



HAL
open science

Classical molecular dynamics simulations of the complex between the RAD51 protein and the BRC hairpin loops of the BRCA2 protein

Nick Buis, Chris-Kriton Skylaris, Guy Grant, Eeson Rajendra, Mike C Payne, Ashok R. Venkitaraman

► To cite this version:

Nick Buis, Chris-Kriton Skylaris, Guy Grant, Eeson Rajendra, Mike C Payne, et al.. Classical molecular dynamics simulations of the complex between the RAD51 protein and the BRC hairpin loops of the BRCA2 protein. *Molecular Simulation*, 2008, 34 (08), pp.749-759. 10.1080/08927020802213281 . hal-00515042

HAL Id: hal-00515042

<https://hal.science/hal-00515042>

Submitted on 4 Sep 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

**Classical molecular dynamics simulations of the complex
between the RAD51 protein and the BRC hairpin loops of the
BRCA2 protein**

Journal:	<i>Molecular Simulation/Journal of Experimental Nanoscience</i>
Manuscript ID:	GMOS-2008-0073
Journal:	Molecular Simulation
Date Submitted by the Author:	08-Mar-2008
Complete List of Authors:	Buis, Nick; University of Southampton, Chemistry Skylaris, Chris-Kriton; University of Southampton, Chemistry Grant, Guy; University of Cambridge, Chemistry Rajendra, Eeson; University of Cambridge, Oncology Payne, Mike; University of Cambridge, Physics Venkitaraman, Ashok; University of Cambridge, Oncology
Keywords:	molecular dynamics, proteins, BRCA2, RAD51

SCHOLARONE™
Manuscripts

Classical molecular dynamics simulations of the complex between the RAD51 protein and the BRC hairpin loops of the BRCA2 protein

Nick Buis^{a,b}, Chris-Kriton Skylaris^a, Guy H. Grant^b, Eeson Rajendra^c, Mike C. Payne^d, and Ashok R. Venkitaraman^c

^a School of Chemistry, University of Southampton, Highfield, Southampton, SO17 1BJ, UK.

^b Unilever Centre for Molecular Informatics, Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge, CB2 1EW, UK.

^c Hutchison/MRC Research Centre, University of Cambridge, Hills Road, Cambridge CB2 2XZ, UK.

^d TCM Group, Cavendish Laboratory, University of Cambridge, J. J. Thomson Avenue, Cambridge CB3 0HE, UK.

27 February 2008

Abstract:

In the repair of double-strand breaks of DNA by homologous recombination the recombinase protein RAD51 has its functions controlled by the breast cancer susceptibility protein BRCA2. BRCA2 can bind to RAD51 via the BRC repeats BRC1-BRC8, which are eight conserved sequence motifs in BRCA2 of about 35 amino acids. We have carried out a series of extensive unrestrained atomistic molecular dynamics (MD) simulations in explicit water for a total time period of 248 ns, in order to study the dynamical behaviour and conformations of the complexes between the hairpin loop region of the BRC repeats and RAD51. Our simulations have allowed us to investigate the conformations adopted by the BRC repeats both while bound to RAD51 and while isolated. These conformations are rationalised through an analysis of the inter- and intra-molecular backbone and side chain bonding interactions in all the eight human BRC repeats as well as in a single-point mutation of BRC4. The differences in sequence result in differences in the interactions between the BRC repeats and the RAD51 protein but these do not appear to disrupt the binding in any of the BRC-RAD51 complexes as there are always a number of key residues remaining which allow a sufficient number of interactions to stabilise the complexes.

1 Introduction

The preservation of genomic integrity is an essential part of successful cell proliferation (1). DNA double-strand breaks (DSB) can occur in cells as a result of endogenous and exogenous sources but cells have evolved repair processes to repair such lesions. If the resulting lesions are not repaired this leads to a predisposition to diseases, including cancer, and for this reason it is vital to understand mechanisms of repairing DNA damage. The most important mechanism of DSB repair is the process of homologous recombination (HR) that uses an intact homologous DNA molecule as a template for repair of the broken molecule. Unlike other methods of repair, this process is error-free and thus is essential for maintaining genomic integrity (2).

HR is a process conserved in all organisms and the core proteins involved in HR are similarly conserved. The key protein in eukaryotic HR is the recombinase RAD51, a member of the RecA

family of recombinases that includes the RadA protein in archaea. In conjunction with various accessory and regulatory proteins, RAD51 mediates homology-directed repair of DSBs (3). RAD51 is composed of two globular domains (4). The C-terminal core domain (amino acids 98-339 in human RAD51), accounts for many of its catalytic properties and is important for oligomerisation and binding to single-stranded DNA, found at the sites of a DSB. This domain is homologous to the core domain of *Escherichia coli* RecA (5). The N-terminal domain (amino acids 1-82) is anchored to the core domain via a flexible linker, and has putative double-stranded DNA binding ability (6).

The functions of RAD51 in the cell are known to depend on the interplay with other proteins and *in vivo* studies have shown that in many organisms, the BRCA2 protein is the key regulator of RAD51 function (7). BRCA2 mutations are associated with a strong predisposition to breast and ovarian cancer and functionally correlate with defective HR resulting in a reduced efficiency and fidelity of DSB repair (1,8); coupled with the compensatory use of error-prone pathways of repair, the result is gross chromosomal instability (2,9). BRCA2 uses two distinct motifs to interact with RAD51: the BRC repeat and a motif in the C-terminus of the protein (10). The BRC repeat is the primary site through which BRCA2 binds RAD51. This 35-40 amino acid motif interacts with RAD51 and allows BRCA2 to deliver the bound RAD51 to sites of damaged DNA. It is known that in this form RAD51 catalyses ATP-dependent homologous recombination which includes the process of pairing and strand exchange for the repair of DNA (3,11). Further study of precisely how BRCA2 directs the loading of RAD51 to DNA is needed to better understand how the genome is maintained in a stable state and how defects in these processes lead to cancer predisposition.

The multi-domain BRCA2 protein has orthologues in numerous organisms (but not yeast, prokaryotes and archaea – although all have RAD51 orthologues) and has little sequence conservation with other proteins (12,13). A number of distinct conserved domains have been characterised in these orthologues and have been the subject of genetic, biochemical and structural study, most notably the RAD51-interaction motifs (13). The complex nature of the interactions between BRCA2 and RAD51 has prompted a structure-based approach in its analysis.

BRC1	1002-NHSFGGSFRTASNKEIKLSEHNKSKSMFFKDIEE-1036
BRC2	1212-NEVGFRGFYSAHGTKLNVSTEALQKAVKLFSDIEN-1246
BRC3	1421-FETSDTFFQTASGKNISVAKESFNKIVNFFDQKPE-1455
BRC4	1517-KEPTLLGFHTASGKVKIAKESLDKVKNLFDEKEQ-1551
BRC5	1664-IENSALAFYTSCSRKTSVSQTSLLLEAKKWLREGIF-1698
BRC6	1837-FEVGPPAFRIASGKIVCVSHETIKKVKDIFTDSFS-1871
BRC7	1971-SANTCGIFSTASGKSVQVSDASLQARQVFSEIED-2005
BRC8	2051-NSSAFSGFSTASGKQVSILESSLHKVKGVLLEFFDL-2085
BRC4TA	1517-KEPTLLGFHAASGKVKIAKESLDKVKNLFDEKEQ-1551
	* : . : : : :

Figure 1. Sequence alignment of the eight human BRC repeats and a BRC4TA mutation. The core 35 amino acid sequence is shown with the β -hairpin loop region used in our MD simulations (see section 2) highlighted in red and the FX₁X₂A motif underlined. The alignment was generated with ClustalW. The symbols on the bottom row denote the degree of conservation observed in each column: "*" denotes that the residues or nucleotides in that column are identical in all sequences in the alignment, ":" denotes that conserved substitutions have been observed and "." denotes that semi-conserved substitutions are observed.

