and LiRad51, an affinity co-precipitation assay was done using proteins expressed in *E. coli*. The two BRC repeats were cloned into the pPEPT1 expression vector, harbouring an N-terminal StrepTag-GB1 fusion and a C-terminal His$_8$-tag. All GB1 fusion constructs were expressed in 10 ml *E. coli* cultures and their bacterial lysates were loaded on Ni-NTA resin, followed by application of similarly prepared LiRad51$^{\text{ATPase}}$ lysate. Imidazole-eluted fractions were analysed by SDS-PAGE (**Figure 39B**). It is evident that LiBRC1 pulls down LiRad51$^{\text{ATPase}}$ in stoichiometric amounts, suggesting a strong interaction, whereas only trace amounts of LiRad51$^{\text{ATPase}}$ are seen in the LiBRC2 pull-down.

To enable the routine measurement of affinities for ligands binding at the LiRad51 ATPase domain, I developed a fluorescence polarisation (FP) assay based on the LiBRC1-LiRad51 interaction. For this, I prepared a fluorescently-labelled LiBRC1 peptide by conjugating a fluorescein fluorophore using cysteine-maleimide chemistry. Direct FP titration of LiRad51$^{\text{ATPase}}$ into this probe gave a K_D of 0.165 μM (**Figure 39C**). Competition experiments using purified GB1 fusion constructs were then set up, giving K_D values of 0.29 μM and 13.55 μM for GB1-LiBRC1 and GB1-LiBRC2, respectively (**Figure 39D**). The competition FP measurements together with the pull-down data strongly indicate that LiBRC1 is a higher affinity repeat, and LiBRC2 a much weaker binder. Moreover, the affinity for LiBRC1 was significantly lower than the nanomolar K_D values reported for high affinity human BRC repeats. To exclude fusion partner-induced effects, I also prepared LiBRC1 as a free peptide and measured its binding to LiRad51 using ITC, which gave a K_D of 0.65 μM, excluding the possibility that the GB1 tag is detrimental to binding.

A

	FxxA module	LFDE module	FP competition	
			$IC_{50} \pm SE$ (μM)	K_D (μM)
*Hs*BRC4[1519-1551]	PTLLG**FHTA**SGKKVKIAKES**L**DKVKN**LFDE**KEQ			
*Li*BRC1[108-140]	LVPTL**FSTA**SGKPVTVRRES**L**QKVAE**RLGD**LAA		1.57 ± 0.08	0.29
*Li*BRC1.1	LVPTL**FST**G**SGKPVTVRRES**L**QKVAE**RLGD**LAA		4.19 ± 0.19	0.94
*Li*BRC2[146-178]	RVPTL**FETG**RGKAVTVQKRS**L**DKAEA**SMDS**LEA		54.96 ± 3.30	13.55
*Li*BRC2.1	RVPTL**FET**A**RGKAVTVQKRS**L**DKAEA**SMDS**LEA		43.05 ± 2.36	10.59

B

C

D

E

Figure 39. *Li*BRC1 is a more potent binder of *Li*Rad51 than *Li*BRC2. (**A**) Sequence alignment of the two *L. infantum* BRC repeats with *Hs*BRC4, and point mutants *Li*BRC1.1 and *Li*BRC2.1. Competition FP binding results are shown on the left. (**B**) Coomassie stained SDS-PAGE gel analysis of *L. infantum* BRC repeat affinity pull-down of LiRad51[ATPase] (**C**) Direct FP titration of *Li*Rad51[ATPase] into fluorescein-tagged *Li*BRC1 (5 nM). Data shown are the means of triplicate measurements ± SD. (**D**) Competition FP titrations of GB1-fused *Li*BRC1 and *Li*BRC2. Fluorescently-labelled LiBRC1 probe (5 nM) was pre-incubated with 500 nM *Li*Rad51[ATPase], to which GB1-LiBRC dilution series were added. Data shown are the means of triplicate measurements ± SD. (**E**) ITC binding measurement of *Li*BRC1 peptide binding to LiRad51[ATPase].

With an K_D of 13.55 μM, *Li*BRC2 is a much weaker binder than *Li*BRC1. Previous studies on human repeats have shown that BRC5 is the lowest affinity repeat within BRCA2, which has been rationalised by the mutation of an alanine to a serine in the FxxA motif of BRC5. I hypothesised that the same effect may be responsible for the decreased affinity in *Li*BRC2, which harbours a glycine instead of the alanine at the equivalent position. To test this,

a rescue mutant was prepared with the alanine re-introduced (*Li*BRC2.1, **Figure 39A**). However, this did not bring about any significant increase in affinity, resulting in a similar K_D of 10.39 μM. This lack of effect prompted me to further investigate the role of the FxxA alanine in binding, and an FxxG mutant of *Li*BRC1 was prepared (*Li*BRC1.1, **Figure 39A**). A threefold decrease in K_D was observed for this mutant, suggesting that, while the methyl group of the FxxA alanine contributes to binding, it is not essential in some contexts. It is thus reasonable to suggest that other factors besides the loss of alanine from the FxxA motif are responsible for the low affinity of *Li*BRC2. For example, *Li*BRC1 Ser117 is likely involved in the stabilisation of the β-turn at the FxxA module, analogously to *Hs*BRC4 and *Hs*BRC8-2, and the equivalent position corresponds to an arginine in *Li*BRC2, which would not be able to mediate such an intra-molecular contact.

5.5 *Li*BRC1 binding studied by X-ray crystallography

To date, there are only two BRC:Rad51 complex structures available in the public domain: the BRC4 complex reported by Pellegrini *et al.*[80] and the BRC8-2 structure presented in **Section 2.3** *X-ray crystallographic study of BRC8-2*. Only a handful of reports investigate BRC repeat binding in organisms other than humans and no structural or biophysical studies have been presented. Motivated by the potential for targeting the BRCA2:Rad51 axis in protozoan diseases, I set out to determine the *Li*BRC1:*Li*Rad51 complex structure by X-ray crystallography.

To reduce the flexibility and conformational heterogeneity of the complex, I designed a deletion mutant of the *Li*Rad51 ATPase domain by removing the DNA-binding loop L2, which in the absence of DNA is typically disordered (*Li*Rad51$^{ATPase,\Delta L2}$, **Figure 38A**). The L2 loop is not involved in the binding of BRC repeats by Rad51, as it is located distal to the interaction interface. The *Li*BRC1:*Li*Rad51 complex was purified from *E. coli* by IMAC co-precipitation, in an analogous manner to the HsBRC8-2 complex presented in **Section 2.3**. Crystallisation trials were performed in 96-well MRC plate format using a variety of commercial sparse-matrix screens and varying conditions such as protein:precipitant volumes and addition of ADP/Mg^{2+} cofactors. A single crystal was obtained, for which a full diffraction dataset was collected to 2.15 Å resolution. Initial phase determination of the structure factors was done using molecular replacement with *Pyrococcus furiosus* RadA ATPase domain as a search model (PDB: 4A6P). After several rounds of refinement, positive electron density

difference was clearly observed for the *Li*BRC1 peptide and the binding mode could be determined (**Figure 40A**).

The overall fold of the *Li*BRC1 peptide is similar to what has been reported for *Hs*BRC4 (**Figure 40B**). The Phe113 and Ala116 hot-spot residues bind the two hydrophobic FxxA site pockets on the ATPase domain in an identical manner to BRC4. The β-turn, mediated by 115-TASGK-119, is also preserved and closely resembles the human BRC4 in its hydrogen-bonding pattern, with a Thr115 side-chain stabilising the turn through hydrogen bonding. Remarkably, an extended β-hairpin fold, akin to the one determined for the BRC8-2 repeat (**Section 2.3**), is also formed by the *Li*BRC1 peptide (**Figure 40B**). A Thr111 residue, positioned at -2 to the FxxA motif, mediates a hydrogen bonding network that allows the peptide to fold back on itself, similarly to Ser2056 in the *Hs*BRC8-2 structure (**Figure 40C**). The Thr111 side-chain hydroxyl forms stabilising hydrogen bonds with the Phe113 and Val121 backbone amides from the two anti-parallel strands of the hairpin. The methyl group of Thr111 side-chain is accommodated by a less bulky Val123 in *Li*BRC1, compared to BRC4 Ile1534 (**Figure 40C**). Remarkably, the extended β-hairpin fold promotes the *Li*BRC1 peptide to fold into a small hydrophobic core mediated by the side-chains of *Li*BRC1 Val109, Thr111, Val123 and Leu128, as well as *Li*Rad51 surface residues Leu241, Gln242, Ala245, Met246 (**Figure 40C**). Leu112, which precedes Phe113, forms additional hydrophobic contacts with a hydrophobic cleft formed by *Li*Rad51 His235, Leu239 and Gln242. The sum of these observations suggest that *Li*BRC1 forms considerably more hydrophobic interactions at the FxxA site compared to the previously elucidated human repeats.
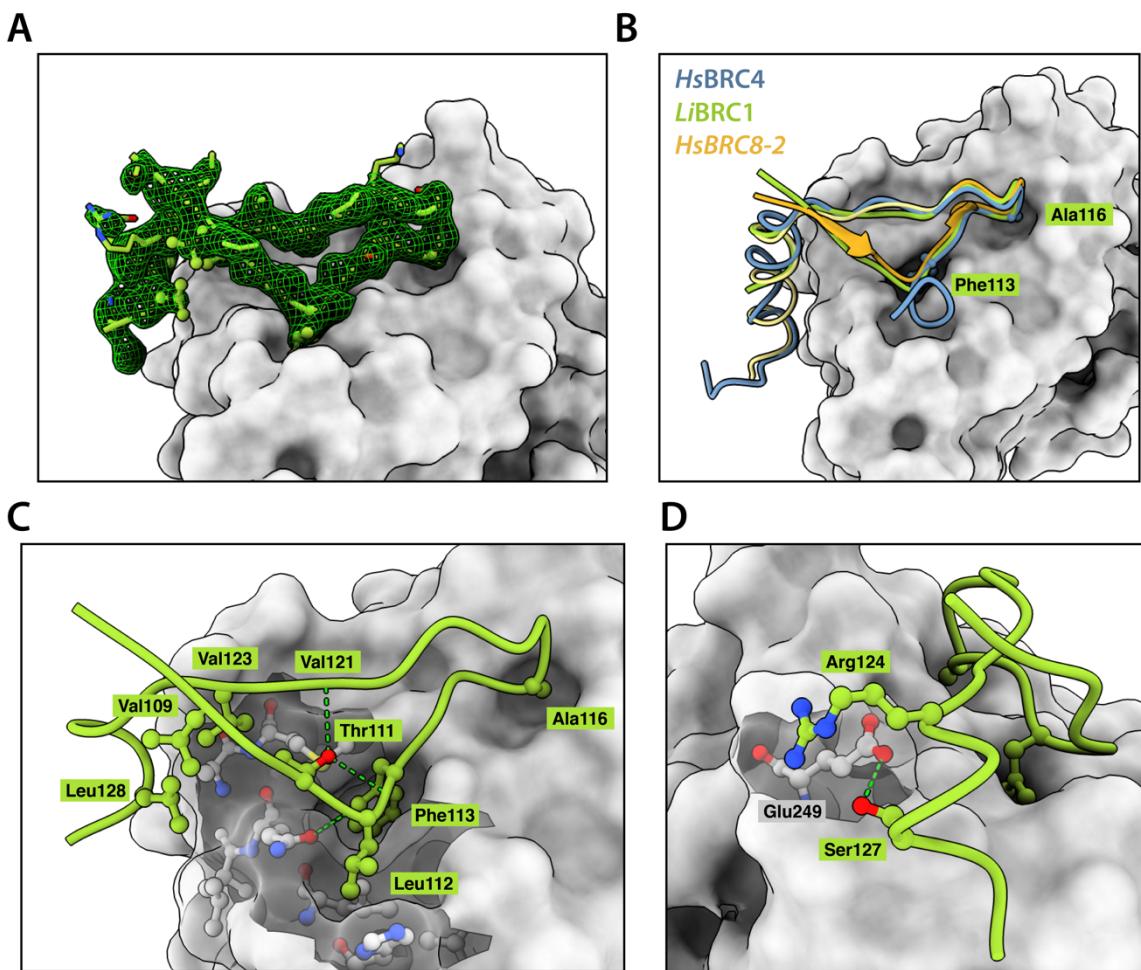
Figure 40. Crystal structure of *Li*BRC1 in complex with *Li*Rad51$^{\text{ATPase}}$. (**A**) *Li*BRC1 peptide 2Fo-Fc electron density map at 1σ (**B**) Overall binding mode of *Li*BRC1 superposed with HsBRC4 and *Hs*BRC8-2. (**C**) Detailed view of the extended β-hairpin fold formed by *Li*BRC1 (**D**) Arg124-mediated electrostatic contacts with *Li*Rad51, further stabilised by a hydrogen bond with Ser127.

The α-helix of the *Li*BRC1 LFDE module starts with a cationic Arg124 residue that forms electrostatic contacts with *Li*Rad51 Glu249 side-chain (**Figure 40D**). A similar interaction has not been observed for *Hs*BRC4 or *Hs*BRC8-2, despite *Hs*RAD51 also containing a glutamate at the equivalent position. Ser127, which is one helical turn away from Arg124, also interacts with Glu249, in a manner that resembles Ser1538 from *Hs*BRC4.

Remarkably, the *Li*BRC1 LFDE module has no electron density defined beyond residue Gln129. This observation was not anticipated, as previous work has stressed the importance of the C-terminal LFDE module for BRC repeat binding.[81] In particular, the hot-spot residues of the LFDE motif, corresponding to RLGD in *Li*BRC1, have been shown to be critical for the binding of BRC4 *in vitro* and the function of BRCA2 in cells,[81] yet, in the case of LiBRC1, these residues have no discernible electron density even after the rest of the peptide has been modelled and several rounds of refinement done. In a crystallographic system there is always a possibility that the binding mode is distorted by lattice packing artefacts, for example, to

accommodate steric clashes with nearby molecules. Upon closer examination of the packing environment around the LFDE interface, a sufficiently large solvent channel is formed around the expected LFDE-binding region was observed and no unmodelled electron density seen elsewhere, suggesting that the peptide C-terminus is disordered. To ensure that the peptide was not degraded by bacterial proteases during purification, the complex was analysed by protein LCMS, and the full-length species were identified (**Appendix 8.3** *Protein LCMS spectrum of LiBRC1:LiRad51 complex*).
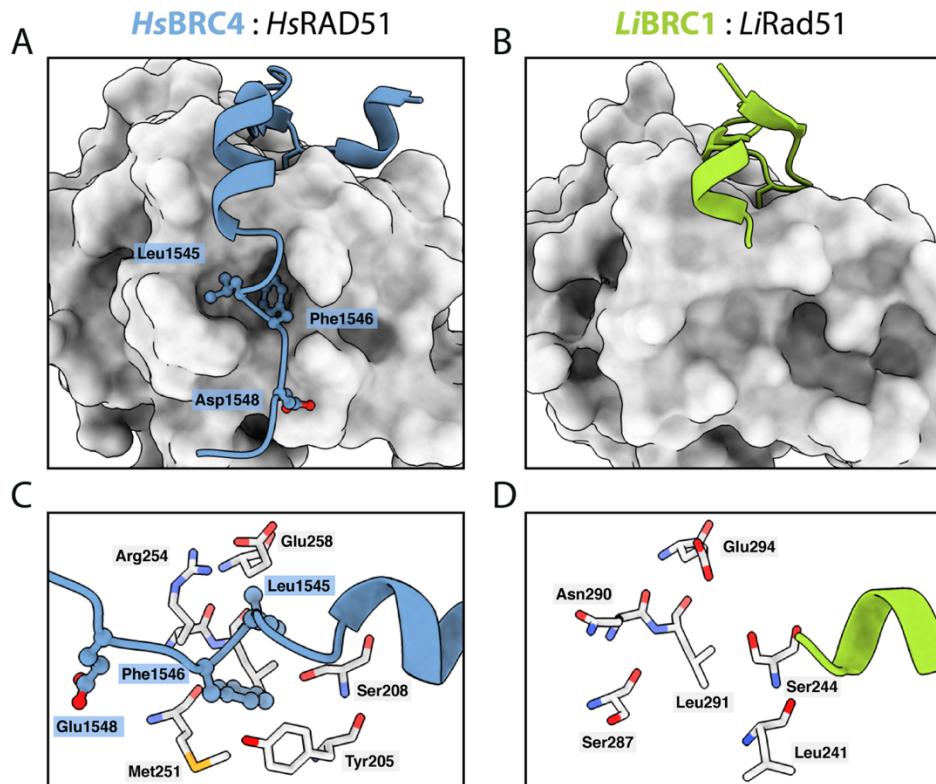
Figure 41. LFDE motif binding is not observed in the *Li*BRC1:*Li*Rad51 crystal structure. (**A,B**) Comparison of the binding mode of C-terminal LFDE motif residues in (**A**) HsBRC4 and (**B**) *Li*BRC1. The difference in hydrophobic pocket depth is clearly apparent. (**C,D**) Comparison of the residues involved in the formation of the LFDE-binding cognate hydrophobic pockets in (**C**) *Hs*RAD51 and (**D**) *Li*Rad51.

The crystal structure suggests a binding mode for *Li*BRC1 in which the LFDE motif does not form critical contacts with the ATPase domain. Comparison of the LiRad51$^{\text{ATPase}}$ surface at the LFDE interface with that of the human RAD51 reveals structural features that further corroborate this hypothesis (**Figure 41**). The *Li*Rad51 surface region corresponding to the pocket on *Hs*RAD51 where BRC4 Leu1545 and Phe1546 bind, presents a significantly more shallow cavity compared to *Hs*RAD51(**Figure 41A,B**). The different topology results from several mutations in the *Li*Rad51 protein relative to *Hs*RAD51. A Tyr205 in the human protein stacks on top of the Phe1548 side-chain, creating a tight pocket for the phenyl group.

This is mutated to Leu241 in *Li*Rad51, a side-chain that is unable to provide equivalent buried area (**Figure 41C,D**). Instead, Leu241 interacts with the hydrophobic core described above, formed by the more N-terminal *Li*BRC1 residues. Similarly, Met251 in *Hs*RAD51 corresponds to Ser287 in LiRad51, leading to a less pronounced hydrophobic cavity at the LFDE site.

In sum, the crystallographic structure of the *Li*BRC1:*Li*Rad51 complex reveals a BRC repeat binding-mode defined by the following novel structural features: an extended β-hairpin at the N-terminus of FxxA module, a small hydrophobic core formed by a number of sequence-distant residues, and a lack of any interaction with *Li*Rad51 at the C-terminal LFDE motif.

5.6 C-terminal *Li*BRC1 residues are not critical for binding

The crystallographic model prompted me to further investigate the importance of C-terminal *Li*BRC1 residues for binding. For this, a set of *Li*BRC1 mutant constructs were designed and purified as fusions to an N-terminal GB1-tag and a C-terminal His_8-tag (**Figure 42**). These were tested in the FP competition assay described above. Two mutants with extended N- or C- termini containing additional residues from the full length *Li*BRCA2 protein were first evaluated to account for the possibility that the initial choice of cut-off failed to capture enough of the repeat sequence (*Li*BRC1.2 and 1.3). Little change in affinity was observed for these longer repeats, which validates the initially delineated repeat boundaries.

	FxxA module	LFDE module	FP competition	
			$IC_{50} \pm SE$ (µM)	K_D (µM)
*Hs*BRC4	PTLLGFHTASGKKVKIAKESLDKVKNLFDEKEQ			
*Li*BRC1	LVPTLFSTASGKPVTVRRESLQKVAERLGDLAA		1.57 ± 0.08	0.29
*Li*BRC1.2	LVPTLFSTASGKPVTVRRESLQKVAERLGDLAAPDME		1.20 ± 0.06	0.19
*Li*BRC1.3	AAAPPLVPTLFSTASGKPVTVRRESLQKVAERLGDLAA		1.67 ± 0.07	0.31
*Li*BRC1.4	LVPTLFSTASGKPVTVRRESLQKVAERLGD		2.08 ± 0.10	0.41
*Li*BRC1.5	LVPTLFSTASGKPVTVRRESLQKVAERLG		1.94 ± 0.12	0.38
*Li*BRC1.6	LVPTLFSTASGKPVTVRRESLQKVAER		4.80 ± 0.34	1.09
*Li*BRC1.7	LVPTLFSTASGKPVTVRRESLQKVAE		1.99 ± 0.08	0.39
*Li*BRC1.8	LVPTLFSTASGKPVTVRRESLQKVA		8.80 ± 0.65	2.09
*Li*BRC1.9	LVPTLFSTASGKPVTVRRESLQKV		11.01 ± 0.64	2.63
*Li*BRC1.10	LVPTLFSTASGKPVTVRRESLQK		40.99 ± 10.39	10.08
*Li*BRC1.11	LVPTLFSTASGKPVTVRRESLQ		36.46 ± 4.99	8.70
*Li*BRC1.12	LVPTLFSTASGKPVTVR		>500	>124

Figure 42. *L. infantum* BRC repeat constructs were purified and tested for LiRad51[ATPase] binding in an FP competition assay. Columns on the left show fitted FP assay IC_{50} values ± SE of fit and calculated K_D values.

I then proceeded with a step-wise deletion approach, in which 1-2 residue fragments were sequentially removed from the C-terminus to map the binding contributions of the LFDE module, a significant part of which lacks any electron density in the crystal structure of the

complex. Interestingly, C-terminal deletions were tolerated without significant loss of affinity up to, and including, removal of Arg134 (*Li*BRC1.7, **Figure 42**). This implies that the 134-RLGD-137 tetrad, whose residue positions correspond to the canonical LFDE motif in humans, is not critical for the binding of *L. infantum* BRC1 repeat to the *Li*Rad51 ATPase domain.

