

**‘Structural Dynamics of a Nanomachine: Investigating
Structure-to-Function Properties of Large, Flexible,
DNA-Associating Proteins’**

‘De Structurele Dynamiek van een Nanomachine: Onderzoek naar
Structure-to-Function Eigenschappen van grote, flexibele, DNA-
verbindende Proteïnen’

Thesis
to obtain the degree of Doctor from the
Erasmus University Rotterdam
by command of the
rector magnificus

Prof.dr. R.C.M.E. Engels

and in accordance with the decision of the Doctorate Board.
The public defence shall be held on

29 January 2019 at 15.30 hrs

**By Małgorzata Grosbart
Born in Warsaw, Poland**



Erasmus University Rotterdam

Doctoral Committee:

Promotors: Prof.dr. C. L. Wyman
Prof.dr. R. Kanaar

Other members: Dr. J. Lebbink
Prof.dr. W. Vermeulen
Dr. M.E. Aubin-Tam

Table of Contents

Chapter 1 General Introduction: Cracking the tools.....	5
--	---

Chapter 2 The bacterial condensin MukB compacts DNA by sequestering supercoils and stabilizing topologically isolated loops.....	25
---	----

Chapter 3 Investigating the structure-to-function relationships of MRE11-RAD50 complex.....	79
--	----

Chapter 4 Imaging of DNA and protein by SFM and combined SFM-TIRF microscopy.....	113
--	-----

Chapter 5 Architectural plasticity of human BRCA2-RAD51 complexes in DNA break repair.....	153
---	-----

Chapter 6 Nanoscale mapping: uncovering the structure of BRCA2 protein by imaging individual domains.....	201
--	-----

Appendix

Summary.....	227
--------------	-----

Curriculum Vitae.....	236
-----------------------	-----

Portfolio.....	237
----------------	-----

List of publications.....	239
---------------------------	-----

Acknowledgments.....	240
----------------------	-----

Chapter one

General Introduction: Cracking the tools

*M. Grosbart*¹

¹Department of Molecular Genetics, Erasmus University Medical Center, P. O. Box 2040, 3000CA Rotterdam, The Netherlands

From basic instructions to a nanomachine

Deoxyribonucleic acid, or DNA, serves as an information bank of every cell by the means of four chemically distinctive monomers, organized linearly in a tri-partitioned code. Combinations of these triplets (called codons) form information to be read and executed within the cell called genes (Crick et al., 1961). In this way, with only four components, human DNA cell contains over 20,000 distinct protein-encoding genes (Ponomarenko et al., 2016). These, in turn, serve as templates for construction of amino-acid chains (Crick, 1970), which, after complex folding and modification events, form active proteins (Ellgaard et al., 2016).

Essentially, we can imagine the role of DNA as an instruction book. And if the DNA is an instruction book, proteins can be viewed as tools that perform, maintain, and coordinate all cellular functions.

One protein molecule (or one kind of protein molecule) is not enough to perform even the easiest cellular processes (often called pathways), therefore many different proteins cooperate to complete each task in this process. However, one type of protein is often not enough to complete *one task*. This problem is solved by creating protein complexes—associations of various proteins with complimenting functions—to perform a task (or tasks) in a complex cellular pathway. A combination of various tools, each designed to fulfill a specific function in a process, can be considered a biological machine. Therefore, it is justified that specialized protein complexes, performing intricate, multi-step

operations on a nanometer scale within cellular pathways, should be called *nanomachines*.

Vulnerabilities of the system: DNA damage and repair mechanisms

Alterations or loss of continuity in the DNA tripartite code will necessarily alter how the code is read and interpreted. Changes, whether big (large-scale deletions, multiplications, strand brakes) or small (point mutations, missing bases, deaminations, Thymine dimers) can have potentially disastrous consequences for the cell (Hoeijmakers, 2009). DNA is chemically stable (Lesnik and Freier, 1995) and it is tightly packed and covered with proteins within the nuclear matrix (Green and Almouzni, 2002), which can make it seem as if the genetic information can be stored and used with safety. This, however, is not the case.

DNA's integrity is constantly threatened with both external and internal factors. Toxins, free radicals, replication errors, radiation are all potential causes of DNA damage. Loss of continuity in genetic information carried by the DNA can have catastrophic consequences for the cell and, in some cases, the whole organism. Therefore, a whole array of DNA damage repair mechanisms is available, carried out by specialized sets of proteins and each specific to a different kind of lesion (Giglia-Mari et al., 2011). A very dramatic type of lesions are double strand breaks, where the DNA is snapped in two somewhere along the helix, creating loose ends. Ironically, this kind of damage is often the result of natural cellular processes (replication, meiosis), and occurs almost daily (Kanaar et al., 1998; Wyman and Kanaar,

2006). The toxicity and high frequency of occurrence of double strand breaks lead to development of two repair mechanisms devoted to these particular lesions. Non-homologous End Joining (NHEJ) is a straightforward process in which two ends of broken DNA are ligated together, at the risk of loss of some nucleotide content (genetic information) at the side of the break. Homologous Recombination (HR) is a more complex, multi-step operation, involving recreation of genetic information at the site of the damage with great fidelity (Wright et al., 2018; Wyman et al., 2004).