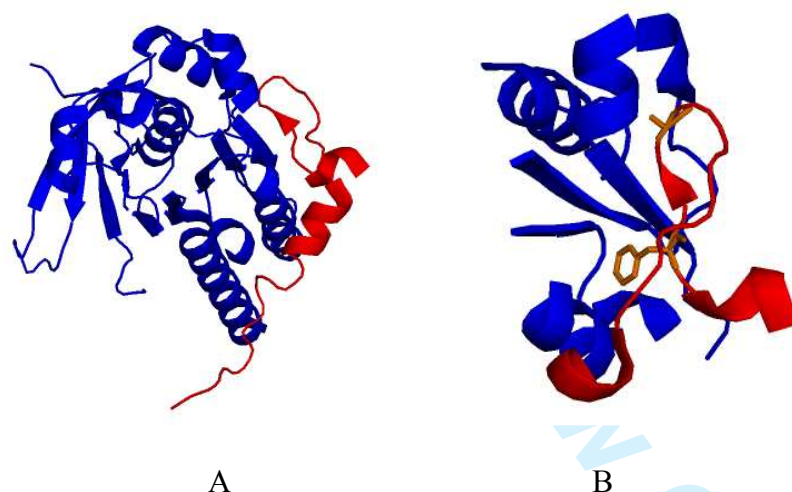
The core BRC repeat, albeit with significant sequence variation, is present in all orthologues but interestingly the number of these repeats varies between organism from one in *Ustilago maydis* and *Caenorhabditis elegans* to fifteen in *Trypanosoma brucei* (12). Human BRCA2 has 3418 amino acids and contains eight BRC repeats (BRC1-8). Each core repeat has 35 amino acids (Figure 1).

The crystal structure of the human RAD51 C-terminal core domain, in complex with BRC repeat 4 (BRC4) is shown in Figure 2 (14). This structure, at a 1.7 Å resolution shows that BRC4 makes contact with the core domain of RAD51. The binding of BRC4 to RAD51 extends the core parallel β -sheet structure of RAD51 by two short anti-parallel strands and also packs an amphipathic α -helix (amino acids 1536–1542) against the RAD51 exterior. This continuous contact between BRC4 and RAD51 is brought about by a strip of 28 amino acids (1521-1548) of BRC4 which include

1
2
3 hydrophobic contacts as well as a number of charge–charge interactions along the interface (12,14).
4 The crystal structure of the human BRC4–RAD51 complex provides a starting point for our study of
5 the interactions between the different BRC repeats and RAD51.
6

7 The crystal structure of the BRC4–RAD51 complex provided the first mechanistic insight into
8 how a BRC repeat binds RAD51 and subsequently can regulate its function. Within BRC4 the highly
9 conserved BRC motif denoted by the consensus FX_1X_2A (1524–FHTA–1527) was shown to bind to the
10 oligomerisation surface of RAD51 (strongly conserved in many RecA orthologues) (7,14,15). The
11 FHTA motif in BRC4 appears stabilised within the context of a β -hairpin loop when bound to RAD51.
12 F1524 and A1527 bind into two hydrophobic pockets in the core domain of RAD51. BRC repeats in
13 BRCA2 have clearly evolved to mimic the interaction seen between RAD51 protomers in a sequence-
14 specific manner (human RAD51 uses the motif 86–FTTA–89 in its inter-domain flexible linker to bind
15 to a neighbouring RAD51) and structural mutations in the FX_1X_2A consensus abrogate binding to
16 RAD51 (7,16).
17

18 *In vitro* studies suggest that the eight human BRC repeats are non-equivalent in their ability to
19 bind to RAD51 (13,16). The BRC repeats BRC5 and BRC6 are known to be inefficient binders
20 whereas BRC3 and BRC4 are good at binding RAD51 (10). With the determination of the atomic
21 structure of BRC4 in a complex with the catalytic core of RAD51, it has been possible to expand our
22 knowledge of the molecular features of the binding and identify the key residues for the interaction
23 between BRCA2–RAD51. However, crystal structures of the other BRC repeats either bound or
24 unbound are still unavailable, considerably limiting efforts to increase our understanding of the
25 interactions and the conformational changes which control binding.
26



51
52
53
54
55
56
57
58
59
60

Figure 2. The crystal structure of the human BRC4–RAD51 complex. (A) Cartoon diagram illustrating the secondary structure elements with RAD51 in blue and the BRC4 repeat in red. The BRC4 extends the core β -pleated sheet, and forms an orthogonal clamp on RAD51. (B) Hydrophobic interactions between RAD51 and the β -hairpin loop of BRC4. The side chains of F1524 and A1527 (shown in orange) of BRC4 are inserted into separate hydrophobic pockets.

In order to advance further our understanding of the conformations and BRCA2–RAD51 interactions in the region of the hairpin loop of the different repeats in solution and ambient temperature we have carried out a series of Molecular Dynamics (MD) simulations. We have studied each of the eight BRC repeats of human BRCA2 both while bound to the core domain of RAD51 and while unbound in solution and focusing on the region of strong sequence conservation and functional importance – the hairpin loop. We have also studied a structural point mutation in BRC4 (T1526A) which has been predicted to weaken the binding of BRC4 to RAD51 (16). In addition to the residues involved in direct inter-molecular interactions between the BRCA2 and RAD51, we also investigated the residues responsible for stabilising the observed conformations of the BRC peptides through intra-molecular bonding. In Section 2 we describe the methods and procedures used for setting up and running the simulations. In Section 3 we present and discuss our results including a detailed analysis

of the various inter- and intra-molecular interactions for the complex and the isolated BRC repeats. Finally, our conclusions are presented in section 4.

2 Methods

2.1 Initial structures

Proteins with the sequences of the eight human BRC repeats and the single point mutation BRC4TA were modelled starting from the crystal structure of the BRC4–RAD51 complex (14) (Protein Data Bank (PDB) code: 1n0w). The BRC4–RAD51 crystal structure was used as the starting point for generating the structures of the other complexes. This was done with the IC PARA command in CHARMM (17) that replaced the residues for the different sequences (as shown in Figure 1) based on the BRC4 backbone structure after editing the residue names in the original PDB file according to each repeat. The models for the simulations of the unbound BRC repeats were also built from the same crystal structure of the BRC4–RAD51 (14) complex by removing the residues corresponding to RAD51. The BRC repeat sequence was truncated in each case to the minimal structure which contained the hairpin loop. For BRC4 this included the 17 residues from PRO1519 to ILE1535, as shown in the sequences of Figure 1, where the 17 residue sequences for all of the repeats are highlighted. The backbone positions of the eight BRC repeats were built by interpolations between the closest known residue positions from the crystal structure of BRC4. The unknown side-chain heavy atom coordinates were built using the generate seg setup and ic build modules of CHARMM. Hydrogen atom coordinates were constructed using the HBUILD module of CHARMM (18), followed by a short geometry optimisation with the heavy atom coordinates kept fixed. Then a full structural relaxation of all the newly replaced residues was carried out, while the coordinates of the residues which were the same as those in the crystal structure of BRC4 were kept fixed.

2.2 Simulation set-up and equilibration

To allow a realistic description of the solvent, all our simulations were carried out in explicit water described by the TIP3P model (19). To avoid the high computational cost of simulations in periodic boundary conditions with explicit solvent a stochastic boundary approach was used to confine the water in a sphere centred at the protein-protein interface. This approach does not have a charge neutrality requirement as in the case of calculations with periodic boundary conditions. In our calculations we assumed that all Asp and Glu residues were negatively charged while Lys and Arg residues were positive. The His residues were represented in their neutral form. To prepare each simulation, each protein was first immersed in a 25Å radius sphere of pre-equilibrated water and any water molecules overlapping with the protein were deleted. This resulted in about 4,000 water molecules being included in the simulation when the initial PSF file was created. The sphere was then divided into two regions (Figure 3): an inner 23Å radius *reaction region* where Newtonian dynamics apply and an outer spherical shell defined by radii from 23Å to 25Å which acted as the *buffer region* where the molecular dynamics follows Langevin's equation of motion. A spherical boundary potential was employed to maintain the correct average distribution of water molecules and to prevent water from escaping into the vacuum. All the atoms in the reservoir region outside of the boundary potential were fixed as the system was partitioned. Before each simulation, the entire system (protein plus water) within the sphere was relaxed to remove bad contacts. An initial dynamics run was performed to shake-up the waters with the protein fixed, at 1000K for 2.5ps followed by an equilibration simulation of the whole system at 300K for 5ps. A new pre-equilibrated water sphere was then overlaid on the system to fill any holes that might have been created during the shake-up. A final equilibration dynamics run was performed with the protein fixed for 2.5ps at 300K and then on the whole system at 300K for 112.5ps. Then the production runs followed, which were carried out at 300

K. All the simulations were performed using the CHARMM (18) package with the version-22 force fields (17) and a time step of 1 fs.