The most C-terminal residue observed in the crystal structure is Gln129. Further truncations up until this residue result in a gradual decrease in affinity, reaching a K_D of 8.7 µM for LiBRC1.11, signifying an important contribution to binding by the 130-KVAE-133 tetrad, which is likewise not seen in the crystal structure. Removal of *Li*BRC1 Glu133 causes a more than four-fold drop in K_D, implying that this residue makes a significant contribution to binding (*Li*BRC1.8, **Figure 42**). It is not immediately apparent from the complex structure how this residue may contribute to binding, as there are no nearby *Li*Rad51 side-chains bearing a positive charge to form salt bridges with. It is possible that, rather than interacting with *Li*Rad51, it stabilises the repeat conformation, for example, by interacting with the cationic Lys130 on the same helical face or by stabilising the overall charge of the peptide.

Bulk removal of the remaining helical residues 125-RESLQ-129 diminished binding even further, resulting in only residual displacement of the FP probe at the highest ligand concentrations in the assay. It is likely that this drop is to a large extent caused by the loss of of Leu128, which contributes to the hydrophobic core that is formed at the edge between the β-hairpin and the α-helix of *Li*BRC1.

To confirm these sequence-activity relationships in the context of the full-length *Li*Rad51 protein, I performed electrophoretic mobility shift assays (EMSA), in which *Li*BRC1 and its truncation constructs *Li*BRC1.7, 1.9 and 1.11 were tested for their ability to inhibit the formation of *Li*Rad51-ssDNA nucleoprotein filament. Both *Li*BRC1 and *Li*BRC1.8 inhibited NF formation in a dose-dependent manner to comparable levels (**Figure 43**, top). In line with the FP measurements, *Li*BRC1.10 and 1.12 were much less potent inhibitors of NF formation (**Figure 43**, bottom).
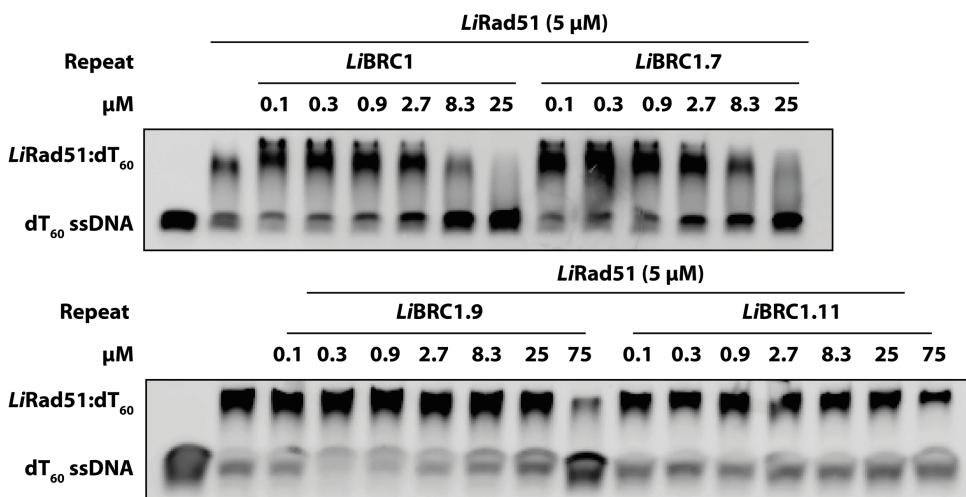
Figure 43. EMSA competition assay evaluating the ability of *Li*BRC1 truncation mutants to inhibit LiRad51 nucleoprotein filament formation. 5 μM *Li*Rad51 was pre-incubated with GB1-fused *Li*BRC1 or its truncation mutants, after which FAM-labelled ssDNA (dT$_{60}$, 100 nM) was added to the reaction. Products were resolved on a 1xTBE native 2% agarose gel.

5.7 Discussion

In this chapter, I have shown that the two BRC repeats from the *L. infantum* BRCA2 ortholog bind the *Li*Rad51 ATPase domain, with *Li*BRC1 manifesting an almost 50 times higher affinity than *Li*BRC2, as measured by an FP competition assay. This represents the first instance where the BRC repeat:Rad51 interaction is confirmed for a non-human BRCA2 ortholog. Depending on the assay, *Li*BRC1 bound *Li*Rad51^{ATPase} with K$_D$ in the range of 165-650 nM, which is at least several times weaker than what has been previously reported for the tightest-binding human repeat BRC4. A crystal structure of the *Li*BRC1:*Li*Rad51^{ATPase} complex was determined and revealed that *Li*BRC1 residues 134-RLGD-137, corresponding to the conserved LFDE motif in humans, do not form ordered contacts with *Li*Rad51^{ATPase}. Subsequent truncation mutagenesis experiments confirmed that, indeed, this C-terminal region of *Li*BRC1 is not critical for binding.

Comparative analysis with the human proteins can help rationalise the lack of interaction observed for the 134-RLGD-137 tetrad from the point of view of the repeat sequence, or, alternatively, by looking at the complementary surfaces formed by the ATPase domains. The human BRC repeat LFDE motifs are defined by two strongly conserved features. First, two bulky hydrophobic residues, such as Trp, Leu, Phe and Val, are conserved at tthe first two positions of the motif in all the human repeats. Secondly, an acidic residue at the last position forms a salt-bridge with Arg250 or other nearby arginines on human RAD51. The

shape of the first two side-chains appears to be less critical than their hydrophobic nature, as evidenced by the different combinations seen in the eight human BRC repeats. Moreover, Rajendra and Venkitaraman showed that exchange of one hydrophobic residue for another is not disruptive for binding, and can in fact bring about improved affinity, such as when Leu1545 is replaced by a tryptophan in *Hs*BRC4.[81] In *Li*BRC1, on the other hand, there is just a single hydrophobic residue present in the tetrad that corresponds to the LFDE motif, which drastically reduces the buried hydrophobic surface area attainable upon binding. Buried hydrophobic contacts tend to contribute significantly to free energy of binding in protein-protein interactions, therefore it is reasonable to assume that the RLGD tetrad would result in a weaker energetic contribution, even if compensatory contacts, for example, a salt bridge involving Arg134, were present. The cognate surface of the *Li*Rad51 ATPase domain also appears less conducive to the binding of an LFDE-like moiety, as the hydrophobic pockets are less pronounced.

Remarkably, the repeat manifests nanomolar binding despite lacking the LFDE-like interaction. Unlike *Hs*BRC4, the N-terminus of *Li*BRC1 peptide forms an extended β-hairpin, which results in a hydrophobic core folding on top of the ATPase domain surface, as well as hydrophobic contacts formed by a nearby Leu112. I propose that these additional hydrophobic interactions partially compensate for the lack of a functional LFDE motif and ensure a high affinity interaction for the *L. infantum* BRCA2 ortholog to localise Rad51 to the sites of DNA damage and stimulate nucleofilament formation on resected ssDNA. The case of *Li*BRC1 therefore presents a distinct mode of BRC repeat binding for the evolutionary distant *L. infantum*. Interestingly, the sequence distance between the two BRC repeats in the *L. infantum* BRCA2 ortholog is much smaller than in human BRCA2, that is, ~6 residues, depending on where repeat boundaries are defined. This means that, in order for the protein to engage more than one Rad51 molecule simultaneously, as has been previously shown for BRCA2, the C terminus of the repeat may need to be vacant, explaining the lack of interaction for the 134-RLGD-137 tetrad.

6 Conclusions and future directions

6.1 Structural insights into the RAD51:BRCA2 axis

Mechanistic details of how BRC repeats exert their function in the context of the full-length BRCA2 protein are not clear, but the interaction has been described for isolated repeats. In evolving eight anciently divergent BRC repeats in mammals, Nature has provided a pool of sequences with varying binding properties. In work leading to this dissertation, systematic shuffling of these sequences revealed module-specific contributions to affinity and yielded chimeric repeats with improved binding affinities over any of the natural sequences.

Intrigued by these findings, in this work I have elucidated novel structural features that define the binding of BRC repeats, advancing our understanding of homologous recombination. In the crystal structure of BRC8-2, an extended β-hairpin, mediated by a Ser2056 at -2 positions to the FxxA motif, encompasses a much larger N-terminal boundary of the repeat to what has been previously delineated as important for binding.[80] Remarkably, an almost identical structural feature, mediated by a Thr111 at the analogous -2 position, was also observed in my subsequent investigation of the *L. infantum* BRC1:Rad51 complex. In the case of *L. infantum*, the β-hairpin permits the formation of additional hydrophobic contacts by the repeat, and I propose that this structural feature achieves high-affinity binding in the absence of a functional LFDE motif. It is possible that other repeats harbouring a hydroxyl-containing Ser/Thr at -2 to FxxA will form such a structural feature in other organisms. For example, the human BRC3 carries a threonine at the analogous position and its FxxA$^{\text{BRC3}}$ module has the second-highest average energetic contribution to binding after FxxA$^{\text{BRC8}}$ (ΔΔG = -0.35 and -0.96 kCal/mol, respectively), which further substantiates a role for this residue in high-affinity binding. In *Arabidopsis*, two out of four repeats have a Thr at this position, four out of six in *P. falciparum*, while trypanosomatids (*Leishmania* and *Trypanosoma* species) have it in all of their BRC repeats.

By investigating the BRC:Rad51 interaction in *L. infantum*, I have demonstrated that the C-terminal region of BRC repeats, and in particular the LFDE motif, is not a universal pre-requisite for high-affinity binding, despite previous reports demonstrating that it is indispensable for a functional RAD51:BRCA2 interaction in human cells.[81] It is intriguing why such differences have evolved in these phylogenetically distant organisms. Homologous recombination has to be carefully balanced, and perhaps by evolving alternative determinants of BRC binding affinity, an organism can fine-tune BRCA2 ortholog activity at the protein

level. An important question then arises: is affinity the only determinant of BRC repeat function? Carreira and Kowalczykowski have shown that human BRC1-4 have a distinct set of functions from BRC5-8.[107] Several reports, including the microfluidic FP experiments discussed in **Section 1.7**, have confirmed that BRC1-4 have higher affinities over BRC5-8. It therefore appears that affinity is indeed the main determinant of function. The BRC8-2 repeat would provide an interesting case for testing this, as it is a structurally and biophysically described high affinity binder with constituent modules from both groups of repeats. If affinity is the only determinant of function, then BRC8-2 is expected to behave like BRC repeats 1-4, whereas the low affinity BRC2-8 will functionally resemble repeats 5-8.

Previous work has shown that BRC repeats as individual molecules are inhibitory to nucleofilament formation and strand exchange but become stimulatory when part of a larger multivalent system, such as the BRCA2 protein or dimeric GST fusions. Shivji *et al*. have demonstrated that a BRCA2 fragment encompassing the eight BRC repeats (BRCA2$_{BRC1-8}$) binds and stabilises the Rad51-ssDNA NF, and stimulates strand exchange.[165] It is not clear how such a multivalent system binds the NF without disrupting it by displacement of the oligomerisation epitope. Perhaps the interaction is mediated via a different set of high-avidity interactions that are not in direct competition with FxxA-mediated RAD51 oligomerisation. One possibility is that the LFDE module residues may mediate this interaction, since they bind outside the FxxA site. Superposing the BRC4:RAD51 complex with the nucleofilament structure suggests that this binding mode may be permissible and would only require a shift in the flexible N-terminal domain (**Figure 9D**). Interestingly, in the study by Carreira and Kowalczykowski, BRC5 was found to be the most potent stimulator of Rad51-ssDNA complex formation, while the repeat shuffling experiments show that it has the weakest-binding FxxA module.[107] BRC8, on the other hand, was the least potent NF stimulator in the weak-affinity BRC5-8 set, while having the tightest-binding FxxA module and the second-weakest LFDE module, as shown by the shuffling experiments. This suggests that the nucleoprotein filament-stabilising function is LFDE-module dependent and is in counteracted by inhibitory nature of the repeat as a whole. It is possible that the two BRC repeats in *L. infantum* have similarly differentiated roles. LiBRC1 may be the high affinity repeat, equivalent to *Hs*BRC1-4, that sequesters Rad51 to ssDNA and increases nucleation, the weaker-binding *Li*BRC2 may then stabilise the nascent nucleofilament.

The work presented here provides an improved structural understanding of modular contributions to binding. Future studies should investigate the functional consequences of BRC repeat modules in more detail. For example, it would be interesting to see if the high-affinity

LFDE[BRC2] module in isolation or as multimeric species may be able to bind and stabilise the NF in the absence of an FxxA module.

6.2 A novel way of screening stapled peptides

Understanding the structural basis of protein-protein interaction can inform drug discovery efforts, and the binding epitopes can themselves serve as templates for inhibitor design. In this work, I have utilised the high affinity BRC8-2 repeat for the development of stapled peptide RAD51 inhibitors. In order to test different stapling architectures, I developed a work-flow for the recombinant preparation and early-stage screening of stapled peptides and tested its applicability to RAD51. The approach demonstrates that stapled peptides can be prepared and evaluated in a laboratory that lacks organic synthesis facilities. Moreover, this method allows parallelisation, since HPLC purification, which can be a bottleneck for SPPS, is not required. The overall cost of consumables and reagents is also likely to be smaller. The described methodology can be applied to a range of therapeutic targets. It will be most suitably applied to similar cases, where a high affinity linear epitope can be extracted from a PPI.

More work is needed to evaluate the scope of the stapled peptide screening approach. For example, is it translatable to other types of Cys-reactive chemistries? Dibromoxylene linkers have been utilised to introduce a shorter, more constrained cysteine staple that is most optimal for a single helical turn ($i,i+4$),[166] yet it is unclear if one can achieve the same product purity with these alternative linkers as was done with divinyls. It is also not clear if non-helical moieties can be stapled efficiently with other linkers, as was shown here. Depending on the target, the range of downstream screening assays can likewise be expanded. For example, *in vitro* assays such as Förster resonance energy transfer (FRET), homogenous time-resolved fluorescence (HTRF) and AlphaScreen (PerkinElmer), among others, should in principle be compatible with the work-flow.

6.3 Targeting RAD51 with a new modality

One of the main challenges of contemporary drug discovery is the so called "undruggable" proteome, defined by targets that are not easily addressable by conventional small-molecule approaches. Until recently, RAD51 could be classified as undruggable due to the physicochemical nature of the interactions that need to be disrupted in order to inhibit it. The recent report of a small-molecule compound targeting the FxxA site to inhibit RAD51 function represents a ground-breaking step in drugging this target.[117] Here, I have shown for

the first time that an alternative modality – stapled peptides – can be employed to this end. The *in vitro* and structural analyses reported here support the application of stapled peptide based on BRC repeats towards disrupting RAD51:BRCA2 and RAD51:DNA interactions. Cellular experiments with the reported peptides are still in progress and will be the ultimate measure of utility for these molecules.

I anticipate that the results presented here will motivate further efforts of targeting RAD51 with macrocyclic peptides. Follow-up work should focus on optimisation of the various pharmacological properties of stapled BRC repeats to ensure target occupancy *in vivo*. Increasing affinity is the most obvious direction, because low- or sub-nanomolar affinities can allow even poorly permeable peptides to achieve high target engagement. The molecules developed in this work present plenty of avenues for increasing affinity, such as those discussed in **Section 3.9**. Structural information from orthologous repeats can also be utilised to design additional interactions. For example, a salt bridge was observed between *Li*BRC1 Arg124 and *Li*Rad51 Glu249, which may be re-capitulated in BRC8-2 by mutating the equivalent Ser1230 to an arginine, since there exists a reciprocal surface glutamic acid on *Hs*RAD51. Moreover, decreasing size of the peptides can improve their membrane permeability. The observation that *Li*BRC1 does not need the LFDE motif for high affinity binding is encouraging, since it suggests that BRC8-2 may be further truncated at the C-terminus, granted that compensatory binding features are introduced elsewhere, resulting in a significantly smaller molecule. For example, one may attempt to recreate the additional hydrophobic contacts observed in *Li*BRC1. Double stapling may be pursued further to improve proteolytic resistance and even allow for oral delivery *in vivo*. For example, in a study of HIV fusion inhibitors, hydrocarbon double-stapling of 37-residue peptides resulted in remarkable oral bioavailability profiles in mice.[167]

Generalised rules have been applied for hydrocarbon stapling of helical motifs, such as which aliphatic side-chain lengths are most appropriate for each inter-residue distance.[149] For less regular structural motifs, however, bespoke design is required. In this study, structure-guided design was paramount to the successful application of stapling to a diverse set of staple architectures. In future development of RAD51-targeting peptides, *in silico* methods, such as molecular dynamics simulations, may be applied to evaluate in advance a much larger initial library of peptides, which can then be triaged into a smaller set to be prepared synthetically. This would allow for the early identification of stapling architectures that are conducive to high affinity binding.

The increasing amount of structural and biophysical data on targeting Rad51 and other poorly-druggable proteins may be utilised for a more multi-faceted approach of probe design.

For example, previously identified fragment hits could be integrated into bespoke unnatural amino acids and incorporated into stapled peptides for more efficient binding. Combining the expanding toolkit of different chemical and biological methods will enable the development of more sophisticated ways of targeting the high-hanging fruit of the proteome.

7 Materials and methods

7.1 Vectors for bacterial expression of peptides

Throughout this work, peptides were recombinantly expressed in *E. coli* as fusions to a GB1 tag, a protein domain derived from the streptococcal protein G that manifests high solubility and superior expression levels. For this purpose, three expression vectors were employed: pEXP-NHis-GB1 (Dr Aleksei Lulla, unpublished, Addgene #112565), pOP3BT (Dr Marko Hyvönen, unpublished, Addgene #112603) and pPEPT1. pOP3BT and pEXP-Nhis-GB1 were previously developed at our lab and available in-house. pPEPT1 was designed and constructed during the project. Plasmid sequence and map for pPEPT1 is shown in **Appendix 8.4**.

7.2 Sequence and ligation independent cloning (SLIC) of BRC repeat peptides

Peptide-coding DNA inserts were designed using the DNAWorks online application,[168] while optimising codon usage for *E. coli* expression. DNA oligonucleotides for the assembly of these fragments were generated using the same software and 15-20 nt linkers were appended to the 5' ends of the outermost forward and reverse oligos for sequence and ligation independent cloning (SLIC). The inserts were then synthesised by assembly PCR using Pfusion DNA polymerase (New England Biolabs) with standard reaction conditions. Each assembly PCR reaction contained 1 µM of the outermost oligos and 0.02 µM of each internal oligo. Inserts were then purified by gel extraction using the GeneJet gel extraction kit (Thermo Scientific). For each of the peptides, all of the oligos as well as the recipient vector digest, are provided in **Appendix 8.5** *Oligonucleotides for assembly PCR and* cloning.

10 µl SLIC reactions were prepared as follows: purified insert at 2-10 ng/µl (final), cut vector at 4-20 ng/µl (final), 1x NEB Buffer 2.1 (New England Biolabs). Then, 0.2 µL (0.6 U) of T4 DNA polymerase (New England Biolabs) was added to each reaction and incubated for

1-5 min at RT, after which 1 µl of 100 mM of either dGTP or dCTP was added to stop the exonuclease activity. Reaction was further incubated for 1 minute at RT, after which it was heated for 5 minutes at 65 °C to deactivate the T4 polymerase, after which the PCR tubes were left at room temperature for 10-20 minutes for the complementary resected ends of the insert and vector to anneal. Whole reaction mixtures were then used to transform chemicompetent T7Express *E. coli* cells (prepared in-house from a New England Biolabs stock). Correct clones were then determined by colony PCR and Sanger sequencing.

7.3 Sequence and ligation independent cloning of *Li*Rad51 constructs

*Li*Rad51-coding (Uniprot: A4I3C9_LEIIN) gene was codon optimised for expression in *E. coli* and purchased as a synthetic gene from ThermoFisher. The full sequence of the synthetic gene is provided in **Section 8.6** *LiRad51 synthetic gene. Li*Rad51 construct inserts were prepared by PCR, using the synthetic gene as template. All of the vectors, primers and restriction enzymes are provided in **Appendix 8.5** *Oligonucleotides for assembly PCR and cloning*. Full-length *Li*Rad51 was cloned into pEXP-MBP as fusion to N-terminal His8-MBP. *Li*Rad51$^{\text{ATPase}}$ was cloned into pHAT2 as fusion to an N-terminal His8-tag. *Li*Rad51$^{\text{ATPase,}\Delta\text{L2}}$ was cloned into pBAT4, lacking any fusion tags. SLIC cloning was done in an identical manner to **Section 7.2** *Sequence and ligation independent cloning (SLIC) of BRC repeat peptides*.

7.4 Purification of monomeric HumRadA22

Humanised archaeal surrogate protein HumRadA22 is based on the monomeric ATPase domain of *Pf*RadA. A glycerol stock of BL21(DE3) *E. coli* carrying a pBAT4 vector with a HumRadA22 insert lacking any expression tags was provided by Dr Marko Hyvönen. The construct has been described previously.[111] These cells were plated on LB agar supplemented with ampicillin (100 µg/ml) and grown overnight at 37 °C. Next day, cells were scraped and used to inoculate 1L 2xYT medium. Liquid media cultures were grown for 3-5 h until OD$_{600}$ = 1, after which expression was induced with IPTG (400 µM) for 3 hours at 37 °C. After protein expression, cells were centrifuged at 4000 g for 25 min and resuspended in 25 ml of buffer A: 20 mM MES pH 6.0, 1 mM EDTA and stored at -20 °C for later use.

Frozen cells were thawed and lysed either using an Emulsiflex homogeniser (Avestin) or by sonication. Lysates were kept on ice. Lysate was centrifugated at 35000 x*g* for 40 min and soluble fraction was removed with a 25 ml serological pipette, and was filtered using a 40 µm syringe filter. The lysate soluble fraction was loaded on a HiTrap SP HP 5 ml column

(Cytiva), which was then washed with 5 CV buffer A. 10-20 CV gradient of buffer B (20 mM MES pH 6.0, 1 mM EDTA, 1 M NaCl) was used to elute the bound HumRadA22 protein.

Cation-exchange fractions were analysed by SDS-PAGE and those containing the protein were pooled and concentrated to 2 ml volume using a 10 kDa MWCO spin concentrator. The sample was then centrifuged for 2 min at 15000 x*g* and loaded on a Superdex 200 16/60 size exclusion column (GE Lifesciences) equilibrated with 20 mM CHES pH 9.5, 100 mM NaCl, 1 mM EDTA HumRadA22. The monomeric protein eluted as single peak at about 75 ml elution volume. Fractions containing the purified protein were pooled, concentrated and stored at -80 °C.

7.5 Purification of monomeric *Li*Rad51$^{\text{ATPase}}$

T7Express *E. coli* cells carrying pHAT2 plasmid (Dr Marko Hyvönen, unpublished, Addgene #112583) expressing N-terminally His$_8$-tagged *Li*Rad51$^{\text{ATPase}}$ were plated from glycerol stocks on LB agar supplemented with ampicillin (100 μg/ml) and grown overnight at 37 °C. Next day, cells were scraped and used to inoculate flasks containing 1 L of 2x YT medium supplemented with 100 μg/mL ampicillin. Cultures were grown at 37°C until OD$_{600}$ of ~0.5, after which temperature was decreased to 15 °C. Once the desired temperature was reached, expression was induced with IPTG (400 μM) overnight. Next day, cells were resuspended in 25 mL of nickel buffer A (50 mM Tris-HCl pH 8.0, 150 mM NaCl, 20 mM imidazole) and frozen. Later, cells were thawed and supplemented with DNase I (100 μL, 2 mg/mL) and AEBSF (1 mM), and lysed on an Emulsiflex C5 ho3mogenizer (Avestin) or by sonication. Cell lysate was centrifuged at 40000 x*g* for 30 min and supernatant collected. Lysate was loaded on a 3 mL Ni-NT A agarose matrix (Cube Biotech), after which column matrix was washed with 10 CV Nickel Buffer A. *Li*Rad51$^{\text{ATPase}}$ was eluted with 12 ml nickel buffer B (50 mM Tris-HCl pH 8.0, 150 mM NaCl, 200 mM imidazole). Protein was concentrated to 2 ml on a centrifugal filter (Amicon, MWCO 10 000 Da) and loaded on a Superdex 75 16/60 HiLoad size exclusion column equilibrated with 20 mM Tris pH 8.0, 100 mM NaCl, 100 mM Li$_2$SO$_4$. Fractions containing monomeric *Li*Rad51$^{\text{ATPase}}$ were pooled, concentrated and flash-frozen.

7.6 Purification of BRC repeat:HumRadA22 peptide complexes

E. coli BL21(DE3) or T7Express cells carrying the pOP3BT-BRC and pBAT4-HumRadA22 plasmids were separately plated on LB agar supplemented with ampicillin (100 μg/ml) and grown overnight at 37 °C. Next day, cells were scraped and used to inoculate

separate flasks containing 1 L of 2x YT medium supplemented with 100 µg/mL. Cultures were grown at 37°C until OD_{600} of ~1, after which expression was induced with IPTG (400 µM) for three hours. Cells were resuspended in 25 mL of nickel buffer A (50 mM Tris-HCl pH=8.0, 150 mM NaCl, 20 mM imidazole) and frozen. Later, cells were thawed and supplemented with DNase I (100 µL, 2 mg/mL) and AEBSF (1 mM), and lysed on an Emulsiflex C5 homogenizer (Avestin) or by sonication. Cell lysate was centrifuged at 40 000 xg for 30 min and supernatant collected. GB1-BRC lysates were loaded on a 3 mL Ni-NT A agarose matrix (Cube Biotech), followed by the application of HumRadA22 lysate from equal culture volume. Column matrix was washed with 10 column volumes Nickel Buffer A. Complex was eluted with nickel buffer B (50 mM Tris-HCl pH 8.0, 150 mM NaCl, 200 mM imidazole) into 2 ml fractions. Fractions containing the complex were pooled (~10 ml total) and buffer exchanged back into nickel buffer A on a PD-10 desalting column (Cytiva). Buffer exchanged nickel output was incubated with 100 µL of 2 mg/ml TEV protease overnight at 4°C. GB1 fusion partner was then removed from the solution by a second Ni-NTA affinity step, collecting the flow-through that contains the complex. Flow-through was concentrated on a centrifugal filter (Amicon, MWCO 3000 Da) to 2 ml and loaded onto a Superdex 75 16/60 HiLoad size exclusion column (Cytiva), previously equilibrated with 20 mM CHES pH 9.5, 100 mM NaCl, 1 mM EDTA. The complex eluted at ~75 ml, fractions were analysed by SDS-PAGE. Fractions containing the complex were pooled and concentrated to 0.45 mM for crystallisation trials.

7.7 Purification of the *Li*BRC1: *Li*Rad51$^{\text{ATPase},\Delta\text{L2}}$ complex

*Li*BRC1:*Li*Rad51$^{\text{ATPase},\Delta\text{L2}}$ complex for crystallisation was prepared in an otherwise identical manner to **Section 7.6** *Purification of BRC repeat:HumRadA22 peptide complexes*, except for the following differences. *Li*Rad51$^{\text{ATPase},\Delta\text{L2}}$ expression was induced overnight at 15°C. Ni-NTA buffers were supplemented with 100 mM Li_2SO_4. SEC was done using the following buffer: 20 mM Tris pH 8.0, 100 mM NaCl, 100 mM Li_2SO_4, 1 mM EDTA.

7.8 Purification of GB1-*Li*BRC fusions

GB1-*Li*BRC fusions were purified using a two-stage process. A C-terminal His tag is used in the first capture step, followed by a second anion exchange step that takes advantage of the highly negatively charged N-terminal GB1, ensuring that only species containing both terminal fusions are recovered and no significant BRC repeat degradation products are present in the final product.

T7Express *E. coli* cells carrying pPEPT1 plasmids expressing GB1-fused *Li*BRC repeat constructs were plated directly from glycerol stocks onto LB agar supplemented with ampicillin (100 µg/ml) and grown overnight at 37°C. Next day, cells were scraped and used to inoculate separate flasks containing 1 L of 2x YT medium supplemented with 100 µg/mL ampicillin. Cultures were grown at 37 °C until OD_{600} of ~1, after which expression was induced with IPTG (400 µM) for three hours. Cells were resuspended in 25 mL of nickel buffer A (50 mM Tris-HCl pH 8.0, 150 mM NaCl, 20 mM imidazole) and frozen. Later, cells were thawed and supplemented with DNase I (100 µL, 2 mg/mL) and AEBSF (1 mM), and lysed on an Emulsiflex C5 homogenizer (Avestin) or by sonication. Cell lysate was centrifuged at 40 000 x*g* for 30 min and supernatant collected. GB1-*Li*BRC lysate was loaded on a 3 mL Ni-NTA agarose matrix (Cube Biotech), after which column matrix was washed with 10 column volumes Nickel Buffer A. Bound protein was eluted with 10 ml nickel buffer B (50 mM Tris-HCl pH 8.0, 150 mM NaCl, 200 mM imidazole). The sample was diluted to 80 ml Q-A buffer (20 mM Tris-HCl, pH 8.0, 1 mM EDTA) and loaded on a HiTrap Q 5 ml column, which was washed with 8 CV of the same buffer, after which the GB1-*Li*BRC fusion was eluted with a linear 0-100% gradient of Q-B buffer (20 mM Tris-HCl, pH 8.0, 1 M NaCl, 1 mM EDTA). Peak fractions containing the desired product were pooled, concentrated and flash-frozen for storage.

7.9 Preparation of stapled peptide:HumRadA22 complexes for crystallography

Stapled peptide complexes were re-constituted from purified peptides and HumRadA22, rather than co-precipitated from lysate, like was done for linear BRC repeat complexes. Peptides were added at a 1.5 stoichiometric excess to HumRadA22 in its size-exclusion buffer, to a final concentration of 0.75 and 0.5 mM for the peptide and protein, respectively.

7.10 Purification of full-length human RAD51

Full-length *Hs*RAD51 was prepared based on a protocol developed at the lab of Prof Luca Pellegrini (Department of Biochemistry, Universiy of Cambridge). *E. coli* BL21(DE3) Rosetta2 cells carrying a pRSF-Duet plasmid co-expressing wild-type *Hs*RAD51 and a BRC4 sequence fused to an N-terminal His-MBP tag were kindly provided by Dr Joseph Maman. Cells were plated from a glycerol stock on LB agar supplemented with kanamycin (25 µg/mL)

and chloramphenicol (34 µg/mL), and grown overnight at 37°C. Next day, cells were scraped and used to inoculate 1 L of 2xYT medium supplemented with same antibiotics. Cells were grown at 37°C with shaking at 200 RPM until an OD_{600} = 0.6, afer which they were cooled down to 18°C and expression induced with IPTG (400 µM) overnight. Cells were resuspended in 25 mL of buffer Ni-A-300 (50 mM Tris-HCl, pH 8.0, 300 mM NaCl, 20 mM imidazole, 1 mM TCEP) and frozen. Later, cells were thawed and supplemented with DNase I (100 µL, 2 mg/mL) and AEBSF (1 mM), and lysed on an Emulsiflex C5 homogenizer (Avestin) or by sonication. Cell lysate was spun down at 40 000 g for 30 min, after which the soluble fraction was loaded on a HisTrap HP 5 ml column (Cytiva). The column was washed with 8 CV Ni-A-300 buffer, after which MBP-BRC4:RAD51 complex was co-eluted with buffer Ni-B-300 (50 mM Tris-HCl, pH 8.0, 300 mM NaCl, 200 mM imidazole, 1 mM TCEP).

The sample was then diluted with Heparin-A buffer (20 mM HEPES pH 7.4, 50 mM NaCl, 1 mM EDTA, 1 mM TCEP) and loaded on a HiTrap Heparin HP 5 ml column. During this step, RAD51 oligomerises on the heparin matrix, which acts as a DNA mimic, and dissociates from the MBP-BRC4 fusion, which is removed in flow-through and wash steps. Column was washed with 8 CV Heparin-A buffer, after which the protein was eluted with a 20 CV, 0-100% linear gradient of Heparin-B buffer (20 mM HEPES pH 7.4, 1 M NaCl, 1 mM EDTA, 1 mM TCEP). RAD51 was concentrated, flash-frozen with liquid nitrogen and stored for future use.

7.11 Crystallography of RAD51:BRC repeat complexes

Crystallisation condition screens of linear and stapled BRC repeat complexes were performed in a generally similar manner. Protein solutions were maintained at 0.5 mM concentration, which amounts to around 10 mg/ml, in their respective SEC buffers. ADP/MgCl2 was added to the protein solution in some of the screens to a final concentration of 20 mM. It is important to note that during complex preparations, in the final SEC steps, all of the buffering agents were kept at a low concentration of 20 mM to allow for effective screening of different pH values in the crystallisation screens employed, which typically employ 100 mM buffer concentrations.

Crystallisation screening was done in 96-well MRC plates using the sitting-drop vapor diffusion technique and a variety of commercial crystallisation screens. A Mosquito liquid handling robot (TTP Labtech) was used to dispense protein and reservoir solutions in sub-microlitre volumes. Typical drops contained 200:200 and 400:200 nl of protein:crystallisation

solutions, while the reservoir contained 80 µl of crystallisation solution. Plates were stored at 17°C in a RockImager crystallisation hotel (Formulatrix) and imaged regularly.

Crystal hits were fished with mounted loops and flash-frozen in liquid nitrogen. Additional cryoprotectant was not added before freezing of crystal hits. Diffraction data were collected on Diamond Light Source (Harwell, UK) beamlines i03, i04, i04-1 and i24. Full native datasets with goniometer sweeps of at least 180° were collected to ensure completeness of diffraction data. For HumRadA22 complexes, molecular replacement phasing method was used with the apo HumRadA22 structure (PDB: 5KDD) as a search model. For LiRad51$^{ATPase,\Delta L2}$ complexes, human Rad51 ATPase domain was used as search model (PDB: 1N0W). Molecular replacement was done with Phaser.[169] The structures were refined without BRC repeats first and the peptides were built into the clearly visible electron density manually. Manual refinement was done in Coot[170] and automated refinement with phenix.refine[169] and autoBUSTER[171]. Crystallisation conditions, as well as data collection and refinement statistics, are provided in **Appendix 8.7** *Crystallographic conditions and data collection/refinement statistics*.

7.12 Small-scale preparation of stapled peptides

T7Express *E. coli* cells carrying pPEPT1 vectors with GB1-fused BRC repeat bis-cysteine mutants were plated directly from glycerol stocks of sequence-verified clones onto LB agar supplemented with ampicillin (100 µg/ml) and grown overnight at 37°C. Next day, enough cells to cover a plastic inoculation loop were scraped and transferred to 10 ml of 2xYT medium supplemented with ampicillin (100 µg/ml) in a 50 ml centrifuge tube. The tubes were left only partially sealed for aeration and autoclave tape was applied to keep the lid in a fixed position. Cultures were incubated at 37°C with shaking at 200 RPM for 4 h or until they were very turbid (OD_{600} = 3-4). Alternatively, rather than growing the cells in the morning, a single colony can be picked from the agar plate using a pipette tip and used to inoculate the 10 ml culture, which is then grown overnight (~16 h) at 37°C with shaking. Protein expression was induced by the addition of IPTG (400 µM) for three hours at 37°C. After expression, cells were centrifugated and pellets frozen at -20°C.

Cells were thawed at room temperature by resuspension in 1 ml of lysis buffer: PBS, 20 mM imidazole, 1 mM TCEP, 1 mM EDTA, 0.1% Triton X-100, 0.2 mg/ml lysoszyme, 1 mM AEBSF, 10 µg/ml DNase I. Lysate was incubated for 10min at room temperature on a rotating mixer. Lysates were spun down in a 2 ml tube on a bench-top centrifuge at 15 000 g for 10 min and supernatant collected by aspiration. 200 µl of re-suspended Ni-NTA resin (100

µl of set resin, Cube Biotech) was washed twice with 1 ml of MilliQ water and resuspended in 200 µl PBS. The resin was mixed with the soluble lysate and incubated on a vertical rotating shaker for 5 min at room temperature, after which it was applied in two portions to a 0.5 ml spin column and centrifugated for 1 min at 1000 x*g* to remove the buffer. Same centrifuge settings were also used for subsequent wash and elution steps. The resin was washed with a total of 1 ml of PBS + 20 mM imidazole containing 1 mM TCEP, followed by a 0.5 ml wash using the same buffer without any reducing agent. The second wash step is crucial to remove any residual TCEP prior to elution, as the nucleophilic reducing agent will form undesired side-products upon reaction with the divinyl linker. The GB1-BRC repeat was eluted with 0.5 ml PBS + 200 mM imidazole, and the elution immediately used for subsequent stapling reactions.

Two otherwise identical solutions of stapled and free thiol control peptide were prepared. To this end, the Ni-NTA elution output was split into 2x250 µl halves. 2 mM divinyl linker solution in DMSO was gradually titrated into the stapled sample to achieve pseudo-dilution conditions and drive single linker macrocycle closing. At the same time, an identical volume of DMSO control without any linker was added to the other 250 µl bis-cysteine peptide solution. 1 mM TCEP was added to the control reaction but not the stapling reaction to maintain free sulfhydryl groups in the control peptides. At the end of the titration, reactions were quenched with 10 mM TCEP.

Both the stapled and control reactions were concentrated to 100 µl on a 3 kDa MWCO spin filter (Amicon) and used in subsequent experiments. Concentration of the unstapled GB1-BRC control reaction product was spectrophotometrically measured on a Nanodrop 1000 or Nanodrop One device using absorbance at 280 nm. Concentration measurements of stapled peptides are complicated by the presence of the aromatic pyrimidine or triazine moieties in the linker molecules that also absorb at 280 nm. Considering this, matching concentrations were assumed for the corresponding stapling/control reaction products, since both reaction inputs come from the same Ni-NTA resin elution and identical final volumes were obtained during linker or DMSO control additions and subsequent concentration steps.

7.13 Solid-phase peptide synthesis (SPPS)

Solid-phase peptide synthesis was done using Fmoc chemistry on a CEM Liberty Automated Microwave Peptide Synthesiser. Merck Rink Amide resin LL (0.29-0.39 mmol/g, 286 mg per reaction) was used. Peptide couplings were conducted with Fmoc-protected amino acids (5 equiv in DMF), DIC/Oxyma Pure (5/10 equiv) in DMF as the coupling reagent, and N,N-diisopropylethylamine (10 equiv) in NMP as the base. Double coupling was used for all

amino acids, with 25 W power at 75 °C over 15 min (50 °C for cysteine amino acids). Fmoc deprotection was carried out using 20% piperidine in DMF, with 45 W power at 75 °C over 3 min. N-termini were capped with acetic anhydride. Cleavage from resin and global deprotection was carried out with a cocktail of 92.5% trifluoroacetic acid, 2.5% water, 2.5% DCM and 2.5% triisopropylsilane for 2h at 45 °C in a sand bath. Cleavage solution was evaporated under nitrogen and reaction products precipitated using ice-cold diethyl ether, after which the products were resuspended in varying water/acetonitrile solutions.

7.14 Liquid chromatography mass spectrometry (LC-MS)

Protein intact mass was measured on a Xevo G2-S Q-TOF mass spectrometer coupled to an Acquity UPLC system using an Acquity UPLC BEH300 C4 column. 0.1% formic acid in ultrapure water was used as buffer A, with a 5-95% linear gradient of buffer B (95% acetonitrile, 0.1% formic acid) at a flow rate of 0.2 mL min^{-1}. The electrospray source was operated with a capillary voltage of 2.0 kV and a cone voltage of 40 V. Nitrogen was used as the desolvation gas at a total flow of 850 L h^{-1}. Total mass spectra were deconvoluted from the ion series using the MaxEnt algorithm pre-installed on MassLynx software (v.4.1 from Waters) according to the manufacturer's instructions.

LC-MS for peptides having molecular weight less than 5 kDa was done on a waters Acquity H-Class UPLC with an ESCi Multi-Mode ionisation Waters SQ Detector 2 spectrometer. Buffer A: 2 mM ammonium acetate, 0.1% formic acid in 95/5% water/acetonitrile. Buffer B: 100% acetonitrile, 0.1% formic acid. Column: AQUITY UPLC CSH C18, 2.1*50 mm, 1.7 μm, 130 Å; gradient: 5-95% B over 3 min.

7.15 Peptide concentration measurements

Many of the purified peptides lacked side chains suitable for conventional UV-Vis concentration determination at 280 nm. Instead, vacuum-dried peptides were resuspended in MilliQ water and their concentration was determined by measuring 205 nm absorbance using the Scopes method[172] on a Nanodrop One UV-VIS spectrophotometer (ThermoFisher).

7.16 Recombinant expression and purification of linear peptides

T7Express *E. coli* cells carrying pOP3BT plasmids expressing GB1-fused BRC repeat were plated directly from glycerol stocks of sequence-verified clones onto LB agar supplemented with ampicillin (100 μg/ml) and grown overnight at 37 °C. Next day, cells were

scraped and used to inoculate separate flasks containing 1 L of 2x YT medium supplemented with 100 µg/mL ampicillin. Cultures were grown at 37 °C until OD_{600} of ~1, after which expression was induced with IPTG (400 µM) for three hours. Cells were resuspended in 25 mL of nickel buffer A (50 mM Tris-HCl pH=8.0, 150 mM NaCl, 20 mM imidazole) and frozen. Later, cells were thawed and supplemented with DNase I (100 µL, 2 mg/mL) and AEBSF (1 mM), and lysed on an Emulsiflex C5 homogenizer (Avestin) or by sonication. Cell lysate was centrifuged at 40000 x*g* for 30 min and supernatant collected. GB1-BRC lysate was loaded on a 3 mL Ni-NT A agarose matrix (Cube Biotech), after which column matrix was washed with 10 column volumes Nickel Buffer A. GB1-BRC repeat was eluted with 12 ml nickel buffer B (50 mM Tris-HCl pH 8.0, 150 mM NaCl, 200 mM imidazole). The eluent was buffer exchanged back into nickel buffer A on a PD-10 desalting column (Cytiva). Buffer exchanged GB1-BRC fusion (~18 ml) was incubated with 100 µL of 2 mg/ml TEV protease overnight at 4°C. The GB1 tag was then removed from the solution by a second Ni-NTA affinity step, collecting the flow-through that contains the BRC peptide. The flow-through was acidified with HCl to pH 2-4 and acetonitrile was added to 10%, after which the solution was centrifuged at 10000 x*g* for 15 min to remove aggregates. The acidified flow-through was then applied to an ACE C8 300 4.6 x 250 mm semi-prep RP-HPLC column equilibrated with RPC buffer A (10% acetonitrile, 0.1% TFA) and peptides were eluted with a 20 CV gradient of RPC buffer B (90% acetonitrile, 0.1% TFA). BRC repeat peptides typically elute at 20-40% buffer B. Peak fractions were analysed by LCMS and pooled for drying under vacuum. If necessary, an additional purification step on a ACE C18 300 4.6 x 250 mm semi-prep RP-HPLC column was included at the end, using identical buffers to the C8 step.