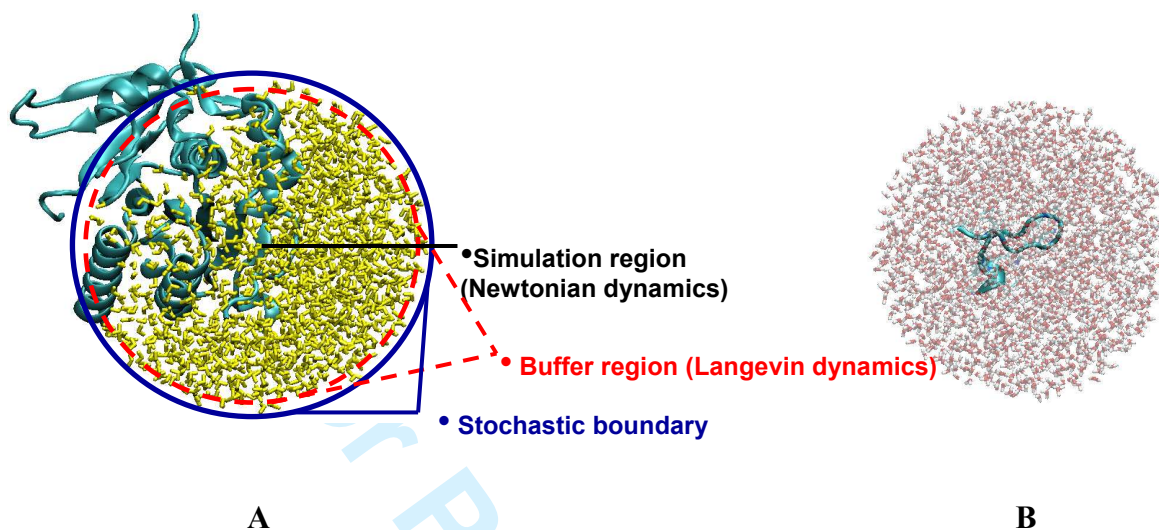
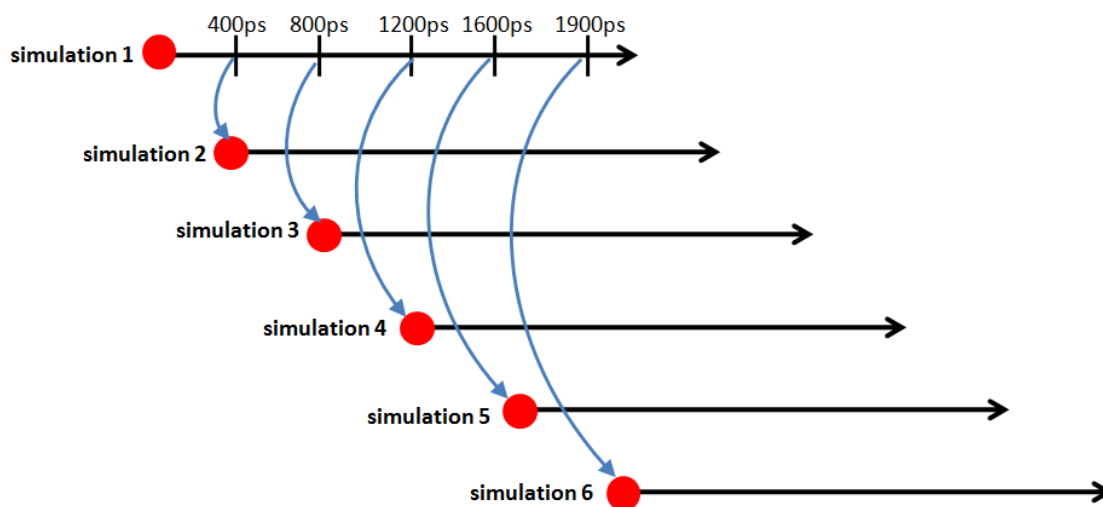


Figure 3. A) The human BRC4–RAD51 complex (cyan) inside the water sphere (yellow) which is divided into two regions; an inner *reaction region* (red broken line) and an outer *buffer region* (blue solid line). Outside the buffer region the atoms are fixed. B) An isolated BRC repeat (cyan) in the water sphere (red and white).

2.3 Production runs

We carried out 12ns of molecular dynamics simulations in total for each of the 9 BRC-RAD51 complexes and for each of the 9 isolated BRC proteins, using the following procedure: first one 2ns simulation was performed (except in the cases of isolated BRC4, BRC4TA and their complexes with RAD51 where this first simulation was run for 10ns) and then five further simulations of 2ns duration were run starting from snapshots taken at 400, 800, 1200, 1600 and 1900 ps intervals of the first simulation. This procedure is summarised in the diagram of Figure 4 from which it can be seen that none of the simulations were run as continuations of the first simulation but they were instead first equilibrated according to the procedure we described in the previous section.



1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Figure 4. Schematic representation of the simulations we have carried out for each protein. The horizontal arrows represent production simulations of 2ns while the red dots at the beginning of each arrow represent the set-up and equilibration procedure described in section 2.2. The initial structures for simulation 1 were generated using the methods described in section 2.1 while the initial structures for simulations 2-6 were provided from snapshots extracted from simulation 1 at the time intervals indicated.

We have followed this procedure in order to enhance our sampling of conformational space and to examine how the different BRC repeats may evolve away from their starting structures. All our simulations were performed in pure water and therefore our observations will not be directly relevant for the interpretation of any salt concentration-dependent effects.

Also, the simulation techniques and the subsequent analysis employed in this work are only suitable for examining the bonding interactions and conformational flexibility of the isolated BRC repeats and of their complexes with RAD51, and are not suitable for drawing (even qualitative) conclusions about the relative binding affinities of the BRC repeats. Other types of approaches can be used in order to address this issue, and this is the focus of ongoing work on which we intend to report in the future.

3 Results and Discussion

In this section we present an analysis of the conformations, bonding and flexibility of the proteins during the simulations. We focus on the similarities, and importantly differences, in the hairpin loops of the BRC repeats as they all use a conserved sequence motif that mimics the oligomerisation of RAD51. The section is divided into two subsections. In the first subsection we discuss and analyse the simulations of the isolated repeats and in the second subsection we discuss the simulations of the complexes where we also observe the effect that the complex formation has on the rigidity of the repeats. The insights obtained are important, more so as to our knowledge no experimental results (e.g. from NMR [15N-HSQC] or CD data) into the motions and flexibility of the repeats are available.

3.1 Isolated BRC repeats

The hairpin loop of the BRC repeats is stabilised by two intramolecular hydrogen bonds between its backbone atoms, as shown in Figure 5.

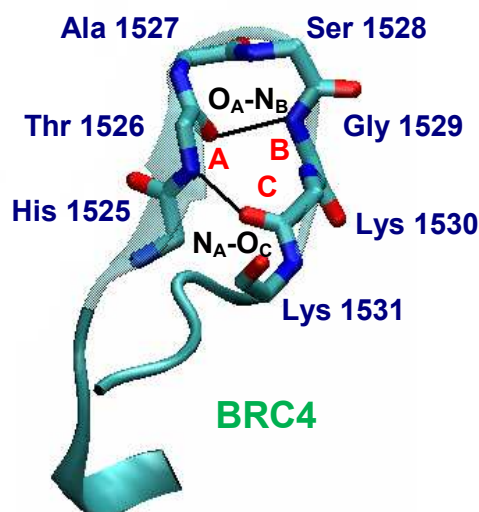


Figure 5. Structure of the backbone of the BRC4 motif around its β -hairpin loop, showing the two hydrogen bonds (black lines) between backbone atoms that stabilise the loop (N_A - O_C , between Thr1526-Lys1530 and O_A - N_B between Thr1526-Gly1529).