Dried peptides were resuspended either in MilliQ water or a buffer of interest. Because most of the BRC repeat peptides lack any tryptophan/tyrosine residues, estimation of concentration using absorbance at 280 nm is not possible, therefore concentration was determined using 205 nm absorbance using the Stokes method on a Nanodrop One UV-VIS spectrophotometer (ThermoFisher).

7.17 Stapling of recombinantly prepared BRC repeat free peptides

Cysteine-stapled BRC repeat peptides lacking any fusion tags were first expressed recombinantly and purified in an identical fashion to the **Section 7.16** *Recombinant expression and purification of linear peptides*. 1 mM TCEP was included in all purification steps to maintain cysteine thiols in a reduced state. After the C8 RPC step, rather than drying the peptide under vacuum, peak fractions corresponding to the cleaved peptide were pooled and diluted 5x

in PBS + 10 mM EDTA in a 50 ml centrifuge tube. A small stir rod was added to the tube, which was then placed on a magnetic stirrer. A syringe was filled with 20 mM linker in DMSO which was then gradually added to the mixture by piercing the centrifuge tube lid. To maintain pseudo-dilution conditions, linker was injected in 50 µl increments every two minutes, to a final concentration of 2 mM, ensuring at least 2x stoichiometric excess of linker over peptide. The reaction mixture was then quenched with 5 mM TCEP, filtered through a 0.45 µm filter and acidified with HCl to pH ~3. Stapled peptide was then purified by a on a ACE C18 300 4.6 x 250 mm semi-prep RP-HPLC column with a 0-100% gradient of A: 0.1% TFA, B: 90 % MeCN + 0.1% TFA. Peak fractions containing the desired product were pooled and dried under vacuum.

7.18 Isothermal titration calorimetry (ITC) of BRC repeat binding

Peptides were resuspended in MilliQ water to 10 times the desired concentration in the syringe. This was then diluted 10x with the ITC buffer to obtain the final titrant solution. HumRadA22 was buffer-exchanged on a NAP-5 desalting column into ITC buffer and protein concentration was adjusted to 10:9 of the desired final value. One ninth volume of MilliQ water was added to the solution to bring the protein concentration to the desired final value, while maintaining identical buffer:MilliQ volume proportions in both the syringe and the cell. ITC was carried out using a Microcal ITC200 instrument at 25°C with a 5.00 µCal reference power DP value, stirring speed of 500-750 rpm, 2 sec filter period. Injection spacing, speed and volume, cell/syringe concentrations as well as the number of injections were adjusted for each peptide and its binding properties. ITC data were fitted using a single-site binding model using the Microcal ITC data analysis program in the Origin 7.0 package. Data points affected by baseline spikes were omitted from the analysis. ITC buffer: 20 mM CHES pH 9.5, 150 mM NaCl, 0.1% Tween-20.

7.19 HumRadA22 fluorescence polarisation (FP) competition assay

HumRadA22 competition assay was a modified version of a protocol described previously.[111] Fluorescein-labelled HsBRC4 probe for this assay was kindly provided by Dr Laurens Lindenburg (Hollfelder group, Department of Biochemistry, University of Cambridge). Black 384-well flat-bottom microplates (Corning #3821) were used with a 40 µl final reaction volume in all measurements. Following buffer conditions were used: 20 mM CHES pH 9.5, 150 mM NaCl, 0.1% BSA, 0.1% Tween-20. Each reaction contained 100 nM

HumRadA22 and 10 nM BRC4-fluorescein. Two-fold serial dilutions of the peptide were added to the reactions. A free probe control reaction containing only 10 nM BRC4-fluorescein was used to calibrate gain and focal height. FP measurements were performed on a Pherastar FX (BMG Labtech) plate reader equipped with an FP 485-520-520 optic module. Binding curves were fitted using the four-parameter logistic model with a variable Hill slope using Prism software (Graphpad). Regression fitting was performed using the least squares optimisation algorithm. K_D values were estimated from the fitted IC_{50} parameters using a previously reported equation.[173]

7.20 *Li*Rad51$^{\text{ATPase}}$ fluorescence polarisation (FP) competition assay

Experiments were prepared in an identical manner to **Section 7.19** *HumRadA22 fluorescence polarisation (FP) competition assay*, except for the differences described here. Following FP buffer was used: 50 mM Tris pH 8.0, 150 mM NaCl, 100 mM Li_2SO4, 1% BSA, 0.1% Tween-20. Each reaction contained 5 nM of Fluor-NCys-LiBRC1 probe and 500 nM of *Li*BRC1$^{\text{ATPase}}$.

7.21 Circular dichroism (CD) spectroscopy

For CD measurements, dried peptides were dissolved in MilliQ water to 0.3 mg/ml, and then two-fold diluted in 20 mM sodium phosphate, pH 7.4, giving a final solution of 0.15 mg/ml peptide in 10 mM sodium phosphate. CD spectra of selected peptides were recorded on an AVIV 410 circular dichroism spectropolarimeter using a 1 mm path length quartz cuvette. Measurements were done at 25 °C, with a 185-260 nm range, 1 nm bandwidth, 5 s averaging time and 0.3 s settling time. Spectra were prepared as smoothed average of three scans and normalised against blank solvent.

7.22 Electrophoretic mobility shift assay (EMSA)

The ability of linear and stapled BRC repeat peptides to dissociate RAD51-ssDNA nucleofilament was evaluated using an electrophoretic mobility shift assay (EMSA). RAD51 DNA-binding reactions (40 µl) were set up in 50 mM HEPES pH 7.4, 150 mM NaCl, 10 mM MgAc$_2$, 2 mM CaCl$_2$, 1 mM TCEP, 1 mM ATP. 5 µM full-length human RAD51 was incubated with varying concentrations of BRC repeats for 10 min at room temperature, followed by the addition of 100 nM fluorescently labelled FAM-dT60 oligonucleotide, and further incubation at 37 °C for 10 min. Control reactions were set up with free FAM-dT60 probe and FAM-dT60

+ 5 µM RAD51. 10 µl of reactions were then loaded on a 1xTBE non-denaturing acrylamide gel (5%) and run at 100 V for 1:30 h at 4 °C. The gel was directly visualized on a Typhoon FLA 9000 imager (GE Healthcare) using FAM channels.

EMSAs with *Li*Rad51 were performed in identical manner but were resolved on a 2% agarose gel, running for 6 min at 250 V / 4°C, rather than on an acrylamide gel. This is because it is impossible to resolve the *Li*Rad51 NF on the smaller pores of an acrylamide gel, presumably due to the highly disordered N-terminal region of *Li*Rad51, which increases the hydrodynamic radius of the complex.

7.23 Peptide serum stability assay

Peptides were diluted to 600 µM in PBS. Human serum (Sigma-Aldrich, #H4522) was diluted to 30% in PBS and incubated at 37 °C for 10 min. 200 µl of peptide solution were added to 400 µl of 30% serum, giving a final concentration of 200 µM and 20% for peptide and serum, respectively. Mixtures were incubated in 1.5 ml tubes at 37 °C on a Thermomixer R (Eppendorf). 50 µl aliquots were taken at specific time points and quenched with 100 µl of 1:1 ethanol:DMSO, incubated for 3 min and then centrifugated at 13000x*g* for 5 min to remove precipitating serum proteins. The supernatant was then two-fold serially diluted and used in FP competition reactions as described in **7.19** *HumRadA22 fluorescence polarisation (FP) competition assay*. Changes in IC$_{50}$ over different incubation time-points were monitored as an indicator of detrimental proteolytic degradation.

Bibliography

(1)   Negrini, S.; Gorgoulis, V. G.; Halazonetis, T. D. Genomic Instability — an Evolving Hallmark of Cancer. *Nat Rev Mol Cell Biol* **2010**, *11* (3), 220–228. https://doi.org/10.1038/nrm2858.

(2)   Ciccia, A.; Elledge, S. J. The DNA Damage Response: Making It Safe to Play with Knives. *Molecular Cell* **2010**, *40* (2), 179–204. https://doi.org/10.1016/j.molcel.2010.09.019.

(3)   Alexander, J. L.; Orr-Weaver, T. L. Replication Fork Instability and the Consequences of Fork Collisions from Rereplication. *Genes Dev.* **2016**, *30* (20), 2241–2252. https://doi.org/10.1101/gad.288142.116.

(4)   Pâques, F.; Haber, J. E. Multiple Pathways of Recombination Induced by Double-Strand Breaks in Saccharomyces Cerevisiae. *Microbiol. Mol. Biol. Rev.* **1999**, *63* (2), 349–404. https://doi.org/10.1128/MMBR.63.2.349-404.1999.

(5)   Roth, D. B. V(D)J Recombination: Mechanism, Errors, and Fidelity. *Microbiology Spectrum* **2014**, *2* (6), 1–11. https://doi.org/10.1128/microbiolspec.MDNA3-0041-2014.

(6)   Pannunzio, N. R.; Watanabe, G.; Lieber, M. R. Nonhomologous DNA End-Joining for Repair of DNA Double-Strand Breaks. *J. Biol. Chem.* **2018**, *293* (27), 10512–10523. https://doi.org/10.1074/jbc.TM117.000374.

(7)   Sallmyr, A.; Tomkinson, A. E. Repair of DNA Double-Strand Breaks by Mammalian Alternative End-Joining Pathways. *J. Biol. Chem.* **2018**, *293* (27), 10536–10546. https://doi.org/10.1074/jbc.TM117.000375.

(8)   Bhargava, R.; Onyango, D. O.; Stark, J. M. Regulation of Single-Strand Annealing and Its Role in Genome Maintenance. *Trends in Genetics* **2016**, *32* (9), 566–575. https://doi.org/10.1016/j.tig.2016.06.007.

(9)   Ferretti, L. P.; Lafranchi, L.; Sartori, A. A. Controlling DNA-End Resection: A New Task for CDKs. *Front. Genet.* **2013**, *4* (99), 1–7. https://doi.org/10.3389/fgene.2013.00099.

(10)   Ait Saada, A.; Lambert, S. A. E.; Carr, A. M. Preserving Replication Fork Integrity and Competence via the Homologous Recombination Pathway. *DNA Repair* **2018**, *71*, 135–147. https://doi.org/10.1016/j.dnarep.2018.08.017.

(11)   Kass, E. M.; Jasin, M. Collaboration and Competition between DNA Double-Strand Break Repair Pathways. *FEBS Letters* **2010**, *584* (17), 3703–3708. https://doi.org/10.1016/j.febslet.2010.07.057.

(12)   Jasin, M.; Rothstein, R. Repair of Strand Breaks by Homologous Recombination. 18.

(13)   Kadyk, L. C.; Hartwell, L. H. Sister Chromatids Are Preferred over Homologs as Substrates for Recombinational Repair in Saccharomyces Cerevisiae. *Genetics* **1992**, *132* (2), 387–402. https://doi.org/10.1093/genetics/132.2.387.

(14)   Fernandez, J.; Bloomer, H.; Kellam, N.; LaRocque, J. R. Chromosome Preference During Homologous Recombination Repair of DNA Double-Strand Breaks in *Drosophila Melanogaster*. *G3* **2019**, *9* (11), 3773–3780. https://doi.org/10.1534/g3.119.400607.

(15)   Mao, Z.; Bozzella, M.; Seluanov, A.; Gorbunova, V. DNA Repair by Nonhomologous End Joining and Homologous Recombination during Cell Cycle in Human Cells. *Cell Cycle* **2008**, *7* (18), 2902–2906. https://doi.org/10.4161/cc.7.18.6679.

(16)   Lamarche, B. J.; Orazio, N. I.; Weitzman, M. D. The MRN Complex in Double-Strand Break Repair and Telomere Maintenance. *FEBS Letters* **2010**, *584* (17), 3682–3695. https://doi.org/10.1016/j.febslet.2010.07.029.

(17)   Garcia, V.; Phelps, S. E. L.; Gray, S.; Neale, M. J. Bidirectional Resection of DNA Double-Strand Breaks by Mre11 and Exo1. *Nature* **2011**, *479* (7372), 241–244. https://doi.org/10.1038/nature10515.

(18)   Möckel, C.; Lammens, K.; Schele, A.; Hopfner, K.-P. ATP Driven Structural Changes of the Bacterial Mre11:Rad50 Catalytic Head Complex. *Nucleic Acids Research* **2012**, *40* (2), 914–927. https://doi.org/10.1093/nar/gkr749.

(19)   Nimonkar, A. V.; Genschel, J.; Kinoshita, E.; Polaczek, P.; Campbell, J. L.; Wyman, C.; Modrich, P.; Kowalczykowski, S. C. BLM-DNA2-RPA-MRN and EXO1-BLM-RPA-MRN Constitute Two DNA End Resection Machineries for Human DNA Break Repair. *Genes & Development* **2011**, *25* (4), 350–362. https://doi.org/10.1101/gad.2003811.

(20)   Cannavo, E.; Cejka, P.; Kowalczykowski, S. C. Relationship of DNA Degradation by Saccharomyces Cerevisiae Exonuclease 1 and Its Stimulation by RPA and Mre11-Rad50-Xrs2 to DNA End Resection. *Proceedings of the National Academy of Sciences* **2013**, *110* (18), E1661–E1668. https://doi.org/10.1073/pnas.1305166110.

(21)   Chung, W.-H.; Zhu, Z.; Papusha, A.; Malkova, A.; Ira, G. Defective Resection at DNA Double-Strand Breaks Leads to De Novo Telomere Formation and Enhances Gene Targeting. *PLoS Genet* **2010**, *6* (5), e1000948. https://doi.org/10.1371/journal.pgen.1000948.

(22)   Levikova, M.; Pinto, C.; Cejka, P. The Motor Activity of DNA2 Functions as an SsDNA Translocase to Promote DNA End Resection. *Genes Dev.* **2017**, *31* (5), 493–502. https://doi.org/10.1101/gad.295196.116.

(23)   Symington, L. S. End Resection at Double-Strand Breaks: Mechanism and Regulation. *Cold Spring Harbor Perspectives in Biology* **2014**, *6* (8), a016436–a016436. https://doi.org/10.1101/cshperspect.a016436.

(24)   Myler, L. R.; Gallardo, I. F.; Soniat, M. M.; Deshpande, R. A.; Gonzalez, X. B.; Kim, Y.; Paull, T. T.; Finkelstein, I. J. Single-Molecule Imaging Reveals How Mre11-Rad50-Nbs1 Initiates DNA Break Repair. *Molecular Cell* **2017**, *67* (5), 891-898.e4. https://doi.org/10.1016/j.molcel.2017.08.002.

(25)   Wang, H.; Shi, L. Z.; Wong, C. C. L.; Han, X.; Hwang, P. Y.-H.; Truong, L. N.; Zhu, Q.; Shao, Z.; Chen, D. J.; Berns, M. W.; Yates, J. R.; Chen, L.; Wu, X. The Interaction of CtIP and Nbs1 Connects CDK and ATM to Regulate HR–Mediated Double-Strand Break Repair. *PLoS Genet* **2013**, *9* (2), e1003277. https://doi.org/10.1371/journal.pgen.1003277.

(26)   Anand, R.; Ranjha, L.; Cannavo, E.; Cejka, P. Phosphorylated CtIP Functions as a Co-Factor of the MRE11-RAD50-NBS1 Endonuclease in DNA End Resection. *Molecular Cell* **2016**, *64* (5), 940–950. https://doi.org/10.1016/j.molcel.2016.10.017.

(27)   Tsukamoto, Y.; Mitsuoka, C.; Terasawa, M.; Ogawa, H.; Ogawa, T. Xrs2p Regulates Mre11p Translocation to the Nucleus and Plays a Role in Telomere Elongation and Meiotic Recombination□D. *Molecular Biology of the Cell* **2005**, *16*, 12.

(28)   Lavin, M. F. ATM and the Mre11 Complex Combine to Recognize and Signal DNA Double-Strand Breaks. *Oncogene* **2007**, *26* (56), 7749–7758. https://doi.org/10.1038/sj.onc.1210880.

(29)   Cheng, Q.; Chen, J. Mechanism of P53 Stabilization by ATM after DNA Damage. *Cell Cycle* **2010**, *9* (3), 472–478. https://doi.org/10.4161/cc.9.3.10556.

(30)   Qin, Z.; Bi, L.; Hou, X.-M.; Zhang, S.; Zhang, X.; Lu, Y.; Li, M.; Modesti, M.; Xi, X.-G.; Sun, B. Human RPA Activates BLM's Bidirectional DNA Unwinding from a Nick. *eLife* **2020**, *9*, e54098. https://doi.org/10.7554/eLife.54098.

(31)   Gasior, S. L.; Olivares, H.; Ear, U.; Hari, D. M.; Weichselbaum, R.; Bishop, D. K. Assembly of RecA-like Recombinases: Distinct Roles for Mediator Proteins in Mitosis and

Meiosis. *Proceedings of the National Academy of Sciences* **2001**, *98* (15), 8411–8418. https://doi.org/10.1073/pnas.121046198.

(32)   Sugiyama, T.; Zaitseva, E. M.; Kowalczykowski, S. C. A Single-Stranded DNA-Binding Protein Is Needed for Efficient Presynaptic Complex Formation by the *Saccharomyces Cerevisiae* Rad51 Protein. *J. Biol. Chem.* **1997**, *272* (12), 7940–7945. https://doi.org/10.1074/jbc.272.12.7940.

(33)   Wang, X.; Haber, J. E. Role of Saccharomyces Single-Stranded DNA-Binding Protein RPA in the Strand Invasion Step of Double-Strand Break Repair. *PLoS Biol* **2004**, *2* (1), e21. https://doi.org/10.1371/journal.pbio.0020021.

(34)   Ma, C. J.; Gibb, B.; Kwon, Y.; Sung, P.; Greene, E. C. Protein Dynamics of Human RPA and RAD51 on SsDNA during Assembly and Disassembly of the RAD51 Filament. *Nucleic Acids Res* **2017**, *45* (2), 749–761. https://doi.org/10.1093/nar/gkw1125.

(35)   Symington, L. S. Role of RAD52 Epistasis Group Genes in Homologous Recombination and Double-Strand Break Repair. *MMBR* **2002**, *66* (4), 630–670. https://doi.org/10.1128/MMBR.66.4.630-670.2002.

(36)   Rijkers, T.; Van Den Ouweland, J.; Morolli, B.; Rolink, A. G.; Baarends, W. M.; Van Sloun, P. P. H.; Lohman, P. H. M.; Pastink, A. Targeted Inactivation of Mouse RAD52Reduces Homologous Recombination but Not Resistance to Ionizing Radiation. *Mol. Cell. Biol.* **1998**, *18* (11), 6423–6429. https://doi.org/10.1128/MCB.18.11.6423.

(37)   Liu, J.; Doty, T.; Gibson, B.; Heyer, W.-D. Human BRCA2 Protein Promotes RAD51 Filament Formation on RPA-Covered Single-Stranded DNA. *Nat Struct Mol Biol* **2010**, *17* (10), 1260–1262. https://doi.org/10.1038/nsmb.1904.

(38)   Feng, Z.; Scott, S. P.; Bussen, W.; Sharma, G. G.; Guo, G.; Pandita, T. K.; Powell, S. N. Rad52 Inactivation Is Synthetically Lethal with BRCA2 Deficiency. *Proceedings of the National Academy of Sciences* **2011**, *108* (2), 686–691. https://doi.org/10.1073/pnas.1010959107.

(39)   Liu, J.; Heyer, W.-D. Who's Who in Human Recombination: BRCA2 and RAD52. *Proceedings of the National Academy of Sciences* **2011**, *108* (2), 441–442. https://doi.org/10.1073/pnas.1016614108.

(40)   Liu, N.; Lamerdin, J. E.; Tebbs, R. S.; Schild, D.; Tucker, J. D.; Shen, M. R.; Brookman, K. W.; Siciliano, M. J.; Walter, C. A.; Fan, W.; Narayana, L. S.; Zhou, Z.-Q.; Adamson, A. W.; Sorensen, K. J.; Chen, D. J.; Jones, N. J.; Thompson, L. H. XRCC2 and XRCC3, New Human Rad51-Family Members, Promote Chromosome Stability and Protect

against DNA Cross-Links and Other Damages. *Molecular Cell* **1998**, *1* (6), 783–793. https://doi.org/10.1016/S1097-2765(00)80078-7.

(41)   Pittman, D. L.; Schimenti, J. C. Midgestation Lethality in Mice Deficient for the RecA-related Gene, Rad51d/Rad51l3. *Genesis* **2000**, *26* (3), 167–173.

(42)   Martino, J.; Brunette, G. J.; Barroso-González, J.; Moiseeva, T. N.; Smith, C. M.; Bakkenist, C. J.; O'Sullivan, R. J.; Bernstein, K. A. The Human Shu Complex Functions with PDS5B and SPIDR to Promote Homologous Recombination. *Nucleic Acids Research* **2019**, *47* (19), 10151–10165. https://doi.org/10.1093/nar/gkz738.

(43)   Chun, J.; Buechelmaier, E. S.; Powell, S. N. Rad51 Paralog Complexes BCDX2 and CX3 Act at Different Stages in the BRCA1-BRCA2-Dependent Homologous Recombination Pathway. *Molecular and Cellular Biology* **2013**, *33* (2), 9.