The hydrogen bonds form between the backbones of residues marked as A, B and C in Figure 5. As only backbone atoms are involved, we expect these interactions to be conserved in all of the repeats, however they will differ in length and strength as a result of the different side chains in the 16-amino acid sequence that is not conserved throughout the BRC repeats. For example, in the case of BRC4 the hydrogen bonding occurs between the residues Thr1526 (A) Gly1529 (B) and Lys1530 (C), while in BRC2 the hydrogen bonding occurs between Ser1221 (A), Gly1224 (B) and Thr1225 (C). Table 1 shows the average bond distances from all the simulations (the 6 simulations described in Figure 4) for each BRC repeat, according to the labelling scheme of Figure 5. Note that in the case of BRC4TA there are three sets of values, as three distinct conformations (“states” 1-3 in the Table) were observed in the simulations.

Table 1. Residues in positions A, B and C for each of the BRC repeats, and the average bond distances (in Å) for the intra-molecular backbone interactions.

Repeat	A	O_A - N_B	N_A - O_C	B	C
BRC1	THR	4.58 ± 0.21	5.36 ± 0.23	ASN	LYS
BRC2	SER	4.46 ± 0.10	4.67 ± 0.12	GLY	THR
BRC3	THR	3.84 ± 0.21	2.96 ± 0.30	GLY	LYS
BRC4	THR	3.62 ± 0.19	2.87 ± 0.12	GLY	LYS
BRC5	THR	3.23 ± 0.12	3.00 ± 0.16	SER	ARG
BRC6	ILE	3.12 ± 0.08	3.02 ± 0.21	GLY	LYS
BRC7	THR	3.90 ± 0.14	3.09 ± 0.09	GLY	LYS
BRC8	THR	4.04 ± 0.05	3.06 ± 0.32	GLY	LYS
BRC4TA (state 1) ALA		3.26 ± 0.22	2.89 ± 0.23	GLY	LYS
BRC4TA (state 2) ALA		6.75 ± 0.09	3.35 ± 0.28	GLY	LYS
BRC4TA (state 3) ALA		6.76 ± 0.10	5.19 ± 0.21	GLY	LYS

Throughout the simulations for BRC3-BRC8, the O_A-N_B and N_A-O_C distances correspond to stable hydrogen bonds and are maintained with only small fluctuations. In contrast, in BRC1, BRC2 and BRC4TA (single point mutation of Thr1526 by Ala, see Figure 1) these interactions appear to be weaker as the O_A-N_B and N_A-O_C distances are longer by about 1 Å or more. It is particularly interesting to compare the behaviour of BRC4 and BRC4TA as a function of time. We observe that the lengths of the O_A-N_B and N_A-O_C hydrogen bonds for BRC4 (Figure 6) remain almost constant, while in the case of BRC4TA there is increased flexibility and variation in both hydrogen bond distances. We can divide the BRC4TA simulation into three stages: The first stage is during the first 3 ns where both bond distances are relatively well-maintained and are similar to those of BRC4. The second stage is from 3 ns to 6 ns where the O_A-N_B distance increases to about 6 Å and the N_A-O_C distance gradually increases to about 5 Å. Finally during the third stage from 6 ns onwards the O_A-N_B and N_A-O_C distances do not vary on average but do fluctuate considerably around their average values.

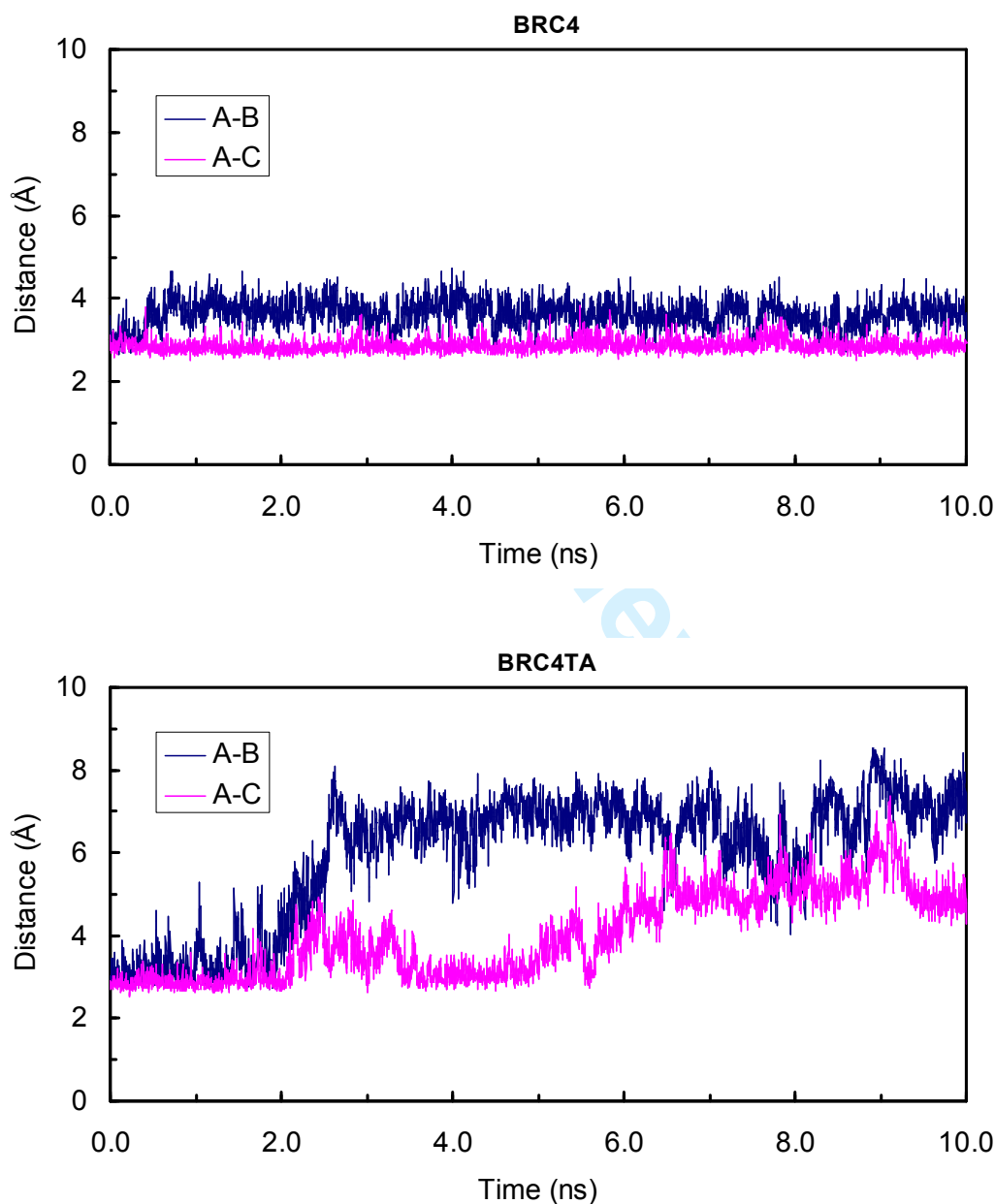


Figure 6. The two backbone hydrogen bond distances (in Å) as functions of time from the simulations of isolated BRC4 (bonds between $O_{Thr}-N_{Gly}$ and $N_{Thr}-O_{Lys}$) and BRC4TA (bonds between $O_{Ala}-N_{Gly}$ and $N_{Ala}-O_{Lys}$).

Further insight into the differences in flexibility and the backbone conformations of the BRC repeats can be gained by examining the structures and by following the torsion angles of the residues that constitute the β -hairpin loop. An analysis of the rotational freedom about the ϕ and ψ angles using Ramachandran plots shows that despite having the same three amino acids (Ala1527, Ser1528 and Gly1529) in their β -hairpin loops BRC4TA and BRC4 behave rather differently as BRC4TA displays a more diffuse distribution of angles than BRC4 (Figure 7).

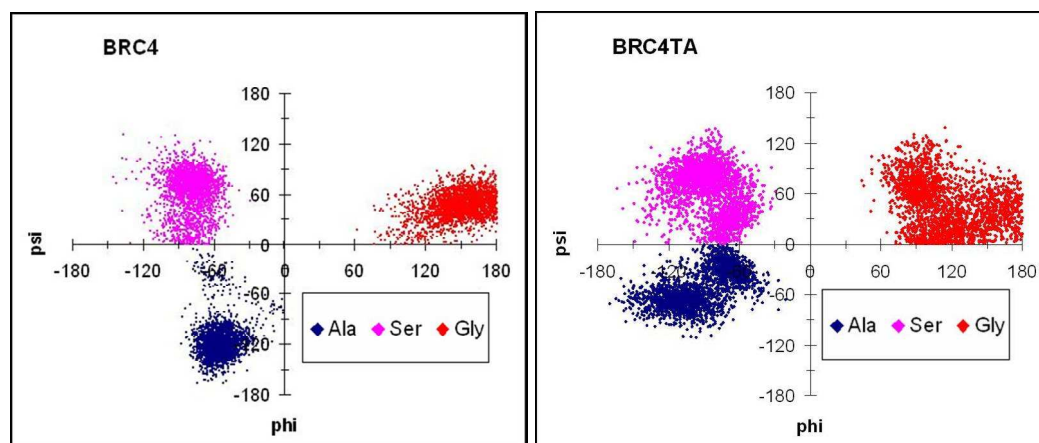


Figure 7. Ramachandran plots for the BRC4 and BRC4TA isolated repeats showing the ϕ and ψ torsion angles of the three residues that form their β -hairpin loops.

For some of the conformations of BRC4TA the angle distributions for Gly1529 are identical to those of BRC4 as can be seen from the Ramachandran plots in Figures 7 and also in Figure 8 where during the first 3 ns of the simulations the two BRC repeats show a similar distribution. However from then onwards the BRC4TA distribution of angles becomes more diffuse and the plots of Figure 8 indicate that it is exchanging between two conformations. The first conformation of BRC4TA is similar to that of BRC4 with $\phi = 60 \pm 30$, $\psi = 150 \pm 30$ while the second one has undergone a 60° rotation about the ψ angle $\phi = 60 \pm 30$, $\psi = 90 \pm 30$ which is still within the allowed region for the left handed helix. This increased flexibility is a consequence of the replacement of Thr by Ala in BRC4TA as this leads to loss of the stabilisation of the loop by the Thr side chain.

The left-handed helix region of the Ramachandran plots is normally only allowed for Gly, and as expected, it is mainly Gly that occupies this position in the BRC repeats. However, non Gly residues are found in this position for BRC1 and BRC5. BRC1 exhibits increased flexibility and has the longest average N_A-O_C distance out of all the BRC repeats. For BRC5 the conformation required to form the loop and the strain placed on the non-Gly residue Ser, appears to be compensated by the rest of the loop as the backbone-backbone hydrogen bonds do not become elongated.

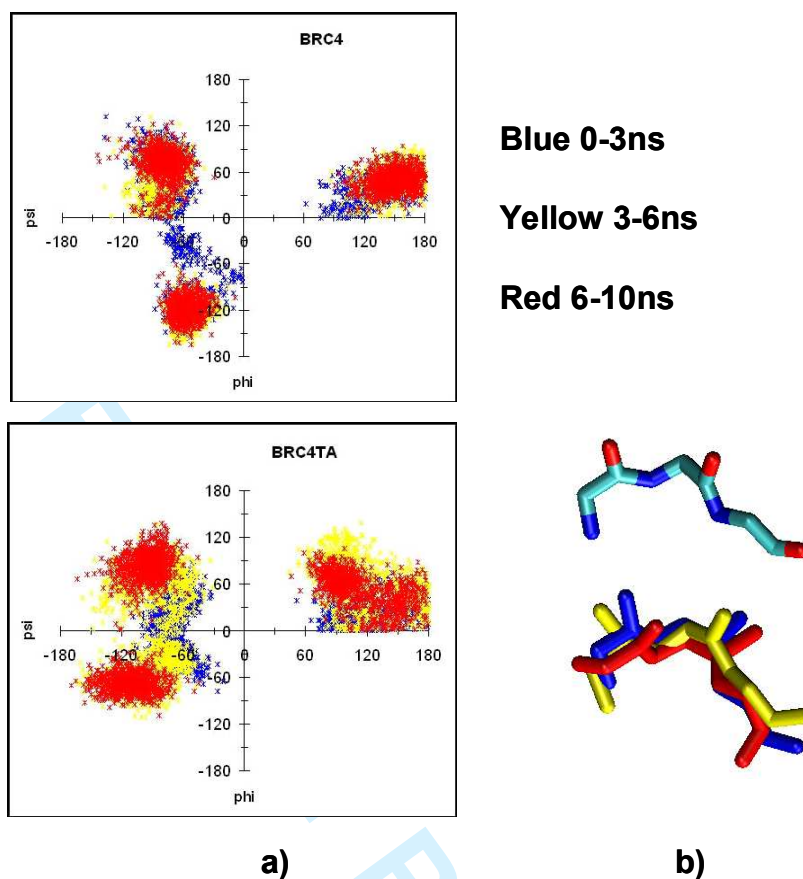


Figure 8. a) Ramachandran plots for the BRC4 and BRC4TA isolated repeats showing the ϕ and ψ torsion angles of the protein backbone, for three time intervals of the simulation. b) The backbone configurations of the hairpin loop residues in BRC4 and BRC4TA. Top: snapshot of BRC4 at 5 ns. Bottom: Snapshots of BRC4TA in the starting configuration (blue), at 4 ns (yellow) and at 9 ns (red), showing the evolution of conformations.

The Ser Ramachandran plots of both BRC4 and BRC4TA lie predominantly as expected in the β -sheet region $\phi = 90 \pm 40$, $\psi = -90 \pm 20$ and show a similar distribution from 4 ns onwards. They differ in the first 4 ns of the simulation, where the distribution of BRC4TA is more diffuse switching between two conformations in the $\phi = 60 \pm 20$, $\psi = -60 \pm 20$ region. The Ala Ramachandran plots show two distinctly different conformations for the two BRC repeats after the first 3 ns of simulation. Both BRC4 and BRC4TA for the Ala plot are distributed in the alpha helix region and start from the same location $\phi = -30 \pm 20$, $\psi = -60 \pm 20$. However after 3 ns BRC4TA has adopted a $\phi = -60 \pm 40$, $\psi = -100 \pm 20$ conformation, while BRC4 has adopted a conformation in the $\phi = -120 \pm 20$, $\psi = -60 \pm 20$ region. It is therefore clear that the conformation of the β -hairpin loop changes during the simulation of BRC4TA. In the BRC4TA Ala1527 and Gly1529 adopt new conformations (Fig 6), while Ser1528 remains close to that found in BRC4, despite showing initial differences. As a result BRC4TA switches to a new more “open” conformation from the one it started while BRC4 retains its initial conformation.

Overall, BRC1, BRC2 and BRC4TA display the greatest degree of flexibility in their conformation during the simulations. The behaviour of BRC1 is a result of the loss of Gly in the third position of the β -hairpin loop, whereas while BRC2 and BRC4TA both retain the Gly in this position they incur the loss of Thr1526, by its replacement by Ser and Ala respectively, and this results in the loss of side chain intramolecular bonds as we will see next. The BRC repeats 3-8 appear to be the most rigid, maintaining their original conformation via a network of intra-molecular bonds.

In addition to the backbone-backbone interactions that we examined above, there are also a number of intra-BRC hydrogen bonds that form between side chains of residues of the hairpin loop. These side chain interactions, which are shown in Figure 9 for the case of BRC4, and are tabulated in Tale 2 for all the repeats, play an important role in keeping the two sides of the loop poised for binding to RAD51.

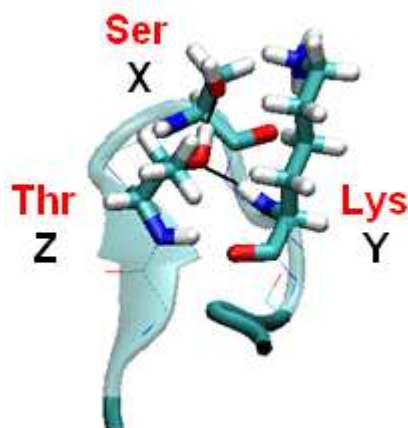


Figure 9. Side chain hydrogen bonds for an isolated BRC repeat, between Z-Y and Z-X (in BRC4 Z = Thr, X = Ser and Y=Lys).