(44)   Amunugama, R.; Groden, J.; Fishel, R. The HsRAD51B–HsRAD51C Stabilizes the HsRAD51 Nucleoprotein Filament. *DNA Repair* **2013**, *12* (9), 723–732. https://doi.org/10.1016/j.dnarep.2013.05.005.

(45)   Stasiak, A.; Egelman, E. H. Structure and Function of RecA-DNA Complexes. *Experientia* **1994**, *50* (3), 192–203. https://doi.org/10.1007/BF01924002.

(46)   Bonilla, B.; Hengel, S. R.; Grundy, M. K.; Bernstein, K. A. RAD51 Gene Family Structure and Function. 24.

(47)   Greene, E. C. DNA Sequence Alignment during Homologous Recombination. *J. Biol. Chem.* **2016**, *291* (22), 11572–11580. https://doi.org/10.1074/jbc.R116.724807.

(48)   McVey, M.; Khodaverdian, V. Y.; Meyer, D.; Cerqueira, P. G.; Heyer, W.-D. Eukaryotic DNA Polymerases in Homologous Recombination. *Annu. Rev. Genet.* **2016**, *50* (1), 393–421. https://doi.org/10.1146/annurev-genet-120215-035243.

(49)   Li, X.; Heyer, W.-D. RAD54 Controls Access to the Invading 3′-OH End after RAD51-Mediated DNA Strand Invasion in Homologous Recombination in Saccharomyces Cerevisiae. *Nucleic Acids Research* **2009**, *37* (2), 638–646. https://doi.org/10.1093/nar/gkn980.

(50)   Wright, W. D.; Heyer, W.-D. Rad54 Functions as a Heteroduplex DNA Pump Modulated by Its DNA Substrates and Rad51 during D Loop Formation. *Molecular Cell* **2014**, *53* (3), 420–432. https://doi.org/10.1016/j.molcel.2013.12.027.

(51)   Nimonkar, A. V.; Sica, R. A.; Kowalczykowski, S. C. Rad52 Promotes Second-End DNA Capture in Double-Stranded Break Repair to Form Complement-Stabilized Joint Molecules. *Proceedings of the National Academy of Sciences* **2009**, *106* (9), 3077–3082. https://doi.org/10.1073/pnas.0813247106.

(52)   Matos, J.; West, S. C. Holliday Junction Resolution: Regulation in Space and Time. *DNA Repair* **2014**, *19*, 176–181. https://doi.org/10.1016/j.dnarep.2014.03.013.

(53)   Swuec, P.; Costa, A. Molecular Mechanism of Double Holliday Junction Dissolution. *Cell Biosci* **2014**, *4* (1), 36. https://doi.org/10.1186/2045-3701-4-36.

(54)   Cejka, P.; Plank, J. L.; Bachrati, C. Z.; Hickson, I. D.; Kowalczykowski, S. C. Rmi1 Stimulates Decatenation of Double Holliday Junctions during Dissolution by Sgs1–Top3. *Nat Struct Mol Biol* **2010**, *17* (11), 1377–1382. https://doi.org/10.1038/nsmb.1919.

(55)   Punatar, R. S.; Martin, M. J.; Wyatt, H. D. M.; Chan, Y. W.; West, S. C. Resolution of Single and Double Holliday Junction Recombination Intermediates by GEN1. 8.

(56)   Kaur, H.; De Muyt, A.; Lichten, M. Top3-Rmi1 DNA Single-Strand Decatenase Is Integral to the Formation and Resolution of Meiotic Recombination Intermediates. *Molecular Cell* **2015**, *57* (4), 583–594. https://doi.org/10.1016/j.molcel.2015.01.020.

(57)   McMahill, M. S.; Sham, C. W.; Bishop, D. K. Synthesis-Dependent Strand Annealing in Meiosis. *PLoS Biol* **2007**, *5* (11), e299. https://doi.org/10.1371/journal.pbio.0050299.

(58)   Game, J. C.; Mortimer, R. K. A Genetic Study of X-Ray Sensitive Mutants in Yeast. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* **1974**, *24* (3), 281–292. https://doi.org/10.1016/0027-5107(74)90176-6.

(59)   Shinohara, A.; Ogawa, H.; Ogawa, T. Rad51 Protein Involved in Repair and Recombination in S. Cerevisiae Is a RecA-like Protein. *Cell* **1992**, *69* (3), 457–470. https://doi.org/10.1016/0092-8674(92)90447-K.

(60)   Sonoda, E. Rad51-Deficient Vertebrate Cells Accumulate Chromosomal Breaks Prior to Cell Death. *The EMBO Journal* **1998**, *17* (2), 598–608. https://doi.org/10.1093/emboj/17.2.598.

(61)   Jeyasekharan, A. D.; Liu, Y.; Hattori, H.; Pisupati, V.; Jonsdottir, A. B.; Rajendra, E.; Lee, M.; Sundaramoorthy, E.; Schlachter, S.; Kaminski, C. F.; Ofir-Rosenfeld, Y.; Sato, K.; Savill, J.; Ayoub, N.; Venkitaraman, A. R. A Cancer-Associated BRCA2 Mutation Reveals Masked Nuclear Export Signals Controlling Localization. *Nat Struct Mol Biol* **2013**, *20* (10), 1191–1198. https://doi.org/10.1038/nsmb.2666.

(62)   Stassen, N. Y.; Logsdon Jr., J. M.; Vora, G. J.; Offenberg, H. H.; Palmer, J. D.; Zolan, M. E. Isolation and Characterization of Rad51 Orthologs from Coprinus Cinereus and Lycopersicon Esculentum, and Phylogenetic Analysis of Eukaryotic RecA Homologs. *Current Genetics* **1997**, *31* (2), 144–157. https://doi.org/10.1007/s002940050189.

(63)   Lin, Z.; Kong, H.; Nei, M.; Ma, H. Origins and Evolution of the RecA/RAD51 Gene Family: Evidence for Ancient Gene Duplication and Endosymbiotic Gene Transfer. *Proceedings of the National Academy of Sciences* **2006**, *103* (27), 10328–10333. https://doi.org/10.1073/pnas.0604232103.

(64)   Liu, J.; Renault, L.; Veaute, X.; Fabre, F.; Stahlberg, H.; Heyer, W.-D. Rad51 Paralogues Rad55–Rad57 Balance the Antirecombinase Srs2 in Rad51 Filament Formation. *Nature* **2011**, *479* (7372), 245–248. https://doi.org/10.1038/nature10522.

(65)   Baumann, P.; Benson, F. E.; West, S. C. Human Rad51 Protein Promotes ATP-Dependent Homologous Pairing and Strand Transfer Reactions In Vitro. *Cell* **1996**, *87* (4), 757–766. https://doi.org/10.1016/S0092-8674(00)81394-X.

(66)   Hilario, J.; Amitani, I.; Baskin, R. J.; Kowalczykowski, S. C. Direct Imaging of Human Rad51 Nucleoprotein Dynamics on Individual DNA Molecules. *Proceedings of the National Academy of Sciences* **2009**, *106* (2), 361–368. https://doi.org/10.1073/pnas.0811965106.

(67)   Forget, A. L.; Kowalczykowski, S. C. Single-Molecule Imaging of DNA Pairing by RecA Reveals a Three-Dimensional Homology Search. *Nature* **2012**, *482* (7385), 423–427. https://doi.org/10.1038/nature10782.

(68)   Morrison, C.; Shinohara, A.; Sonoda, E.; Yamaguchi-Iwai, Y.; Takata, M.; Weichselbaum, R. R.; Takeda, S. The Essential Functions of Human Rad51 Are Independent of ATP Hydrolysis. *Mol. Cell. Biol.* **1999**, *19* (10), 6891–6897. https://doi.org/10.1128/MCB.19.10.6891.

(69)   Chi, P.; Van Komen, S.; Sehorn, M. G.; Sigurdsson, S.; Sung, P. Roles of ATP Binding and ATP Hydrolysis in Human Rad51 Recombinase Function. *DNA Repair* **2006**, *5* (3), 381–391. https://doi.org/10.1016/j.dnarep.2005.11.005.

(70)   Bugreev, D. V.; Mazin, A. V. Ca2+ Activates Human Homologous Recombination Protein Rad51 by Modulating Its ATPase Activity. *Proceedings of the National Academy of Sciences* **2004**, *101* (27), 9988–9993. https://doi.org/10.1073/pnas.0402105101.

(71)   Ristic, D. Human Rad51 Filaments on Double- and Single-Stranded DNA: Correlating Regular and Irregular Forms with Recombination Function. *Nucleic Acids Research* **2005**, *33* (10), 3292–3302. https://doi.org/10.1093/nar/gki640.

(72)   Sawaya, M. R.; Guo, S.; Tabor, S.; Richardson, C. C.; Ellenberger, T. Crystal Structure of the Helicase Domain from the Replicative Helicase-Primase of Bacteriophage T7. *Cell* **1999**, *99* (2), 167–177. https://doi.org/10.1016/S0092-8674(00)81648-7.

(73)   Abrahams, J. P.; Leslie, A. G. W.; Lutter, R.; Walker, J. E. Structure at 2.8 Å Resolution of F1-ATPase from Bovine Heart Mitochondria. *Nature* **1994**, *370* (6491), 621–628. https://doi.org/10.1038/370621a0.

(74)   Marsh, M. E.; Scott, D. E.; Ehebauer, M. T.; Abell, C.; Blundell, T. L.; Hyvönen, M. ATP Half-Sites in RadA and RAD51 Recombinases Bind Nucleotides. *FEBS Open Bio* **2016**, *6* (5), 372–385. https://doi.org/10.1002/2211-5463.12052.

(75)   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 NMR. *Journal of Molecular Biology* **1999**, *290* (2), 495–504. https://doi.org/10.1006/jmbi.1999.2904.

(76)   Xu, J.; Zhao, L.; Xu, Y.; Zhao, W.; Sung, P.; Wang, H.-W. Cryo-EM Structures of Human RAD51 Recombinase Filaments during Catalysis of DNA-Strand Exchange. *Nat Struct Mol Biol* **2017**, *24* (1), 40–46. https://doi.org/10.1038/nsmb.3336.

(77)   Yu, X.; Egelman, E. H. The RecA Hexamer Is a Structural Homologue of Ring Helicases. *Nat Struct Mol Biol* **1997**, *4* (2), 101–104. https://doi.org/10.1038/nsb0297-101.

(78)   Shin, D. S. Full-Length Archaeal Rad51 Structure and Mutants: Mechanisms for RAD51 Assembly and Control by BRCA2. *The EMBO Journal* **2003**, *22* (17), 4566–4576. https://doi.org/10.1093/emboj/cdg429.

(79)   Conway, A. B.; Lynch, T. W.; Zhang, Y.; Fortin, G. S.; Fung, C. W.; Symington, L. S.; Rice, P. A. Crystal Structure of a Rad51 Filament. *Nat Struct Mol Biol* **2004**, *11* (8), 791–796. https://doi.org/10.1038/nsmb795.

(80)   Pellegrini, L.; Yu, D. S.; Lo, T.; Anand, S.; Lee, M.; Blundell, T. L.; Venkitaraman, A. R. Insights into DNA Recombination from the Structure of a RAD51–BRCA2 Complex. *Nature* **2002**, *420* (6913), 287–293. https://doi.org/10.1038/nature01230.

(81)   Rajendra, E.; Venkitaraman, A. R. Two Modules in the BRC Repeats of BRCA2 Mediate Structural and Functional Interactions with the RAD51 Recombinase. *Nucleic Acids Research* **2010**, *38* (1), 82–96. https://doi.org/10.1093/nar/gkp873.

(82)   Klapstein, K.; Chou, T.; Bruinsma, R. Physics of RecA-Mediated Homologous Recognition. *Biophysical Journal* **2004**, *87* (3), 1466–1477. https://doi.org/10.1529/biophysj.104.039578.

(83)   Stasiak, A.; Di Capua, E.; Koller, Th. Elongation of Duplex DNA by RecA Protein. *Journal of Molecular Biology* **1981**, *151* (3), 557–564. https://doi.org/10.1016/0022-2836(81)90010-3.

(84)   Di Capua, E.; Engel, A.; Stasiak, A.; Koller, Th. Characterization of Complexes between RecA Protein and Duplex DNA by Electron Microscopy. *Journal of Molecular Biology* **1982**, *157* (1), 87–103. https://doi.org/10.1016/0022-2836(82)90514-9.

(85)   Yu, X.; Egelman, E. H. Structural Data Suggest That the Active and Inactive Forms of the RecA Filament Are Not Simply Interconvertible. *Journal of Molecular Biology* **1992**, *227* (1), 334–346. https://doi.org/10.1016/0022-2836(92)90702-L.

(86)   Short, J. M.; Liu, Y.; Chen, S.; Soni, N.; Madhusudhan, M. S.; Shivji, M. K. K.; Venkitaraman, A. R. High-Resolution Structure of the Presynaptic RAD51 Filament on Single-Stranded DNA by Electron Cryo-Microscopy. *Nucleic Acids Res* **2016**, gkw783. https://doi.org/10.1093/nar/gkw783.

(87)   Davies, O. R.; Pellegrini, L. Interaction with the BRCA2 C Terminus Protects RAD51–DNA Filaments from Disassembly by BRC Repeats. *Nat Struct Mol Biol* **2007**, *14* (6), 475–483. https://doi.org/10.1038/nsmb1251.

(88)   Tombline, G.; Fishel, R. Biochemical Characterization of the Human RAD51 Protein: I. ATP HYDROLYSIS. *J. Biol. Chem.* **2002**, *277* (17), 14417–14425. https://doi.org/10.1074/jbc.M109915200.

(89)   Chen, J.; Villanueva, N.; Rould, M. A.; Morrical, S. W. Insights into the Mechanism of Rad51 Recombinase from the Structure and Properties of a Filament Interface Mutant. *Nucleic Acids Research* **2010**, *38* (14), 4889–4906. https://doi.org/10.1093/nar/gkq209.

(90)   Amunugama, R.; He, Y.; Willcox, S.; Forties, R. A.; Shim, K.-S.; Bundschuh, R.; Luo, Y.; Griffith, J.; Fishel, R. RAD51 Protein ATP Cap Regulates Nucleoprotein Filament Stability. *J. Biol. Chem.* **2012**, *287* (12), 8724–8736. https://doi.org/10.1074/jbc.M111.239426.

(91)   Andriuskevicius, T.; Kotenko, O.; Makovets, S. Putting Together and Taking Apart: Assembly and Disassembly of the Rad51 Nucleoprotein Filament in DNA Repair and Genome Stability. *CST* **2018**, *2* (5), 96–112. https://doi.org/10.15698/cst2018.05.134.

(92)   Ford, D.; Easton, D. F.; Stratton, M.; Narod, S.; Goldgar, D.; Devilee, P.; Bishop, D. T.; Weber, B.; Lenoir, G.; Chang-Claude, J.; Sobol, H.; Teare, M. D.; Struewing, J.; Arason, A.; Scherneck, S.; Peto, J.; Rebbeck, T. R.; Tonin, P.; Neuhausen, S.; Barkardottir, R.; Eyfjord, J.; Lynch, H.; Ponder, B. A. J.; Gayther, S. A.; Birch, J. M.; Lindblom, A.; Stoppa-Lyonnet, D.; Bignon, Y.; Borg, A.; Hamann, U.; Haites, N.; Scott, R. J.; Maugard, C. M.; Vasen, H.; Seitz, S.; Cannon-Albright, L. A.; Schofield, A.; Zelada-Hedman, M. Genetic Heterogeneity and Penetrance Analysis of the BRCA1 and BRCA2 Genes in Breast Cancer

Families. *The American Journal of Human Genetics* **1998**, *62* (3), 676–689. https://doi.org/10.1086/301749.

(93)   Hakem, R. Developmental Studies of Brca1 and Brca2 Knock-Out Mice. 15.

(94)   Siaud, N.; Dray, E.; Gy, I.; Gérard, E.; Takvorian, N.; Doutriaux, M.-P. Brca2 Is Involved in Meiosis in Arabidopsis Thaliana as Suggested by Its Interaction with Dmc1. *EMBO J* **2004**, *23* (6), 1392–1401. https://doi.org/10.1038/sj.emboj.7600146.

(95)   Jasin, M. Homologous Repair of DNA Damage and Tumorigenesis:The BRCA Connection. *Oncogene* **2002**, *21* (58), 8981–8993. https://doi.org/10.1038/sj.onc.1206176.

(96)   Kojic, M.; Kostrub, C. F.; Buchman, A. R.; Holloman, W. K. BRCA2 Homolog Required for Proficiency in DNA Repair, Recombination, and Genome Stability in Ustilago Maydis. *Molecular Cell* **2002**, *10* (3), 683–691. https://doi.org/10.1016/S1097-2765(02)00632-9.

(97)   Davies, A. A.; Masson, J.-Y.; McIlwraith, M. J.; Stasiak, A. Z.; Stasiak, A.; Venkitaraman, A. R.; West, S. C. Role of BRCA2 in Control of the RAD51 Recombination and DNA Repair Protein. *Molecular Cell* **2001**, *7* (2), 273–282. https://doi.org/10.1016/S1097-2765(01)00175-7.

(98)   Reuter, M.; Zelensky, A.; Smal, I.; Meijering, E.; van Cappellen, W. A.; de Gruiter, H. M.; van Belle, G. J.; van Royen, M. E.; Houtsmuller, A. B.; Essers, J.; Kanaar, R.; Wyman, C. BRCA2 Diffuses as Oligomeric Clusters with RAD51 and Changes Mobility after DNA Damage in Live Cells. *Journal of Cell Biology* **2014**, *207* (5), 599–613. https://doi.org/10.1083/jcb.201405014.

(99)   Tarsounas, M.; Davies, D.; West, S. C. BRCA2-Dependent and Independent Formation of RAD51 Nuclear Foci. *Oncogene* **2003**, *22* (8), 1115–1123. https://doi.org/10.1038/sj.onc.1206263.

(100)   Jensen, R. B.; Carreira, A.; Kowalczykowski, S. C. Purified Human BRCA2 Stimulates RAD51-Mediated Recombination. *Nature* **2010**, *467* (7316), 678–683. https://doi.org/10.1038/nature09399.

(101)   Chen, P.-L.; Chen, C.-F.; Chen, Y.; Xiao, J.; Sharp, Z. D.; Lee, W.-H. The BRC Repeats in BRCA2 Are Critical for RAD51 Binding and Resistance to Methyl Methanesulfonate Treatment. *Proceedings of the National Academy of Sciences* **1998**, *95* (9), 5287–5292. https://doi.org/10.1073/pnas.95.9.5287.

(102)   Bignell, G.; Micklem, G.; Stratton, M. R.; Ashworth, A.; Wooster, R. The BRC Repeats Are Conserved in Mammalian BRCA2 Proteins. 6.

(103)   Mazloum, N.; Zhou, Q.; Holloman, W. K. D-Loop Formation by Brh2 Protein of Ustilago Maydis. *Proceedings of the National Academy of Sciences* **2008**, *105* (2), 524–529. https://doi.org/10.1073/pnas.0707031105.

(104)   Hartley, C. L.; McCulloch, R. Trypanosoma Brucei BRCA2 Acts in Antigenic Variation and Has Undergone a Recent Expansion in BRC Repeat Number That Is Important during Homologous Recombination. *Mol Microbiol* **2008**, *68* (5), 1237–1251. https://doi.org/10.1111/j.1365-2958.2008.06230.x.

(105)   Trenaman, A.; Hartley, C.; Prorocic, M.; Passos-Silva, D. G.; Hoek, M. van den; Nechyporuk-Zloy, V.; Machado, C. R.; McCulloch, R. Trypanosoma Brucei BRCA2 Acts in a Life Cycle-Specific Genome Stability Process and Dictates BRC Repeat Number-Dependent RAD51 Subnuclear Dynamics. *Nucleic Acids Research* **2013**, *41* (2), 943–960. https://doi.org/10.1093/nar/gks1192.

(106)   Filippo, J. S.; Chi, P.; Sehorn, M. G.; Etchin, J.; Krejci, L.; Sung, P. Recombination Mediator and Rad51 Targeting Activities of a Human BRCA2 Polypeptide. *J. Biol. Chem.* **2006**, *281* (17), 11649–11657. https://doi.org/10.1074/jbc.M601249200.

(107)   Carreira, A.; Kowalczykowski, S. C. Two Classes of BRC Repeats in BRCA2 Promote RAD51 Nucleoprotein Filament Function by Distinct Mechanisms. *Proceedings of the National Academy of Sciences* **2011**, *108* (26), 10448–10453. https://doi.org/10.1073/pnas.1106971108.

(108)   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 G2/M Checkpoint Control. *Journal of Biological Chemistry* **1999**, *274* (46), 32931–32935. https://doi.org/10.1074/jbc.274.46.32931.

(109)   Lo, T.; Pellegrini, L.; Venkitaraman, A. R.; Blundell, T. L. Sequence Fingerprints in BRCA2 and RAD51: Implications for DNA Repair and Cancer. *DNA Repair* **2003**, *2* (9), 1015–1028. https://doi.org/10.1016/S1568-7864(03)00097-1.

(110)   Cole, D. J.; Rajendra, E.; Roberts-Thomson, M.; Hardwick, B.; McKenzie, G. J.; Payne, M. C.; Venkitaraman, A. R.; Skylaris, C.-K. Interrogation of the Protein-Protein Interactions between Human BRCA2 BRC Repeats and RAD51 Reveals Atomistic Determinants of Affinity. *PLoS Comput Biol* **2011**, *7* (7), e1002096. https://doi.org/10.1371/journal.pcbi.1002096.