Some of the BRC repeats contain residues that are not capable of making intramolecular hydrogen bonds with their side chains, and in these cases the side chains are oriented to point towards the bulk solvent. The point mutation in BRC4TA (replacement of Thr with Ala) is such an example and has a profound effect on the rigidity of the chain as in this case both of the hydrogen bonds shown in Figure 9 are disrupted.

Table 2. The residues in positions X, Y and Z of the BRC repeats (see Figure 9) and the average bond distance (Å) for the Z-Y and Z-X bonds.

Repeat	X	Y	Z	Z-Y(BB)	Z-X
BRC1	SER	LYS	THR	6.95 ± 0.09	8.66 ± 0.09
BRC2	HIS	THR	SER	6.59 ± 0.18	10.62 ± 0.06
BRC3	SER	LYS	THR	3.06 ± 0.10	3.27 ± 0.19
BRC4	SER	LYS	THR	6.66 ± 0.12	8.45 ± 0.09
BRC5	CYS	ARG	THR	3.03 ± 0.14	5.63 ± 0.13
BRC6	SER	LYS	ILE		
BRC7	SER	LYS	THR	3.02 ± 0.18	3.69 ± 0.10
BRC8	SER	LYS	THR	3.07 ± 0.07	3.47 ± 0.12
BRC4TA	SER	LYS	ALA		

We can see from Table 2 that BRC repeats 3, 5, 7 and 8 form intramolecular hydrogen bonds. The side chains at position Z in the cases of BRC6 and BRC4TA are hydrophobic and thus do not contain atoms capable of forming an intramolecular hydrogen bond network.

3.2 BRC-RAD51 complexes

The BRC repeat motif that we studied here contains two hydrophobic residues that make contact with RAD51: (1) Phe which is present in all eight of the BRC repeats (Phe1524 in BRC4) and in the single point mutation BRC4TA and (2) Ala which is present in seven out of the eight BRC

repeats (Ala1527 in BRC4). There are also some polar contacts along the interface with RAD51 and the combination of these interactions is responsible for the binding between RAD51 and the BRC repeats.

We first examine the two intra-BRC backbone hydrogen bonds that we saw in the previous section, as they are also present in the case of the BRC-RAD51 complex. The hydrogen bonds in question form between the residues in positions A, B and C as shown in Figure 10.

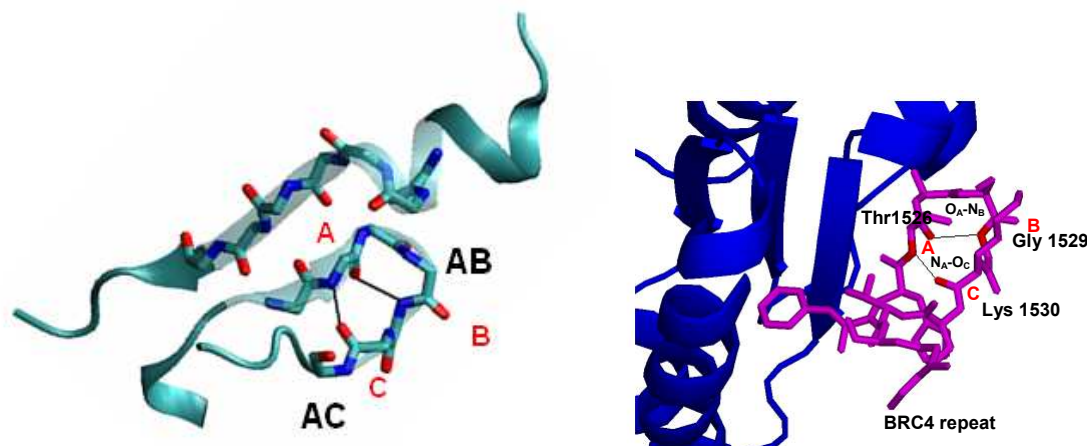


Figure 10. The backbone structure of BRC4 bound to RAD51 (truncated in this picture to only include the section of RAD51 interacting directly with BRC). The intra-BRC hydrogen bonding backbone interactions are shown by black lines (A-C between Thr1526-Lys1530 and A-B between Thr1526-Gly1529 for BRC4).

In the case of the complex, these two intra-BRC backbone-backbone hydrogen bonds are maintained throughout the simulations for all the BRC repeats, as can be seen in Table 3. This is in contrast with the simulations of the unbound BRC repeats where BRC1, 2 and 4TA showed partial disruption of the N_A-O_C bond as the simulation progressed with BRC4TA adopting a new conformation for its β -hairpin loop.

Table 3. The residues in positions A, B and C for each of the BRC repeats, and the average bond distance (\AA) for the intra-BRC backbone interactions when bound to RAD51.

Repeat	A	O_A-N_B	N_A-O_C	B	C
BRC1	THR	3.16 ± 0.12	3.17 ± 0.10	ASN	LYS
BRC2	SER	3.49 ± 0.06	2.92 ± 0.14	GLY	THR
BRC3	THR	3.81 ± 0.13	3.09 ± 0.06	GLY	LYS
BRC4	THR	3.71 ± 0.12	3.17 ± 0.09	GLY	LYS
BRC5	THR	3.06 ± 0.15	2.89 ± 0.12	SER	ARG
BRC6	ILE	3.48 ± 0.09	2.98 ± 0.15	GLY	LYS
BRC7	THR	3.78 ± 0.11	3.30 ± 0.08	GLY	LYS
BRC8	THR	3.93 ± 0.21	3.10 ± 0.09	GLY	LYS
BRC4TA	ALA	4.13 ± 0.18	2.86 ± 0.13	GLY	LYS

The second bond bridging the β -hairpin loop between O_A-N_B also remains rather rigid during the simulation for all the BRC repeats and appears much tighter than in the unbound case: Only BRC4TA demonstrates some degree of fluctuation (Figure 11) which is caused by the fact that the Oxygen of the Ala is less restrained. These observations support the view that when bound to RAD51, the BRC repeat sequences can vary widely without significant changes to their backbone conformation.

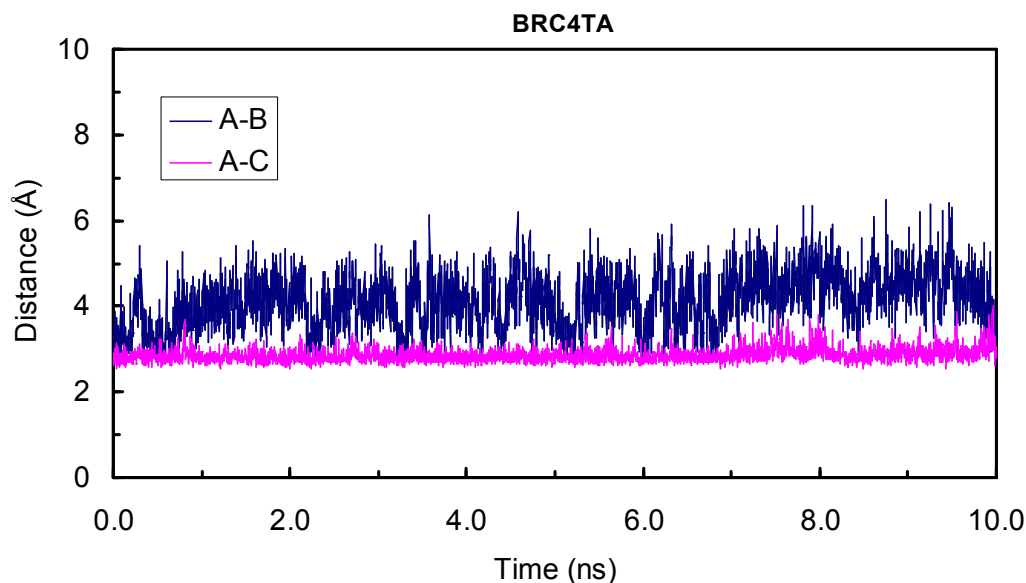


Figure 11. Bond distances (in Å) as a function of time between the $O_{\text{Ala}}-N_{\text{Gly}}$ (A-B) and $N_{\text{Ala}}-O_{\text{Lys}}$ (A-C) atoms in the backbone of the BRC4TA repeat when bound to RAD51.