(111)   Moschetti, T.; Sharpe, T.; Fischer, G.; Marsh, M. E.; Ng, H. K.; Morgan, M.; Scott, D. E.; Blundell, T. L.; R. Venkitaraman, A.; Skidmore, J.; Abell, C.; Hyvönen, M. Engineering Archeal Surrogate Systems for the Development of Protein–Protein Interaction

Inhibitors against Human RAD51. *Journal of Molecular Biology* **2016**, *428* (23), 4589–4607. https://doi.org/10.1016/j.jmb.2016.10.009.

(112)   Donoho, G.; Brenneman, M. A.; Cui, T. X.; Donoviel, D.; Vogel, H.; Goodwin, E. H.; Chen, D. J.; Hasty, P. Deletion OfBrca2 Exon 27 Causes Hypersensitivity to DNA Crosslinks, Chromosomal Instability, and Reduced Life Span in Mice. *Genes Chromosom. Cancer* **2003**, *36* (4), 317–331. https://doi.org/10.1002/gcc.10148.

(113)   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* **2007**, *14* (6), 468–474. https://doi.org/10.1038/nsmb1245.

(114)   Esashi, F.; Christ, N.; Gannon, J.; Liu, Y.; Hunt, T.; Jasin, M.; West, S. C. CDK-Dependent Phosphorylation of BRCA2 as a Regulatory Mechanism for Recombinational Repair. *Nature* **2005**, *434* (7033), 598–604. https://doi.org/10.1038/nature03404.

(115)   Lindenburg, L. H.; Pantelejevs, T.; Gielen, F.; Zuazua-Villar, P.; Butz, M.; Rees, E.; Kaminski, C. F.; Downs, J. A.; Hyvönen, M.; Hollfelder, F. Improved RAD51 Binders through Motif Shuffling Based on the Modularity of BRC Repeats. *bioRxiv* **2020**, 2020.05.14.097071. https://doi.org/10.1101/2020.05.14.097071.

(116)   Chatterjee, G.; Jimenez-Sainz, J.; Presti, T.; Nguyen, T.; Jensen, R. B. Distinct Binding of BRCA2 BRC Repeats to RAD51 Generates Differential DNA Damage Sensitivity. *Nucleic Acids Res* **2016**, *44* (11), 5256–5270. https://doi.org/10.1093/nar/gkw242.

(117)   Scott, D. E.; Francis-Newton, N. J.; Marsh, M. E.; Coyne, A. G.; Fischer, G.; Moschetti, T.; Bayly, A. R.; Sharpe, T. D.; Haas, K. T.; Barber, L.; Valenzano, C. R.; Srinivasan, R.; Huggins, D. J.; Lee, M.; Emery, A.; Hardwick, B.; Ehebauer, M.; Dagostin, C.; Esposito, A.; Pellegrini, L.; Perrior, T.; McKenzie, G.; Blundell, T. L.; Hyvönen, M.; Skidmore, J.; Venkitaraman, A. R.; Abell, C. A Small-Molecule Inhibitor of the BRCA2-RAD51 Interaction Modulates RAD51 Assembly and Potentiates DNA Damage-Induced Cell Death. *Cell Chemical Biology* **2021**, S2451945621000581. https://doi.org/10.1016/j.chembiol.2021.02.006.

(118)   Gielen, F.; Butz, M.; Rees, E. J.; Erdelyi, M.; Moschetti, T.; Hyvönen, M.; Edel, J. B.; Kaminski, C. F.; Hollfelder, F. Quantitative Affinity Determination by Fluorescence Anisotropy Measurements of Individual Nanoliter Droplets. *Anal. Chem.* **2017**, *89* (2), 1092–1101. https://doi.org/10.1021/acs.analchem.6b02528.

(119)   Adl, S. M.; Bass, D.; Lane, C. E.; Lukeš, J.; Schoch, C. L.; Smirnov, A.; Agatha, S.; Berney, C.; Brown, M. W.; Burki, F.; Cárdenas, P.; Čepička, I.; Chistyakova, L.; Campo, J.; Dunthorn, M.; Edvardsen, B.; Eglit, Y.; Guillou, L.; Hampl, V.; Heiss, A. A.; Hoppenrath,

M.; James, T. Y.; Karnkowska, A.; Karpov, S.; Kim, E.; Kolisko, M.; Kudryavtsev, A.; Lahr, D. J. G.; Lara, E.; Le Gall, L.; Lynn, D. H.; Mann, D. G.; Massana, R.; Mitchell, E. A. D.; Morrow, C.; Park, J. S.; Pawlowski, J. W.; Powell, M. J.; Richter, D. J.; Rueckert, S.; Shadwick, L.; Shimano, S.; Spiegel, F. W.; Torruella, G.; Youssef, N.; Zlatogursky, V.; Zhang, Q. Revisions to the Classification, Nomenclature, and Diversity of Eukaryotes. *J. Eukaryot. Microbiol.* **2019**, *66* (1), 4–119. https://doi.org/10.1111/jeu.12691.

(120) Zucca, M.; Savoia, D. Current Developments in the Therapy of Protozoan Infections. *Open Med Chem J* **2011**, *5*, 4–10. https://doi.org/10.2174/1874104501105010004.

(121) McCulloch, R.; Barry, J. D. A Role for RAD51 and Homologous Recombination in Trypanosoma Brucei Antigenic Variation. *Genes & Development* **1999**, *13* (21), 2875–2888. https://doi.org/10.1101/gad.13.21.2875.

(122) Gomes Passos Silva, D.; da Silva Santos, S.; Nardelli, S. C.; Mendes, I. C.; Freire, A. C. G.; Repolês, B. M.; Resende, B. C.; Costa-Silva, H. M.; da Silva, V. S.; Oliveira, K. A. de; Oliveira, C. F. B.; Vilela, L. F. F.; Nagem, R. A. P.; Franco, G. R.; Macedo, A. M.; Pena, S. D. J.; Tahara, E. B.; Sales Junior, P. A.; Moreira, D. S.; Teixeira, S. M. R.; McCulloch, R.; Virgilio, S.; Tosi, L. R. O.; Schenkman, S.; Andrade, L. O.; Murta, S. M. F.; Machado, C. R. The in Vivo and in Vitro Roles of Trypanosoma Cruzi Rad51 in the Repair of DNA Double Strand Breaks and Oxidative Lesions. *PLoS Negl Trop Dis* **2018**, *12* (11), e0006875. https://doi.org/10.1371/journal.pntd.0006875.

(123) Alves, C. L.; Repolês, B. M.; da Silva, M. S.; Mendes, I. C.; Marin, P. A.; Aguiar, P. H. N.; Santos, S. da S.; Franco, G. R.; Macedo, A. M.; Pena, S. D. J.; Andrade, L. de O.; Guarneri, A. A.; Tahara, E. B.; Elias, M. C.; Machado, C. R. The Recombinase Rad51 Plays a Key Role in Events of Genetic Exchange in Trypanosoma Cruzi. *Sci Rep* **2018**, *8* (1), 13335. https://doi.org/10.1038/s41598-018-31541-z.

(124) Laffitte, M.-C. N.; Leprohon, P.; Papadopoulou, B.; Ouellette, M. Plasticity of the Leishmania Genome Leading to Gene Copy Number Variations and Drug Resistance. *F1000Res* **2016**, *5*, 2350. https://doi.org/10.12688/f1000research.9218.1.

(125) Baskar, R.; Lee, K. A.; Yeo, R.; Yeoh, K.-W. Cancer and Radiation Therapy: Current Advances and Future Directions. *Int. J. Med. Sci.* **2012**, *9* (3), 193–199. https://doi.org/10.7150/ijms.3635.

(126) Lhuillier, C.; Rudqvist, N.-P.; Elemento, O.; Formenti, S. C.; Demaria, S. Radiation Therapy and Anti-Tumor Immunity: Exposing Immunogenic Mutations to the Immune System. *Genome Med* **2019**, *11* (1), 40. https://doi.org/10.1186/s13073-019-0653-7.

(127)   Klein, H. L. The Consequences of Rad51 Overexpression for Normal and Tumor Cells. *DNA Repair* **2008**, *7* (5), 686–693. https://doi.org/10.1016/j.dnarep.2007.12.008.

(128)   Short, S. C.; Giampieri, S.; Worku, M.; Alcaide-German, M.; Sioftanos, G.; Bourne, S.; Lio, K. I.; Shaked-Rabi, M.; Martindale, C. Rad51 Inhibition Is an Effective Means of Targeting DNA Repair in Glioma Models and CD133+ Tumor-Derived Cells. *Neuro-Oncology* **2011**, *13* (5), 487–499. https://doi.org/10.1093/neuonc/nor010.

(129)   Park, S.-R. Activation-Induced Cytidine Deaminase in B Cell Immunity and Cancers. *Immune Netw* **2012**, *12* (6), 230. https://doi.org/10.4110/in.2012.12.6.230.

(130)   Mills, K.; Cyr, A.; Maclay, T.; Day, M.; Patrick, D.; Clas, S.-D.; Hasham, M. G.; Chow, K. Synthetic Lethality Targeting the Aid/RAD51 Axis for Treatment of AICDA-Expressing Lymphoma. *Blood* **2017**, *130* (Supplement 1), 4628–4628. https://doi.org/10.1182/blood.V130.Suppl_1.4628.4628.

(131)   Ishida, T.; Takizawa, Y.; Kainuma, T.; Inoue, J.; Mikawa, T.; Shibata, T.; Suzuki, H.; Tashiro, S.; Kurumizaka, H. DIDS, a Chemical Compound That Inhibits RAD51-Mediated Homologous Pairing and Strand Exchange. *Nucleic Acids Research* **2009**, *37* (10), 3367–3376. https://doi.org/10.1093/nar/gkp200.

(132)   Huang, F.; Motlekar, N. A.; Burgwin, C. M.; Napper, A. D.; Diamond, S. L.; Mazin, A. V. Identification of Specific Inhibitors of Human RAD51 Recombinase Using High-Throughput Screening. *ACS Chem. Biol.* **2011**, *6* (6), 628–635. https://doi.org/10.1021/cb100428c.

(133)   Day, M.; Lapierre, J.-M.; O'Shea, T.; Mills, K. Abstract C14: A Novel RAD51 Inhibitor, CYT-0851, Shows Anticancer Activity in Preclinical Models of Pancreatic Cancer. *Cancer Res* **2019**, *79* (24 Supplement), C14. https://doi.org/10.1158/1538-7445.PANCA19-C14.

(134)   Budke, B.; Kalin, J. H.; Pawlowski, M.; Zelivianskaia, A. S.; Wu, M.; Kozikowski, A. P.; Connell, P. P. An Optimized RAD51 Inhibitor That Disrupts Homologous Recombination without Requiring Michael Acceptor Reactivity. *J. Med. Chem.* **2013**, *56* (1), 254–263. https://doi.org/10.1021/jm301565b.

(135)   Zhu, J.; Zhou, L.; Wu, G.; Konig, H.; Lin, X.; Li, G.; Qiu, X.; Chen, C.; Hu, C.; Goldblatt, E.; Bhatia, R.; Chamberlin, A. R.; Chen, P.; Lee, W. A Novel Small Molecule RAD51 Inactivator Overcomes Imatinib-resistance in Chronic Myeloid Leukaemia. *EMBO Mol Med* **2013**, *5* (3), 353–365. https://doi.org/10.1002/emmm.201201760.

(136)   Scott, D. E.; Bayly, A. R.; Abell, C.; Skidmore, J. Small Molecules, Big Targets: Drug Discovery Faces the Protein–Protein Interaction Challenge. *Nat Rev Drug Discov* **2016**, *15* (8), 533–550. https://doi.org/10.1038/nrd.2016.29.

(137)   Santos, R.; Ursu, O.; Gaulton, A.; Bento, A. P.; Donadi, R. S.; Bologa, C. G.; Karlsson, A.; Al-Lazikani, B.; Hersey, A.; Oprea, T. I.; Overington, J. P. A Comprehensive Map of Molecular Drug Targets. *Nat Rev Drug Discov* **2017**, *16* (1), 19–34. https://doi.org/10.1038/nrd.2016.230.

(138)   Carter, A. J.; Kraemer, O.; Zwick, M.; Mueller-Fahrnow, A.; Arrowsmith, C. H.; Edwards, A. M. Target 2035: Probing the Human Proteome. *Drug Discovery Today* **2019**, *24* (11), 2111–2115. https://doi.org/10.1016/j.drudis.2019.06.020.

(139)   Hopkins, A. L.; Keserü, G. M.; Leeson, P. D.; Rees, D. C.; Reynolds, C. H. The Role of Ligand Efficiency Metrics in Drug Discovery. *Nat Rev Drug Discov* **2014**, *13* (2), 105–121. https://doi.org/10.1038/nrd4163.

(140)   Scott, D. E.; Ehebauer, M. T.; Pukala, T.; Marsh, M.; Blundell, T. L.; Venkitaraman, A. R.; Abell, C.; Hyvönen, M. Using a Fragment-Based Approach To Target Protein-Protein Interactions. *ChemBioChem* **2013**, *14* (3), 332–342. https://doi.org/10.1002/cbic.201200521.

(141)   Quianzon, C. C.; Cheikh, I. History of Insulin. *Journal of Community Hospital Internal Medicine Perspectives* **2012**, *2* (2), 18701. https://doi.org/10.3402/jchimp.v2i2.18701.

(142)   Dougherty, P. G.; Qian, Z.; Pei, D. Macrocycles as Protein–Protein Interaction Inhibitors. *Biochemical Journal* **2017**, *474* (7), 1109–1125. https://doi.org/10.1042/BCJ20160619.

(143)   Iegre, J.; Brear, P.; Baker, D. J.; Tan, Y. S.; Atkinson, E. L.; Sore, H. F.; O' Donovan, D. H.; Verma, C. S.; Hyvönen, M.; Spring, D. R. Efficient Development of Stable and Highly Functionalised Peptides Targeting the CK2α/CK2β Protein–Protein Interaction. *Chem. Sci.* **2019**, *10* (19), 5056–5063. https://doi.org/10.1039/C9SC00798A.

(144)   Heinis, C.; Rutherford, T.; Freund, S.; Winter, G. Phage-Encoded Combinatorial Chemical Libraries Based on Bicyclic Peptides. *Nat Chem Biol* **2009**, *5* (7), 502–507. https://doi.org/10.1038/nchembio.184.

(145)   Fairlie, D. P.; Dantas de Araujo, A. Stapling Peptides Using Cysteine Crosslinking: Stapling Peptides. *Biopolymers* **2016**, *106* (6), 843–852. https://doi.org/10.1002/bip.22877.

(146)   Robertson, N. S.; Walsh, S. J.; Fowler, E.; Yoshida, M.; Rowe, S. M.; Wu, Y.; Sore, H. F.; Parker, J. S.; Spring, D. R. Macrocyclisation and Functionalisation of Unprotected

Peptides *via* Divinyltriazine Cysteine Stapling. *Chem. Commun.* **2019**, *55* (64), 9499–9502. https://doi.org/10.1039/C9CC05042F.

(147) Bluntzer, M. T. J.; O'Connell, J.; Baker, T. S.; Michel, J.; Hulme, A. N. Designing Stapled Peptides to Inhibit PROTEIN-PROTEIN Interactions: An Analysis of Successes in a Rapidly Changing Field. *Peptide Science* **2021**, *113* (1). https://doi.org/10.1002/pep2.24191.

(148) Guharoy, M.; Chakrabarti, P. Secondary Structure Based Analysis and Classification of Biological Interfaces: Identification of Binding Motifs in Protein–Protein Interactions. *Bioinformatics* **2007**, *23* (15), 1909–1918. https://doi.org/10.1093/bioinformatics/btm274.

(149) Kim, Y.-W.; Grossmann, T. N.; Verdine, G. L. Synthesis of All-Hydrocarbon Stapled α-Helical Peptides by Ring-Closing Olefin Metathesis. *Nat Protoc* **2011**, *6* (6), 761–771. https://doi.org/10.1038/nprot.2011.324.

(150) Wendt, M.; Bellavita, R.; Gerber, A.; Efrém, N.-L.; van Ramshorst, T.; Pearce, N. M.; Davey, P. R. J.; Everard, I.; Vazquez-Chantada, M.; Chiarpari, E.; Grieco, P.; Hennig, S.; Grossmann, T. N. Bicyclic B-sheet Mimetics That Target the Transcriptional Coactivator B-catenin and Inhibit Wnt Signaling. *Angew. Chem. Int. Ed.* **2021**, anie.202102082. https://doi.org/10.1002/anie.202102082.

(151) Yu, D. S.; Sonoda, E.; Takeda, S.; Huang, C. L. H.; Pellegrini, L.; Blundell, T. L.; Venkitaraman, A. R. Dynamic Control of Rad51 Recombinase by Self-Association and Interaction with BRCA2. *Molecular Cell* **2003**, *12* (4), 1029–1041. https://doi.org/10.1016/S1097-2765(03)00394-0.

(152) Scott, D. E.; Marsh, M.; Blundell, T. L.; Abell, C.; Hyvönen, M. Structure-Activity Relationship of the Peptide Binding-Motif Mediating the BRCA2:RAD51 Protein-Protein Interaction. *FEBS Lett* **2016**, *590* (8), 1094–1102. https://doi.org/10.1002/1873-3468.12139.

(153) Sallman, D. A.; Borate, U.; Cull, E. H.; Donnellan, W. B.; Komrokji, R. S.; Steidl, U. G.; Corvez, M. M.; Payton, M.; Annis, D. A.; Pinchasik, D.; Aivado, M.; Verma, A. Phase 1/1b Study of the Stapled Peptide ALRN-6924, a Dual Inhibitor of MDMX and MDM2, As Monotherapy or in Combination with Cytarabine for the Treatment of Relapsed/Refractory AML and Advanced MDS with TP53 Wild-Type. *Blood* **2018**, *132* (Supplement 1), 4066–4066. https://doi.org/10.1182/blood-2018-99-118780.

(154)   Ravasco, J. M. J. M.; Faustino, H.; Trindade, A.; Gois, P. M. P. Bioconjugation with Maleimides: A Useful Tool for Chemical Biology. *Chem. Eur. J.* **2019**, *25* (1), 43–59. https://doi.org/10.1002/chem.201803174.

(155)   Lee, Y.-J.; Kurra, Y.; Liu, W. R. Phospha-Michael Addition as a New Click Reaction for Protein Functionalization. *ChemBioChem* **2016**, *17* (6), 456–461. https://doi.org/10.1002/cbic.201500697.

(156)   Walsh, S. J.; Omarjee, S.; Galloway, W. R. J. D.; Kwan, T. T.-L.; Sore, H. F.; Parker, J. S.; Hyvönen, M.; Carroll, J. S.; Spring, D. R. A General Approach for the Site-Selective Modification of Native Proteins, Enabling the Generation of Stable and Functional Antibody–Drug Conjugates. *Chem. Sci.* **2019**, *10* (3), 694–700. https://doi.org/10.1039/C8SC04645J.

(157)   Guidotti, G.; Brambilla, L.; Rossi, D. Cell-Penetrating Peptides: From Basic Research to Clinics. *Trends in Pharmacological Sciences* **2017**, *38* (4), 406–424. https://doi.org/10.1016/j.tips.2017.01.003.

(158)   Kale, S. S.; Villequey, C.; Kong, X.-D.; Zorzi, A.; Deyle, K.; Heinis, C. Cyclization of Peptides with Two Chemical Bridges Affords Large Scaffold Diversities. *Nature Chem* **2018**, *10* (7), 715–723. https://doi.org/10.1038/s41557-018-0042-7.

(159)   Levengood, M. R.; Zhang, X.; Hunter, J. H.; Emmerton, K. K.; Miyamoto, J. B.; Lewis, T. S.; Senter, P. D. Orthogonal Cysteine Protection Enables Homogeneous Multi-Drug Antibody–Drug Conjugates. *Angew. Chem. Int. Ed.* **2017**, *56* (3), 733–737. https://doi.org/10.1002/anie.201608292.

(160)   Coin, I.; Beyermann, M.; Bienert, M. Solid-Phase Peptide Synthesis: From Standard Procedures to the Synthesis of Difficult Sequences. *Nat Protoc* **2007**, *2* (12), 3247–3256. https://doi.org/10.1038/nprot.2007.454.

(161)   Frei, J. C.; Lai, J. R. Protein and Antibody Engineering by Phage Display. In *Methods in Enzymology*; Elsevier, 2016; Vol. 580, pp 45–87. https://doi.org/10.1016/bs.mie.2016.05.005.

(162)   Vila-Perelló, M.; Muir, T. W. Biological Applications of Protein Splicing. *Cell* **2010**, *143* (2), 191–200. https://doi.org/10.1016/j.cell.2010.09.031.

(163)   Wang, H.; Dawber, R. S.; Zhang, P.; Walko, M.; Wilson, A. J.; Wang, X. Peptide-Based Inhibitors of Protein–Protein Interactions: Biophysical, Structural and Cellular Consequences of Introducing a Constraint. *Chem. Sci.* **2021**, *12* (17), 5977–5993. https://doi.org/10.1039/D1SC00165E.

(164)   Genois, M.-M.; Mukherjee, A.; Ubeda, J.-M.; Buisson, R.; Paquet, E.; Roy, G.; Plourde, M.; Coulombe, Y.; Ouellette, M.; Masson, J.-Y. Interactions between BRCA2 and RAD51 for Promoting Homologous Recombination in Leishmania Infantum. *Nucleic Acids Research* **2012**, *40* (14), 6570–6584. https://doi.org/10.1093/nar/gks306.

(165)   Shivji, M. K. K.; Davies, O. R.; Savill, J. M.; Bates, D. L.; Pellegrini, L.; Venkitaraman, A. R. A Region of Human BRCA2 Containing Multiple BRC Repeats Promotes RAD51-Mediated Strand Exchange. *Nucleic Acids Research* **2006**, *34* (14), 4000–4011. https://doi.org/10.1093/nar/gkl505.

(166)   Jo, H.; Meinhardt, N.; Wu, Y.; Kulkarni, S.; Hu, X.; Low, K. E.; Davies, P. L.; DeGrado, W. F.; Greenbaum, D. C. Development of α-Helical Calpain Probes by Mimicking a Natural Protein–Protein Interaction. *J. Am. Chem. Soc.* **2012**, *134* (42), 17704–17713. https://doi.org/10.1021/ja307599z.

(167)   Bird, G. H.; Madani, N.; Perry, A. F.; Princiotto, A. M.; Supko, J. G.; He, X.; Gavathiotis, E.; Sodroski, J. G.; Walensky, L. D. Hydrocarbon Double-Stapling Remedies the Proteolytic Instability of a Lengthy Peptide Therapeutic. *Proceedings of the National Academy of Sciences* **2010**, *107* (32), 14093–14098. https://doi.org/10.1073/pnas.1002713107.

(168)   Hoover, D. M. DNAWorks: An Automated Method for Designing Oligonucleotides for PCR-Based Gene Synthesis. *Nucleic Acids Research* **2002**, *30* (10), 43e– 443. https://doi.org/10.1093/nar/30.10.e43.

(169)   Liebschner, D.; Afonine, P. V.; Baker, M. L.; Bunkóczi, G.; Chen, V. B.; Croll, T. I.; Hintze, B.; Hung, L.-W.; Jain, S.; McCoy, A. J.; Moriarty, N. W.; Oeffner, R. D.; Poon, B. K.; Prisant, M. G.; Read, R. J.; Richardson, J. S.; Richardson, D. C.; Sammito, M. D.; Sobolev, O. V.; Stockwell, D. H.; Terwilliger, T. C.; Urzhumtsev, A. G.; Videau, L. L.; Williams, C. J.; Adams, P. D. Macromolecular Structure Determination Using X-Rays, Neutrons and Electrons: Recent Developments in *Phenix*. *Acta Crystallogr D Struct Biol* **2019**, *75* (10), 861–877. https://doi.org/10.1107/S2059798319011471.

(170)   Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K. Features and Development of *Coot*. *Acta Crystallogr D Biol Crystallogr* **2010**, *66* (4), 486–501. https://doi.org/10.1107/S0907444910007493.

(171)   Smart, O. S.; Womack, T. O.; Flensburg, C.; Keller, P.; Paciorek, W.; Sharff, A.; Vonrhein, C.; Bricogne, G. Exploiting Structure Similarity in Refinement: Automated NCS and Target-Structure Restraints in It BUSTER. *Acta Crystallographica Section D* **2012**, *68* (4), 368–380. https://doi.org/10.1107/S0907444911056058.

(172)   Scopes, R. K. Measurement of Protein by Spectrophotometry at 205 Nm. *Analytical Biochemistry* **1974**, *59* (1), 277–282. https://doi.org/10.1016/0003-2697(74)90034-7.

(173)   Nikolovska-Coleska, Z.; Wang, R.; Fang, X.; Pan, H.; Tomita, Y.; Li, P.; Roller, P. P.; Krajewski, K.; Saito, N. G.; Stuckey, J. A.; Wang, S. Development and Optimization of a Binding Assay for the XIAP BIR3 Domain Using Fluorescence Polarization. *Analytical Biochemistry* **2004**, *332* (2), 261–273. https://doi.org/10.1016/j.ab.2004.05.055.

8 Appendix

8.1 RP-HPLC chromatograms and MS spectra of purified peptides

BRC2

Species	m/z calculated	m/z found
M+2H	2289.01	2289.8
M+3H	1526.34	1527.1
M+4H	1145.00	1145.6
M+5H	916.20	916.8

BRC4

Species	m/z calculated	m/z found
M+3H	1578.47	1579.1
M+4H	1184.10	1184.8
M+5H	947.48	947.9
M+6H	789.74	790.2

BRC8

BRC8 C18 RP-HPLC 220 nm

Peak ID	Compound	Time	Mass Found	BPM
6		1.36		1138

1:MS ES+
3.5e+007

Species	m/z calculated	m/z found
M+3H	1516.01	1516.4
M+4H	1137.26	1137.8
M+5H	910.01	910.4
M+6H	758.51	758.9

BRC8-2

BRC8-2 C18 RP-HPLC 220 nm

1:MS ES+
5.0e+007

Species	m/z calculated	m/z found
M+2H	2176.90	2177.2
M+3H	1451.60	1451.9
M+4H	1088.95	1089.4
M+5H	871.36	871.7

BRC8-2^{S2056V}

BRC8-2 ^{S2056V} **C18 RP-HPLC 220 nm**

1:MS ES+
5.0e+006

Species	m/z calculated	m/z found
M+3H	1455.62	1455.9
M+4H	1091.96	1092.6
M+5H	873.77	874.3

BRC8-2^{A1237V}

1:MS ES+
6.7e+007

Species	m/z calculated	m/z found
M+2H	2190.92	2191.0
M+3H	1460.95	1461.1
M+4H	1095.96	1096.4
M+5H	876.97	877.3

BRC8-2$^{S2056V,\ A1237V}$

1:MS ES+
1.2e+008

Species	m/z calculated	m/z found
M+2H	2196.95	2196.8
M+3H	1464.97	1464.7
M+4H	1098.98	1098.8
M+5H	879.38	879.3

142

BRC8-2[S2056A]

BRC8-2 [S2056A] **C18 RP-HPLC 220 nm**

1:MS ES+
5.2e+007

Species	m/z calculated	m/z found
M+2H	2168.43	2168.7
M+3H	1445.95	1446.6
M+4H	1084.71	1085.3

SP2

SP2 C18 RP-HPLC

1:MS ES+
7.4e+007

Species	m/z calculated	m/z found
M+2H	2289.97	2290.7
M+3H	1526.98	1527.5
M+4H	1145.48	1146.0
M+5H	916.59	917.0

143

SP24

SP24 C18 RP-HPLC 220 nm

1:MS ES+
1.4e+008

Species	m/z calculated	m/z found
M+2H	1483.72	1483.3
M+3H	989.48	989.3
M+4H	742.36	742.0
M+5H	594.095276	594.0

SP30

SP30 C18 RP-HPLC 220 nm

1:MS ES+
1.1e+008

Species	m/z calculated	m/z found
M+2H	1527.267276	1526.9
M+3H	1018.513943	1018.2
M+4H	764.137276	764.0
M+5H	611.511276	611.5

SP31

SP31 C18 RP-HPLC 220 nm

1:MS ES+
6.6e+007

* = TFA counterions

Species	m/z calculated	m/z found
M+4H	1144.08	1144.0
M+5H	915.46	915.3
M+6H	764.58	762.9
M+7H	654.84	654.1

SP32

1:MS ES+
9.4e+006

Species	m/z calculated	m/z found
M+4H	1340.48	1340.5
M+5H	1072.58	1072.7
M+6H	895.77	894.0
M+7H	767.18	766.3

Fluor-NCys-LiBRC1

Peak ID	Compound	Time	Mass Found	BPM
8		1.29		1047

1:MS ES+
6.3e+007

Species	m/z calculated	m/z found
M+3H	1396.24	1395.8
M+4H	1047.43	1047.2
M+5H	838.15	838.0
M+6H	700.02	698.4
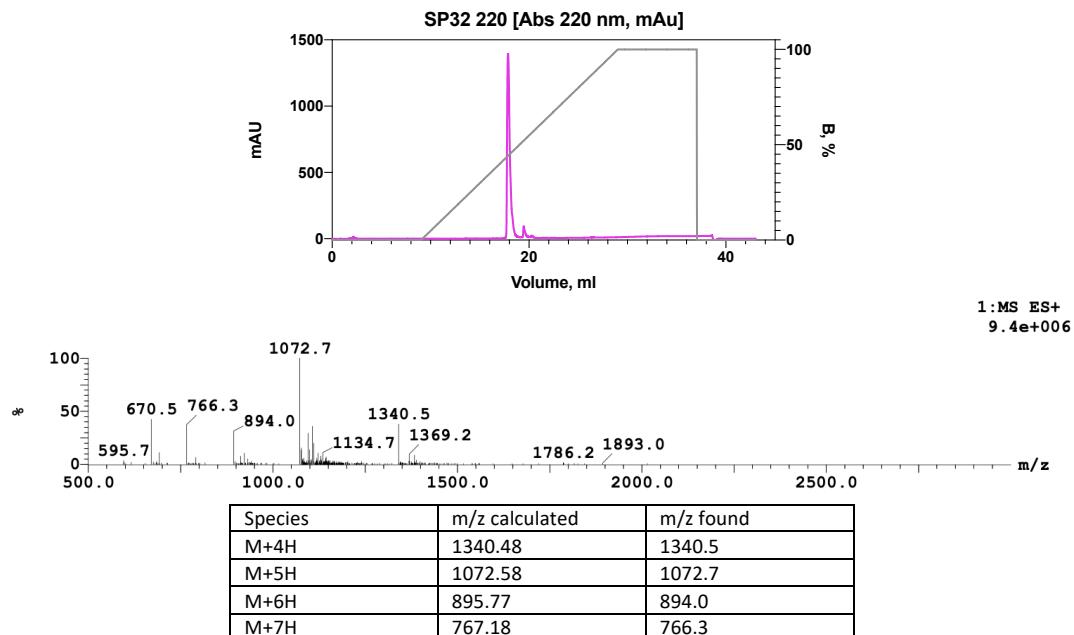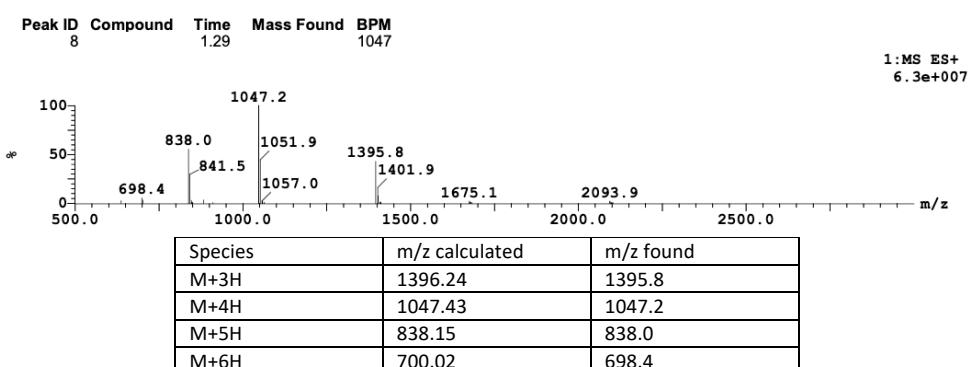
8.2 Representative protein LC-MS spectra of GB1-fused stapled peptides

GB1-SP2

Species	mass (calc), Da	mass (obs), Da
GB1-**SP2**-DVT-Ac (-Met1)	14773.97	14774.00
GB1-**SP2**-DVT-Ac (-Met1, -Ser2)	14686.89	14687.00
GB1-**SP2**-(DVT-Ac)$_2$ (-Met1)	14994.2	14994.5

GB1-SP10

Species	mass (calc), Da	mass (obs), Da
GB1-**SP10**-DVT-Ac (-Met1)	14251.46	14251.5
GB1-**SP10**-DVT-Ac (-Met1, -Ser2)	14164.38	14164.5
GB1-**SP10**-(DVT-Ac)$_2$ (-Met1)	14471.69	14471.5

GB1-SP11

Species	mass (calc), Da	mass (obs), Da
GB1-**SP11**-DVT-Ac (-Met1)	14210.4	14210.2
GB1-**SP11**-DVT-Ac (-Met1, -Ser2)	14123.33	14122.9
GB1-**SP11**-(DVT-Ac)₂ (-Met1)	14430.63	14430.4

GB1-SP12

Species	mass (calc), Da	mass (obs), Da
GB1-**SP12**-DVT-Ac (-Met1)	14500.64	14500.5
GB1-**SP12**-DVT-Ac (-Met1, -Ser2)	14720.87	14720.5
GB1-**SP12**-(DVT-Ac)₂ (-Met1)	14413.56	14413.5

GB1-SP15

Species	mass (calc), Da	mass (obs), Da
GB1-**SP15**-DVT-Ac (-Met1)	14760.97	14760.7
GB1-**SP15**-(DVT-Ac)₂ (-Met1)	14981.2	14981.2

8.3 Protein LCMS spectrum of *Li*BRC1:*Li*Rad51 complex

Species	m/z calculated	m/z found
M+5H (C1)	732.04	732.02
M+6H (C2)	610.20	610.18

8.4 Vector map for the pPEPT1 expression cassette

pPEPT1

```
                                                      XbaI
5′  TAATACGACTCACTATAGGGAATTGTGAGCGGATAACAATTCCTCTAGAAATAATTTTGTTTAAC
    ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++      65
3′  ATTATGCTGAGTGATATCCCTTAACACTCGCCTATTGTTAAGGAGATCTTTATTAAAACAAATTG
```
lac operator   RBS

T7 promoter

```
                          ClaI
                          BspDI
    TTTAAGAAGGAGATATATCCATGTATCGATTAAATAAGGAGGAATAAACCATGTCTGCTTGGAGC
    ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++     130
    AAATTCTTCCTCTATATAGGTACATAGCTAATTTATTCCTCCTTATTTGGTACAGACGAACCTCG
```
RBS   Met Tyr Arg Leu Asn Lys Glu Glu   Trp Ser
minicistron   Strep-Tag II

```
                        SpeI
    CACCCGCAGTTCGAGAAATCTGGCACTAGTGGCACCTACAAACTGATCCTGAACGGTAAAACCCT
    ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++     195
    GTGGGCGTCAAGCTCTTTAGACCGTGATCACCGTGGATGTTTGACTAGGACTTGCCATTTTGGGA
```
His Pro Gln Phe Glu Lys Ser Gly Thr Ser Gly Thr Tyr Lys Leu Ile Leu Asn Gly Lys Thr Leu
Strep-Tag II (in frame with Strep-Tag II) GB1

```
                        AccI
    GAAAGGTGAAACCACCACCGAAGCTGTAGACGCTGCTACTGCTGAAAAAGTTTTCAAACAGTACG
    ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++     260
    CTTTCCACTTTGGTGGTGGCTTCGACATCTGCGACGATGACGACTTTTTCAAAAGTTTGTCATGC
```
Lys Gly Glu Thr Thr Thr Glu Ala Val Asp Ala Ala Thr Ala Glu Lys Val Phe Lys Gln Tyr
GB1

```
    CTAACGACAACGGTGTGGACGGTGAATGGACCTACGACGACGCTACCAAAACCTTCACGGTTACG
    ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++     325
    GATTGCTGTTGCCACACCTGCCACTTACCTGGATGCTGCTGCGATGGTTTTGGAAGTGCCAATGC
```
Ala Asn Asp Asn Gly Val Asp Gly Glu Trp Thr Tyr Asp Asp Ala Thr Lys Thr Phe Thr Val Thr
GB1

```
    Acc65I       KpnI              BamHI    BsaI                      BsaI
    GAAAGCGGTACCGAAAACCTGTACTTCCAGGGATCCGACCCTGGAGACCCGAAGGTCTCTGACCC
    ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++     390
    CTTTCGCCATGGCTTTTGGACATGAAGGTCCCTAGGCTGGGACCTCTGGGCTTCCAGAGACTGGG
```
Glu Ser Gly Thr Glu Asn Leu Tyr Phe Gln Gly Ser Asp Pro Gly Asp Pro Lys Val Ser Asp Pro
GB1   TEV site   (in frame with TEV site)

```
    AvaI
    XhoI
    BsoBI
    PspXI
    PaeR7I   BmeT110I                      HindIII
    TGGCTCGAGCCATCACCATCACCATCACCATCACTAAGCTTGAGTATTCTATAGTGTCACCTAAA
    ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++     455
    ACCGAGCTCGGTAGTGGTAGTGGTAGTGGTAGTGGTAGTGATTCGAACTCATAAGATATCACAGTGGATTT
```
Gly Ser Ser His His His His His His His His   SP6 promoter
8xHis

- Unique cutters are shown in bold

8.5 Oligonucleotides for assembly PCR and cloning

Insert	Vector	Res. enzyme(s)	No	Sequence
BRC8-2^{S2056V}	pOP3BT	*BamHI/XhoI*	1	CCTGTACTTCCAGGGATCCGTTAACTCTTCTGCGTTCGTTGG
			2	GTAGAAACGTTCAGTTTTTTACCAGACGCGGTAGAGAAACCAACGAACGCAGAAGAGTTA
			3	CTGGTAAAAAACTGAACGTTTCTACCGAAGCGCTGCAGAAAGCGGTTAAACTGTTCTCTG
			4	CAAGCTTAGCTCGAGCCAGAGATGTTTTCGATGTCAGAGAACAGTTTAACCGCTTT
BRC8-2^{A1237V}	pOP3BT	*BamHI/XhoI*	1	CCTGTACTTCCAGGGATCCGTTAACTCTTCTGCGTTCTCCG
			2	GTAGAAACGTTCAGTTTTTTACCAGACGCGGTAGAGAAACCGGAGAACGCAGAAGAGTTA
			3	CTGGTAAAAAACTGAACGTTTCTACCGAAGCGCTGCAGAAAGTTGTTAAACTGTTCTCTG
			4	CAAGCTTAGCTCGAGCCAGAGATGTTTTCGATGTCAGAGAACAGTTTAACAACTTTCTG
BRC8-2$^{S2056V, A1237V}$	pOP3BT	*BamHI/XhoI*	1	CCTGTACTTCCAGGGATCCGTTAACTCTTCTGCGTTCGTTGG
			2	GTAGAAACGTTCAGTTTTTTACCAGACGCGGTAGAGAAACCAACGAACGCAGAAGAGTTA
			3	CTGGTAAAAAACTGAACGTTTCTACCGAAGCGCTGCAGAAAGTTGTTAAACTGTTCTCTG
			4	CAAGCTTAGCTCGAGCCAGAGATGTTTTCGATGTCAGAGAACAGTTTAACAACTTTCTG
BRC8-2^{S2056A}	pOP3BT	*BamHI/XhoI*	1	CCTGTACTTCCAGGGATCCGTTAACTCTTCTGCGTTCGCG
			2	GTAGAAACGTTCAGTTTTTTACCAGACGCGGTAGAGAAACCCGCGAACGCAGAAGAGTTA
			3	CTGGTAAAAAACTGAACGTTTCTACCGAAGCGCTGCAGAAAGCGGTTAAACTGTTCTCTG
			4	CAAGCTTAGCTCGAGCCAGAGATGTTTTCGATGTCAGAGAACAGTTTAACCGCTTT
SP1	pPEPT1	*BsaI*	1	TTCCAGGGATCCGACCCTTCTTGTGAAGCGCTGCAGAAGGCGTGTAAACTGTTCTC
			2	ATGGCTCGAGCCAGGGTCAGAGATGTTTTCGATGTCGGGAACAGTTTACACGCCTT
SP2	pOP3BT	*BamHI/XhoI*	1	CCTGTACTTCCAGGGATCCGTTAACTCTTCTGCGTTCTCCG
			2	CAGTTTTTTACCAGACGCGGTAGAGAAACCGGAGAACGCAGAAGAGTTAA
			3	CCGCGTCTGGTAAAAAACTGAACGTTTCTTGCGAAGCGCTGCAGAAAGCG
			4	CAAGCTTAGCTCGAGCCAGAGATGTTTTCGATGTCAGAGAACAGTTTGCACGCTTTCTGCAGCGCTT
	pPEPT1	*BsaI*	1	TTCCAGGGATCCGACCCTGTTAACTCTTCTGCGTTCTCCGGTTTCTCTACCGCGTCTGGTAAAAAACTGAACGTTTCTTGCGA
			2	ATGGCTCGAGCCAGGGTCAGAGATGTTTTCGATGTCAGAGAACAGTTTGCACGCTTTCTGCAGCGCTTCGCAAGAAACGTTCAGTTTT
SP7	pPEPT1	*BsaI*	1	TTCCAGGGATCCGACCCTGTTAACTCTTCTGCGTTCTCTGGTTTCTCTCACCGCGTCTGGTAAAAAACTGAACGTTTCTTGCGA
			2	ATGGCTCGAGCCAGGGTCAGAGATGTTTTCGATGTCGGGAACAGTTTAACCGCCTTCTGCGAGCGCTTCGCAAGAAACGTTCAGTTTT
SP8	pPEPT1	*BsaI*	1	TTCCAGGGATCCGACCCTGTTAACTCTTCTGCGTTCTCTGGTTTCTCTACCGCGTCTGGTAAAAAACTGAACGTTTCTACCTGTG
			2	ATGGCTCGAGCCAGGGTCAGAGATGTTTTCGATGTCGGGAACAGACAAACCGCCTTCTGCGACGCGCACAGGTAGAAACGTTCAGT
SP9	pPEPT1	*BsaI*	1	TTCCAGGGATCCGACCCTGTTAACTCTTCTGCGTTCTCTGGTTTCTCTACCGCGTCTGGTAAAAAACTGAACGTTTCTACCGAAG
			2	ATGGCTCGAGCCAGGGTCAGAGATGTTTTCGATGTCACAGAACAGTTTAACCGCCTTACACAGCGCTTCGGTAGAAACGTTCAGT
SP10	pPEPT1	*BsaI*	1	TTCCAGGGATCCGACCCTGCGTGTGGTTTCTCTACCGCGTCTGGTAAAAAACTGAACGTTTCTTGTGAAGCGCTGCAG
			2	ATGGCTCGAGCCAGGGTCAGAGATGTTTTCGATGTCGGAGAACAGTTTAACCGCCTTCTGCAGCGCTTCACAAGAAA
SP11	pPEPT1	*BsaI*	1	TTCCAGGGATCCGACCCTTCTTGTTTCTCTACCGCGTCTGGTAAAAAACTGAACGTTTCTTGTGAAGCGCTGCAGA
			2	ATGGCTCGAGCCAGGGTCAGAGATGTTTTCGATGTCGGGAACAGTTTAACCGCCTTCTGCAGCGCTTCACAAGAA
SP12	pPEPT1	*BsaI*	1	TTCCAGGGATCCGACCCTTCTTCTGCGTGTTCTGGTTTCTCTACCGCGTCTGGTAAAAAACTGAACGTTTCTACCGAAGCGT
			2	ATGGCTCGAGCCAGGGTCAGAGATGTTTTCGATGTCGGGAACAGTTTAACCGCCTTCTGACACGCTTCGGTAGAAACGTTC
SP13	pPEPT1	*BsaI*	1	TTCCAGGGATCCGACCCTGTTAACTGTTCTGCGTTCTCTGGTTTCTCTACCGCGTCTGGTAAAAAACTGAACGTTTCTACCGAAG
			2	ATGGCTCGAGCCAGGGTCAGAGATGTTTTCGATGTCGGGAACAGTTTAACCGCCTTACACAGCGCTTCGGTAGAAACGTTCAGT
SP14	pPEPT1	*BsaI*	1	TTCCAGGGATCCGACCCTGTTAACTGTTCTGCGTTCTCTGGTTTCTCTACCGCGTCTG
			2	ATGGCTCGAGCCAGGGTCAGAAACACACAGTTTTTTACCAGACGCGGTAGAGAAACCA
SP16	pPEPT1	*BsaI*	1	TTCCAGGGATCCGACCCTGCGTGTTCTGGTTTCTCTACCGCGTCTGGTAAAAA
			2	ATGGCTCGAGCCAGGGTCAGAACAAACGTTCAGTTTTTTACCAGACGCGGTAGA
SP17	pPEPT1	*BsaI*	1	TTCCAGGGATCCGACCCTGCGTGCTCTGGTTTCTCTACCGC
			2	TAGAAACGTTCAGTTTTTTACCAGACGCGGTAGAGAAACCAGAGC
			3	TCTGGTAAAAAACTGAACGTTTCTACCGAAGCGTGCTGCAAAGCG
			4	ATGGCTCGAGCCAGGGTCGTCGCAGAACAGTTTAACCGCTTTGCAGCACGCTT
SP19	pPEPT1	*BsaI*	1	TTCCAGGGATCCGACCCTGCGTGCGGTTTCTCTACC
			2	AGAAACGTTCAGTTTTTTACCAGACGCGGTAGAGAAACCGCACGC
			3	TCTGGTAAAAAACTGAACGTTTCTTGCGAAGCGCTGTGCAAAGCG
			4	ATGGCTCGAGCCAGGGTCGTCGCAGAACAGTTTAACCGCTTTGCACAGCGCTT
SP20	pPEPT1	*BsaI*	1	TTCCAGGGATCCGACCCTGCGTGCGGTTTCTCTACC
			2	AGAAACGTTCAGTTTTTTACCAGACGCGGTAGAGAAACCGCACGC