The BRC backbone also forms two hydrogen bonds directly with the backbone of RAD51 through the FX_1X_2A sequence, which is conserved in 7 out of the 8 BRC repeats, and the FX_1X_2S sequence in the case of BRC5. Here X_1 and X_2 signify residues that are not necessarily conserved between the different repeats. These two hydrogen bonds contribute to the stabilisation of the complex, in conjunction with the two hydrophobic interactions by the sidegroups of F and A. The FX_1X_2A sequence is also encountered in the N-terminal domain of the RecA-like recombinase of many organisms and is believed to be responsible for the binding of the RAD51 proteins with each other in the multimerisation of RAD51 that occurs during filament formation (20). The BRC residue at position X_1 (His1525 in the case of BRC4) forms hydrogen bonds with the backbone of the Tyr191 of RAD51 (through its peptidic amine and carbonyl groups), which strengthen the contact between BRC and RAD51 as shown in Figure 12.

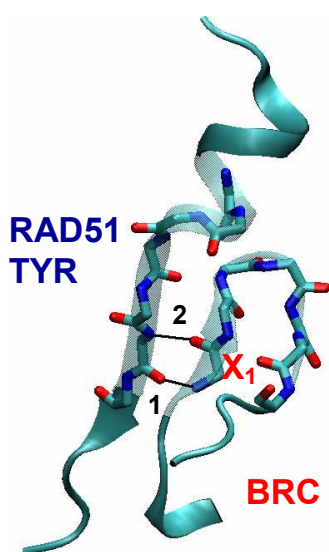


Figure 12. Diagram showing the two intermolecular backbone-backbone hydrogen bonds between a BRC repeat and RAD51; (1) peptidic amine of the BRC4 (His 1525) interacting with the carbonyl oxygen of Tyr on RAD51; (2) the carbonyl oxygen of the BRC interacting with the amine of the same Tyr of RAD51.

Table 4. Residue in position X_1 of the BRC repeats and their average hydrogen bond distances (Å) from the Tyr in the backbone of RAD51

Repeat	X_1	1	2
BRC1	ARG	2.90 ± 0.08	2.83 ± 0.03
BRC2	TYR	2.88 ± 0.10	2.82 ± 0.07
BRC3	GLN	2.95 ± 0.03	2.82 ± 0.09
BRC4	HIS	2.99 ± 0.06	2.95 ± 0.12
BRC5	TYR	2.87 ± 0.07	2.79 ± 0.14
BRC6	ARG	2.90 ± 0.04	2.79 ± 0.06
BRC7	SER	2.84 ± 0.07	2.80 ± 0.04
BRC8	SER	2.89 ± 0.11	2.80 ± 0.07
BRC4TA	HIS	3.06 ± 0.08	2.78 ± 0.09

Irrespective of the nature of the residue in position X_1 in the sequence FX_1X_2A , or FX_1X_2S in the case of BRC5, the observed distances (see Table 4 and Figure 12) suggest that the two hydrogen bonds are well maintained for all of the BRC repeats throughout the simulations.

Our observations so far support the view that the presence of RAD51 in a complex with BRC does also affect the intra-molecular interactions in the BRC repeats. In particular, it appears that the participation of a BRC repeat in a complex with RAD51 results in the repeat being much more rigid with a decreased level of fluctuation from its preferred conformation. This trend also carries over to the side chains of the repeats. In the simulations of the complexes the repeats whose side chains are capable of forming intra-molecular hydrogen bonds, as shown in Figure 13, appear more rigid than in simulations where the same repeats are isolated. Average distances for the intra-BRC side chain bonds are shown in Table 5. In contrast with Table 2 we can observe that in all the repeats with side chains capable of forming hydrogen bonds the distances remain within the lengths expected for hydrogen bonds, with BRC2 being the only exception.

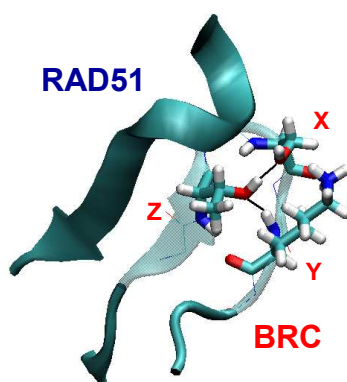
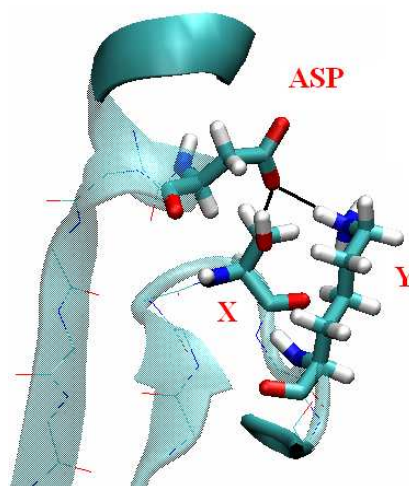


Figure 13. The network of intra-BRC side chain hydrogen bonding interactions in the complex with RAD51 (in BRC4 Z = Thr, Y = Lys and X = Ser).

Table 5. The residues in positions X, Y and Z and the average bond distance (Å) for the Z-Y and Z-X bonds in the BRC-RAD51 complex.

Repeat	X	Y	Z	Z-Y(BB)	Z-X
BRC1	SER	LYS	THR	2.93 ± 0.08	2.77 ± 0.09
BRC2	HIS	THR	SER	3.40 ± 0.19	6.77 ± 0.21
BRC3	SER	LYS	THR	2.96 ± 0.22	2.77 ± 0.11
BRC4	SER	LYS	THR	2.91 ± 0.13	2.79 ± 0.05
BRC5	CYS	ARG	THR	3.15 ± 0.18	4.29 ± 0.12
BRC6	SER	LYS	ILE		
BRC7	SER	LYS	THR	2.92 ± 0.09	2.76 ± 0.09
BRC8	SER	LYS	THR	2.92 ± 0.08	2.76 ± 0.10
BRC4TA	SER	LYS	ALA		

The FX_1X_2A domain of the BRC repeats plays a prominent role in maintaining the contact to RAD51 by forming an anti-parallel two-strand extension of its beta sheet. This is stabilised by two hydrophobic interactions, firstly with the Phe whose side group is buried within a distinct pocket in RAD51, and secondly through the Ala which is buried within a small cavity formed by the side chains of residues Phe166, Pro168, Leu171, Leu186 and Val189 of RAD51.

**Figure 14.** Side chain hydrogen bonding between two residues of the BRC repeats (in BRC4 “X” = Ser1527 and “Y” = Lys1530) and the Asp187 of RAD51.

The BRC-RAD51 complex is further stabilised by two side chain inter-protein polar interactions involving BRC residues not directly in backbone contact with RAD51. As shown in Figure 14, one of these contacts is formed by the opposite charges between the side chains of Asp187 of RAD51 and a Lys, shown as position “Y” (which is Lys1530 for the case of BRC4) which is present in BRC repeats 1, 3, 4, 6, 7, 8, and BRC4TA. The same repeats also form a second contact with RAD51 in the form of a hydrogen bond, also between Asp187 and a Ser, which is shown as position “X” in Figure 14 and is Ser 1527 in the case of BRC4. In the case of BRC2 there is a Thr in position “Y” which is unable to reach an orientation that allows for interaction with the Asp of RAD51. In the case of BRC5 the residue in position “Y” is Arg whose side chain has the correct length and chemical functionality to interact with the Asp of RAD51. However this is not observed because the interaction is only possible if Asp adopts a suitable orientation through hydrogen bonding with the side chain of the residue at position “X” (see Figure 14) in the BRC repeats but BRC5 has a Cys in this position which is not capable of bonding with the RAD51 Asp. Distances for these interactions are given in Table 6.

Table 6. Residues in positions X and Y of the BRC repeats and average bond distances (Å) for their polar interactions with the Asp of RAD51

Repeat	X	X-ASP	Y	Y-ASP
BRC1	SER	2.66 ± 0.06	LYS	2.77 ± 0.07
BRC2	HIS	6.10 ± 0.13	THR	
BRC3	SER	2.97 ± 0.23	LYS	2.85 ± 0.20
BRC4	SER	2.63 ± 0.18	LYS	2.74 ± 0.12
BRC5	CYS	5.94 ± 0.10	ARG	6.02 ± 0.14
BRC6	SER	3.78 ± 0.21	LYS	5.70 ± 0.19
BRC7	SER	2.63 ± 0.12	LYS	2.74 ± 0.13
BRC8	SER	3.22 ± 0.07	LYS	2.90 ± 0.10
BRC4TA	SER	2.67 ± 0.19	LYS	2.99 ± 0.12

It is interesting to note that even though all our simulations were performed in explicit water, we did not observe any insertion of water molecules at the interface between RAD51 and each of the BRC repeats. This indicates that the RAD51 and the BRC repeats are flexible enough to rapidly adopt and retain overall conformations that closely follow the shape of each other.

4 Conclusions

We have studied the complex of the human recombinase enzyme RAD51 with the hairpin loop region of the BRC repeat motifs of the cancer susceptibility protein BRCA2, using molecular dynamics simulations in explicit water, at room temperature. All eight BRC repeats (BRC1-BRC8) that are present in BRCA2 were investigated as well as the repeat BRC4TA which is produced by a mutation of the BRC4 repeat Thr1526 to Ala. Our simulations have enabled us to observe the dynamical behaviour of the complex and of the isolated BRC repeat proteins and to carry out a systematic analysis of its conformations and of the interactions that are responsible for their stabilisation. The differences in conformations between the BRC repeats are caused by differences in both intra and inter-protein hydrogen bonds and affect the shape and rigidity of the hairpin loop. The high degree of conservation of the ASG sequence which is needed for the formation of the hairpin loop and the ability of the BRC repeat side chains to form intra-molecular hydrogen bonds have also as a consequence the orientation of key side chains in a way that allows them to interact favourably with RAD51. In simulations of the isolated repeats, BRC1, BRC2 and BRC4TA show the most flexibility and adopt new conformations; in contrast BRC repeats 3-8 are more rigid and retain a conformation similar to that of BRC4 in the crystal structure of the BRCA2-RAD51 complex. In the complex the repeats appear much more rigid due to several hydrophobic, charge-charge and hydrogen bonding interactions with RAD51 which re-inforce their conformational stability.

5 Acknowledgments

N. B. would like to thank the Medical Research Council for postdoctoral funding through a Discipline-Hopping grant (Grant number: G0502205). C.-K. S would like to thank the Royal Society for a University Research Fellowship. MCP acknowledges financial support from EPSRC (Grant GR/S61263/01).

6 References

- (1) AR Venkitaraman: Cancer susceptibility and the functions of BRCA1 and BRCA2. *Cell* 108 (2002) 171-82.
- (2) A Tutt, D Bertwistle, J Valentine, A Gabriel, S Swift, G Ross, C Griffin, J Thacker, A Ashworth: Mutation in Brca2 stimulates error-prone homology-directed repair of DNA double-strand breaks occurring between repeated sequences. *Embo Journal* 20 (2001) 4704-16.
- (3) P Sung, H Klein: Mechanism of homologous recombination: mediators and helicases take on regulatory functions. *Nature Reviews Molecular Cell Biology* 7 (2006) 739-50.
- (4) DS Shin, L Pellegrini, DS Daniels, B Yelent, L Craig, D Bates, DS Yu, MK Shivji, C Hitomi, AS Arvai, N Volkmann, H Tsuruta, TL Blundell, AR Venkitaraman, JA Tainer: Full-length archaeal Rad51 structure and mutants: mechanisms for RAD51 assembly and control by BRCA2. *Embo Journal* 22 (2003) 4566-76.
- (5) RM Story, IT Weber, TA Steitz: THE STRUCTURE OF THE ESCHERICHIA-COLI RECA PROTEIN MONOMER AND POLYMER. *Nature* 355 (1992) 318-25.
- (6) H Aihara, Y Ito, H Kurumizaka, S Yokoyama, T Shibata: 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* 290 (1999) 495-504.
- (7) AA Davies, JY Masson, MJ McLlwraith, AZ Stasiak, A Stasiak, AR Venkitaraman, SC West: Role of BRCA2 in control of the RAD51 recombination and DNA repair protein. *Molecular Cell* 7 (2001) 273-82.
- (8) MKK Shivji, AR Venkitaraman: DNA recombination, chromosomal stability and carcinogenesis: insights into the role of BRCA2. *DNA Repair* 3 (2004) 835-43.
- (9) V Yu, M Koehler, C Steinlein, M Schmid, LA Hanakahi, AJ van Gool, SC West, AR Venkitaraman: Gross chromosomal rearrangements and genetic exchange between nonhomologous chromosomes following BRCA2 inactivation. *Genes & Development* 14 (2000) 1400-06.
- (10) AKC Wong, R Pero, PA Ormonde, SV Tavtigian, PL Bartel: RAD51 interacts with the evolutionarily conserved BRC motifs in the human breast cancer susceptibility gene brca2. *Journal of Biological Chemistry* 272 (1997) 31941-44.
- (11) P Baumann, FE Benson, SC West: Human Rad51 protein promotes ATP-dependent homologous pairing and strand transfer reactions in vitro. *Cell* 87 (1996) 757-66.
- (12) T Lo, L Pellegrini, AR Venkitaraman, TL Blundell: Sequence fingerprints in BRCA2 and RAD51: implications for DNA repair and cancer. *DNA Repair* 2 (2003) 1015-28.

- 1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
- (13) PL Chen, CF Chen, YM Chen, J Xiao, ZD Sharp, WH Lee: The BRC repeats in BRCA2 are critical for RAD51 binding and resistance to methyl methanesulfonate treatment. *Proceedings of the National Academy of Sciences of the United States of America* 95 (1998) 5287-92.
 - (14) L Pellegrini, DS Yu, T Lo, S Anand, M Lee, TL Blundell, AR Venkitaraman: Insights into DNA recombination from the structure of a RAD51-BRCA2 complex. *Nature* 420 (2002) 287-93.
 - (15) L Pellegrini, A Venkitaraman: Emerging functions of BRCA2 in DNA recombination. *Trends in Biochemical Sciences* 29 (2004) 310-16.
 - (16) CF Chen, PL Chen, Q Zhong, ZD Sharp, WH Lee: Expression of BRC repeats in breast cancer cells disrupts the BRCA2-Rad51 complex and leads to radiation hypersensitivity and loss of G(2)/M checkpoint control. *Journal of Biological Chemistry* 274 (1999) 32931-35.
 - (17) AD MacKerell, D Bashford, M Bellott, RL Dunbrack, JD Evanseck, MJ Field, S Fischer, J Gao, H Guo, S Ha, D Joseph-McCarthy, L Kuchnir, K Kuczera, FTK Lau, C Mattos, S Michnick, T Ngo, DT Nguyen, B Prodhom, WE Reiher, B Roux, M Schlenkrich, JC Smith, R Stote, J Straub, M Watanabe, J Wiorkiewicz-Kuczera, D Yin, M Karplus: All-atom empirical potential for molecular modeling and dynamics studies of proteins. *Journal of Physical Chemistry B* 102 (1998) 3586-616.
 - (18) BR Brooks, RE Bruccoleri, BD Olafson, DJ States, S Swaminathan, M Karplus: CHARMM - A PROGRAM FOR MACROMOLECULAR ENERGY, MINIMIZATION, AND DYNAMICS CALCULATIONS. *Journal of Computational Chemistry* 4 (1983) 187-217.
 - (19) WL Jorgensen, J Chandrasekhar, JD Madura, RW Impey, ML Klein: COMPARISON OF SIMPLE POTENTIAL FUNCTIONS FOR SIMULATING LIQUID WATER. *Journal of Chemical Physics* 79 (1983) 926-35.
 - (20) AB Conway, TW Lynch, Y Zhang, GS Fortin, CW Fung, LS Symington, PA Rice: Crystal structure of a Rad51 filament. *Nature Structural & Molecular Biology* 11 (2004).