150

			3	TCTGGTAAAAAACTGAACGTTTCTACCGAAGCGTGCTGCAAAGCG
			4	ATGGCTCGAGCCAGGGTCGTCGCAGAACAGTTTAACCGCTTTGCAGCACGCTT
SP24-SNAC	pPEPT1	*BamHI/XhoI*	1	ACCTGTACTTCCAGGGATCCGGTTTCTCTACCGCGTCTGGT
			2	GCAGCGCCTGGCAAGAAACGTTCAGTTTTTTACCAGACGCGGTAGAGAAA
			3	CTTGCCAGGCGCTGCAGAAAGCGTGCAAACTGTTCTCTGGTAGCCATCAT
			4	TGGTGATGGTGATGGCTCGAGCCCCAATGATGGCTACCAGAGAACAG
SP30	pEXP-Nhis-GB1	*BsaI/HindIII*	1	TTCTAATACGACTCACTATAGGTACCGAAAACCTGTACTTCC
			2	GCCGCTCGCGGTGCTAAAGCCGCTGCAGCCCTGGAAGTACAGGTTTTCGGTACCTATAGTGAGTC
			3	CCGCGAGCGGCAAAAAACTGAACGTGAGCACCCAGGCGTGCCAGA
			4	CTATAGAATACTCAAGCTTAGCCGCTAAACAGTTTCACCGCTTTCTGGCACGCCTGGGT
SP31	pEXP-Nhis-GB1	*BsaI/HindIII*	1	GGTACCGAAAACCTGTACTTCCAGGGTCGTCGCCGTCGT
			2	AGACGCGGTAGAGAAACCAGAGCAACCACCACGACGACGACGACGACGGCGACGACC
			3	TGGTTTCTCTACCGCGTCTGGTAAAAAACTGAACGTTTCTACCCAGGCGTGCCAGAAAGC
			4	CTATAGAATACTCAAGCTTAACCAGAGAACAGTTTAACCGCTTTCTGGCACGCCTG
LiBRC1	pPEPT1	*BsaI*	1	TTCCAGGGATCCGACCCTCTGGTTCCGACCCTGTTCTCTACCGCGTCTGGTAAACCGGTGACCGTTCGTCGTG
			2	ATGGCTCGAGCCAGGGTCCGCCGCCAGGTCACCCAGACGTTCCGCAACTTTCTGCAGGGATTCACGACGAACGGTCACC
	pOP3BT	*BamHI/HindIII*	1	GAAAACCTGTACTTCCAGGGATCCCTGGTTCCGACCCTGTTCT
			2	ACGAACGGTCACCGGTTTACCAGACGCGGTAGAGAACAGGGTCGGAACCA
			3	AACCGGTGACCGTTCGTCGTGAATCCCTGCAGAAAGTTGCGGAACGTCTG
			4	GGTGACACTATAGAATACTCAAGCTTAGGCGGCCAAGTCACCCAGACGTTCCGCAACTTTCT
NCys-LiBRC1	pOP3BT	*BamHI/HindIII*	1	AACCTGTACTTCCAGGGATCCTGCCGTGTTCCGACCCTGTTC
			2	TGAACGGTAACCGCTTTACCACGACCGGTTTCGAACAGGGTCGGAACACG
			3	GGTAAAGCGGTTACCGTTCAGAAACGTTCTCTGGACAAAGCGGAAGCGTC
			4	ACTATAGAATACTCAAGCTTAGGCTTCTAAAGAGTCCATAGACGCTTCCGCTTTGTCC
LiBRC1.1	pPEPT1	*BsaI*	1	TTCCAGGGATCCGACCCTCTGGTTCCGACCCTGTTCTCTACCGGTTCTGGTAAACCGGTGACCGTTCGTCGTG
			2	ATGGCTCGAGCCAGGGTCCGCCGCCAGGTCACCCAGACGTTCCGCAACTTTCTGCAGGGATTCACGACGAACGGTCACC
LiBRC1.2	pPEPT1	*BsaI*	1	TTCCAGGGATCCGACCCTCTGGTTCCGACCCTGTTCTCTACCGCGTCTGGTAAGCCGGTTACCGTTCGTCGTGAATCTCTC
			2	ATGGCTCGAGCCAGGGTCTTCCATATCAGGTGCGGCCAGGTCACCCAGACGTTCCGCAACTTTCTGGAGAGATTCACGACGAACGG
LiBRC1.3	pPEPT1	*BsaI*	1	TTCCAGGGATCCGACCCTGCTGCGGCTCCGCCGCTGGTTCCGACCCTGTTCTCTACCGCGTCTGGTAAGCCGGTTACCGTTCG
			2	ATGGCTCGAGCCAGGGTCCGCCGCCAGGTCACCCAGACGTTCCGCAACCTTTTGCAGAGATTCACGACGAACGGTAACCGGCTTAC
LiBRC1.4	pPEPT1	*BsaI*	1	TTCCAGGGATCCGACCCTCTGGTTCCGACCCTGTTCTCTACCGCGTCTGGTAAACCGGTGACCGTTC
			2	ATGGCTCGAGCCAGGGTCGTCACCCAGACGTTCCGCAACTTTCTGCAGGGATTCACGACGAACGGTCACCGGTTTACC
LiBRC1.5	pPEPT1	*BsaI*	1	TTCCAGGGATCCGACCCTCTGGTTCCGACCCTGTTCTCTACCGCGTCTGGTAAACCGGTGACCGTTCGTCGTG
			2	ATGGCTCGAGCCAGGGTCAGAACCACCACCCAGACGTTCCGCAACTTTCTGCAGGGATTCACGACGAACGGTCACCG
LiBRC1.6	pPEPT1	*BsaI*	1	TTCCAGGGATCCGACCCTCTGGTTCCGACCCTGTTCTCTACCGCGTCTGGTAAACCGGTGACCGTTCGTCGTG
			2	ATGGCTCGAGCCAGGGTCAGAGCCACCACCACGTTCCGCAACCTTCTGCAGGGATTCACGACGAACGGTCACCGGTT
LiBRC1.7	pPEPT1	*BsaI*	1	TTCCAGGGATCCGACCCTGCTGCGGCTCCGCCGCTGGTTCCGACCCTGTTCTCTACCGCGTCTGGTAAGCCGGTTACCGTTCG
			2	ATGGCTCGAGCCAGGGTCCGCCGCCAGGTCACCCAGACGTTCCGCAACTTTTGCAGAGATTCACGACGAACGGTAACCGGCTTAC
LiBRC1.8	pPEPT1	*BsaI*	1	TTCCAGGGATCCGACCCTCTGGTTCCGACCCTGTTCTCTACCGCGTCTGGTAAGCCGGTTACCGTTCGT
			2	ATGGCTCGAGCCAGGGTCAGAGCCACCGCCACCCGCAACTTTCTGCAGGGATTCACGACGAACGGTAACCGGCTTACC
LiBRC1.9	pPEPT1	*BsaI*	1	TTCCAGGGATCCGACCCTCTGGTTCCGACCCTGTTCTCTACCGCGTCTGGTAAGCCGGTTACCGTTCGT
			2	ATGGCTCGAGCCAGGGTCAGAGCCACCGCCACCAACTTTCTGGAGAGATTCACGACGAACGGTAACCGGCTTACCAG
LiBRC1.10	pPEPT1	*BsaI*	1	TTCCAGGGATCCGACCCTCTGGTTCCGACCCTGTTCT
			2	ACGAACGGTCACCGGTTTACCAGACGCGGTAGAGAACAGGGTCGGAACCA
			3	AACCGGTGACCGTTCGTCGTGAATCTCTCCAGAAAGGTGGCTCTGGTGGT
			4	ATGGCTCGAGCCAGGGTCGGAACCACCAGAGCCACCTTT
LiBRC1.11	pPEPT1	*BsaI*	1	TTCCAGGGATCCGACCCTCTGGTTCCGACCCTGTTCT
			2	ACGAACGGTCACCGGTTTACCAGACGCGGTAGAGAACAGGGTCGGAACCA
			3	AACCGGTGACCGTTCGTCGTGAATCTCTCCAGGGTGGCTCTGGTGGTTCC
			4	ATGGCTCGAGCCAGGGTCGGAACCACCAGAGCCACC
LiBRC1.12	pPEPT1	*BsaI*	1	TTCCAGGGATCCGACCCTCTGGTTCCGACCCTGTTCTCTACCGCG
			2	ATGGCTCGAGCCAGGGTCACGAACGGTTACCGGTTTACCAGACGCGGTAGAGAACAGGGT
LiBRC2	pPEPT1	*BsaI*	1	TTCCAGGGATCCGACCCTCGTGTTCCGACCCTGTTCGAAACCGGTCGGTGGTAAAGCGGTTACCGTTCAGAAACGTTC
			2	ATGGCTCGAGCCAGGGTCCGCTTCCAGAGAGTCCATAGACGCTTCCGCTTTGTCGAGAGAACGTTTCTGAACGGTAACC
LiBRC2.1	pPEPT1	*BsaI*	1	TTCCAGGGATCCGACCCTCGTGTTCCGACCCTGTTCGAAACCGCGCGTGGTAAAGCGGTTACCGTTCAGAAACGTTC
			2	ATGGCTCGAGCCAGGGTCCGCTTCCAGAGAGTCCATAGACGCTTCCGCTTTGTCGAGAGAACGTTTCTGAACGGTAACC

LiRad51 (full-length)	pEXP-MBP	*BsaI/HindII*	1	GAAAACCTGTACTTCCAGTCCGGACAGACCCGTAGCAAAGC
			2	ATAGAATACTCAAGCTTATTAGTCACGTGCGTCACCAAC
LiRad51^{ATPase}	pHAT2	*NcoI/HindIII*	1	CCATCACCATCACTCCATGGCAGAAATTATCATGGTTACCACCGGTAG
			2	ACTATAGAATACTCAAGCTTAGTCACGTGCGTC
LiRad51^{ATPase,ΔL2}	pBAT4	*NcoI/HindIII*	1	AAGGAGATATATCCATGGCAGAAATTATCATGGTTACCACCGGTAG
			2	ACTATAGAATACTCAAGCTTAGTCACGTGCGTC
			3	GTTGTTGCCAATGGTGGTCATATTATGGCACATGCCAGC (inner F)
			4	CATAATATGACCACCATTGGCAACAACCTGATTGGTAACAACAAC (inner R)

8.6 *Li*Rad51 synthetic gene

```
>LiRAD51
AAGAAGGAGATATATCCATGGGACAGACCCGTAGCAAAGCAAAAGGTCGTCGTGGTCGTCCGAGCGCACGTCCGAGTGAAGAAGTTGAAGTTGT
TGAAAGCCAGCCGCAAGAAGCACTGCAGAACGAAGAACAAGAACCGCGTCAGCAGCAGCAACAGAGCACCGATATGGCAGAACCGAATGCAAGC
GGTTTTCGTGTTATTCAGATTCTGGAAAATTATGGTGTTGCCAGCAGCGACATCAAAAAACTGATGGAATGTGGTTTCTATACCGTGGAAAGCG
CAGCATATGCACCGAAAAAAGCAATTCTGGCAGTTAAAGGCATCAGCGAAATAAAGCCGAAAAAATCATGGCCGAATGTGCAAAACTGGTTCC
GATGGGTTTTACCAGCGCAGTTGCATATCACGAAGCACGTAAAGAAATTATCATGGTTACCACCGGTAGCCGTGAAGTTGATAAACTGTTAGGT
GGTGGTATTGAAACCGGTAGCATTACCGAACTGTTTGGTGAATTTCGTACCGGCAAAACCCAGCTGTGTCATACCCTGTGTGTTACCTGTCAGC
TGCCGATTAGCCAAGGTGGTGCAGAAGGTATGGCACTGTATATTGATACCGAAGGCACCTTTCGTCCGGAACGTCTGGTTGCAGTTGCAGAACG
TTATAAACTGGACCCGGAAGATGTTCTGGCAAATGTTGCATGTGCACGTGCCTTTAATACCGATCATCAGCAACAACTGCTGCTGCAGGCAAGC
GCAATGATGGCAGAAAATCGTTTTGCACTGATTGTTGTTGATAGCGCAACCGCACTGTATCGTACCGATTATAGCGGTCGTAATGAACTGGCAG
CACGTCAGATGCACCTGGGTAAATTTCTGCGTAGCCTGCATAATCTGGCCGAAGAATATGGTGTGGCCGTTGTTGTTACCAATCAGGTTGTTGC
CAATGTTGATGGTAGCGCACAGATGTTTCAGGCAGATAGCAAAAAACCGATTGGTGGTCATATTATGGCACATGCCAGCACCACACGTCTGAGC
CTGCGTAAAGGTCGTGGTGAACAGCGTATTATCAAAGTTTATGATAGCCCGTGTCTGGCAGAAGCAGAAGCAATTTTTGGTATTTATGATGATG
GTGTTGGTGACGCACGTGACTAAGCTTGAGTATTCTATAGT
```

8.7 Crystallographic conditions and data collection/refinement statistics

Complex	BRC8-2:HumRadA22	SP2:HumRadA22	SP24:HumRadA22
Protein	0.5 mM BRC8-2:HumRadA22 in 20 mM CHES pH 9.5, 100 mM NaCl, 20 mM ADP/MgCl$_2$	0.5 mM SP2:HumRadA22 in 20 mM CHES pH 9.5, 100 mM NaCl	0.5 mM SP24:HumRadA22 in 20 mM CHES pH 9.5, 100 mM NaCl, 20 mM ADP/MgCl$_2$
Condition	0.2 M NH4Cl (Salt) 20 %w/v PEG 3350 (Precipitant)	8 % w/v PEG 8000 (precipitant) 0.08 M Potassium phosphate pH 5.6 (buffer)	0.1 M Na$_3$Cit 4.2 pH (Buffer) 20 %w/v PEG 1K (Precipitant) 0.2 M Li$_2$SO4 (Salt)
Protein:condition (nl:nl)	200:200	200:200	400:200
PDB	6HQU	-	-
Data collection processing			
Beamline	DLS i03	DLS i03	DLS i04
Wavelength (Å)	0.9762	0.9762Å	0.9795
Space group	P 1 211	P 41 21 2	P 21 21 2
a, b, c (Å)	114.403 75.473 114.792	112.56 112.56 140.79	140.46 38.64 43.53
α, β, γ (°)	90.00 97.06 90.00	90.00 90.00 90.00	90.00 90.00 90.00
Resolution range (high resolution bin) (Å)	85.87 - 1.97 (1.97-1.96)	87.92 - 3.02 (3.02 - 3.07)	70.23 - 1.63 (1.66 - 1.63)
R$_{meas}$	0.091 (2.778)	0.242 (3.751)	0.059 (4.504)
Completeness (%)	99.8 (100.0)	100.0 (100.0)	98.7 (99.2)
reflections	522829 / 137994	170095 / 13512	239726 / 30303
Redundancy	3.8 (3.9)	12.8 (13.8)	7.9 (7.8)
$<I/\sigma(I)>$	7.4 (0.5)	10.0 (0.9)	12.0 (0.3)
CC$_{1/2}$	0.998 (0.272)	1.0 (0.3)	1.0 (0.3)
Refinement			
R$_{cryst}$/R$_{free}$	0.263 / 0.271	0.228 / 0.254	0.260 / 0.270
Resolution range (Å)	85.870 - 1.966	69.29 - 3.02	70.23 - 1.62
set	137876/ 6853	17515 / 897	27946 / 1450
Number of atoms	13996	3834	2041
Mean/Wilson B-factor	68.1 / 45.5	95.6 / 79.5	63.3 / 34.4
favoured/allowed/outliers (%)	99.00 / 1.00 / 0.00	97.22 / 2.57 / 0.21	99.17 / 0.83 / 0.00
RMSD bonds (Å)	0.011	0.017	0.013
RMSD angles (°)	1.28	1.66	1.69

Complex	SP30:HumRadA22	*Li*BRC1:*Li*Rad51
Protein	0.5 mM SP30:HumRadA22 in 20 mM CHES pH 9.5, 100 mM NaCl, 20 mM ADP/MgCl$_2$	0.5 mM LiBRC1 :LiRad51$^{ATPase,\Delta L2}$ in 20 mM Tris pH 8.0, 100 mM NaCl, 100 mM Li2SO4, 20 mM ADP/MgCl$_2$
Condition	14% w/v PEG 4000 (precipitant), 6% v/v MPD (precipitant), 0.1M Na K Phos pH 6.2 (buffer)	32% low MW PEG smear (precipitant, Molecular Dimensions), 0.1M Tris pH 8.5
Protein:condition (nl:nl)	200:200	200:200
PDB	-	-
Data collection processing		
Beamline	DLS	DLS i04-1
Wavelength (Å)	0.9795	0.9159
Space group	P 21 21 2	P 4 2 2
a, b, c (Å)	143.13 38.01 43.92	61.00 61.00 119.22
α, β, γ (°)	90.00 90.00 90.00	90.00 90.00 90.00
Resolution range (high resolution bin) (Å)	1.25 - 41.99 (1.25 - 1.27)	2.15 - 59.61 (2.15 - 2.18)
R_{meas}	0.046 (1.597)	0.119 (8.167)
Completeness (%)	99.6 (95.3)	98.9 (98.9)
reflections	513195 / 67481	296871 / 12826
Redundancy	7.6 (4.6)	23.1 (18.1)
$<I/\sigma(I)>$	15.8 (0.8)	15.6 (0.4)
$CC_{1/2}$	1.0 (0.4)	1.0 (0.5)
Refinement		
R_{cryst}/R_{free}	0.196 / 0.210	0.211 / 0.267
Resolution range (Å)	36.74 - 1.24	54.31 - 2.15
work/test set	64107 / 3361	10597 / 517
Number of atoms	2373	1806
Mean/Wilson B-factor	27.9 / 17.8	67.897 / 50.7
favoured/allowed/outliers	99.15 / 0.85 / 0.00	98.65 / 1.35 / 0.00
RMSD bonds (Å)	0.014	0.012
RMSD angles (°)	1.61	1.62