The goal of HR is to use a homologous DNA strand as a template to faithfully recreate information on the damaged strand. In simplification, it can be seen as a succession of steps that start at two DNA ends (the lesion) and one intact DNA fragment (template) and result in two intact strands (template and repaired strand). The first step is Damage Recognition, where a set of proteins clusters on the loose DNA ends, setting cell cycle signaling pathways in motion. It is closely followed by End Resection, where stretches of one of the DNA strands at the damage site are removed by specialized nucleases to create 3' ss overhangs on each side of the break. A part of these early processes in HR is DNA tethering, where the two ends are being held in close proximity by the MRE11/RAD50/(NBS1) (MR(N))complex, which is also playing a crucial role in both Recognition and Resection steps. Next, in a step called Filament Formation, the newly formed ssDNA at the 3' DNA end is covered in stretches with the recombinase protein (RAD51 in human cells) and subsequently hybridizes with the temporarily opened homologous region in the template (Strand Invasion). Missing

information is then resynthesized into the damaged strand via a DNA polymerase. Lastly, the DNA heteroduplex is resolved and the broken ends of the damaged strand are ligated back together, bringing it back to the exact pre-damage state (Kanaar et al., 1998; Wright et al., 2018; Wyman and Kanaar, 2006; Wyman et al., 2004).

This complicated and not yet fully understood process is carried out by collections of proteins. The roles these proteins play along the steps of HR are usually known by now: there is no doubt that MR(N) takes part in signalling, resection and tethering (Wyman and Kanaar, 2006); similarly, it is known that BRCA2 facilitates RAD51 filament formation (Davies et al., 2001). The obscure part is how they do it. It is one thing to assert that MR(N) is present at the site of damage during the initial steps of HR (Deshpande et al., 2016a) and that it is capable of inducing ATM signalling cascade (Deshpande et al., 2016b), removing one strand of the broken DNA end (Paull and Gellert, 1999), and keep the ends in close proximity (de Jager et al., 2001; Moreno-Herrero et al., 2005), but it doesn't offer answers on how exactly these processes are carried out. One way to approach this problem is to look at the actions performed by proteins as work done with tools.

What makes a tool? Structure-to-function characteristics of proteins

While DNA is the information bank of the cell, proteins can be viewed as its main workforce. Proteins and protein complexes act as coordinators (in signalling pathways), scaffolds (as structural proteins) and precise executors (as enzymes) of all known cellular functions. More often than not, these categories overlap,

especially in the cases of protein complexes where individual components compliment their functions. There are many examples of scaffolding proteins that also have active role in signaling and display enzymatic functions (Smith and Scott, 2013). Human protein complex MRE11/RAD50/NBS1, which is a major interest in this work, is an excellent example. The MRN complex is both an ATPase and a nuclease (enzymatic activities) (Deshpande et al., 2016a), activates the ATM checkpoint regulation (signalling role) (Deshpande et al., 2016b), and ensures the stability of broken DNA strands by physically holding them together in close proximity (structural function) (de Jager et al., 2001; Moreno-Herrero et al., 2005). It is becoming increasingly obvious, that protein and protein complexes have to be regarded as executors of not one, but many functions, often within one task or pathway.

Properties of proteins and protein complexes can be studied using various methods, starting from *in vitro* biochemical assays, up to using genetic animal models. In the light of versatility of functions discussed above, however, the question is not only *what* a protein can do, but also *how* it does it. Function-oriented assays can rarely provide a direct answers to the latter question. To properly understand the mechanisms underlying protein activities, we should approach the problem from both functional and structural perspective. In other words, we should think of a protein as a tool.

What defines a tool and gives it specific properties is the shape. A nutcracker can perform its function owing to its specific design—a niche for the nut and two levers to apply pressure. However, the nutcracker would be still useless as a tool if it

wouldn't be possible to purposefully rearrange its shape, by opening and closing the levers. Similarly, the shape of proteins (conformation) and the ability to change shape (conformational changes) will enable them to perform assigned functions with efficiency and accuracy. The structure-to-function character is what truly defines a protein as a tool.

The architecture of proteins is determined by their amino acid sequence (Uversky, 2016). Depending on composition, amino acid strands will fold and form secondary and tertiary structures, driven mostly by electrostatic and hydrophobic forces. The protein can then undergo a final folding step, by association with protein molecules of the same or different kind (quaternary structure) (Ellgaard et al., 2016). The final result can range from globular or elongated, to highly irregular molecules. The architecture will determine features of the protein-tool, like spatial organization flexibility or persistence length, which will all affect the function. The initial conformation can be changed by a range of factors: post-translational modifications, ligand binding, binding of partner protein, nucleotide hydrolysis, or even change in pH, red-ox state or temperature of the environment (Gerstein et al., 1994; Grant et al., 2010).

Examples of structure-to-function dependency can be found in literature in great abundance. Cyclical conformational changes of Myosin, fuelled by ATP hydrolysis, allow "walking" on Actin fibres and lead to cellular filament contraction (Molloy et al., 1995; Sun et al., 2007). Estrogen receptor activation requires conformational change triggered by substrate binding (Fritsch et al., 1992). DNA polymerase, in order to ensure faithful replication

of the genetic information, undergoes a series of dNTP-driven conformation changes before and immediately after incorporating a nucleotide into newly synthesized DNA strand (Doublie et al., 1999; Doublie et al., 1998; Meli et al., 2016). DNA mismatch protein MutS is an elegant, well studied example of structure-to-function relation. MutS, upon finding a mismatch base pair binds to it, forming an immobile DNA-protein complex. The protein can then bind ATP and undergo a conformational change, changing into a sliding clamp that can move freely along the DNA strand and recruit other components of the machinery required to repair the lesion (Groothuizen et al., 2015).

Function without structure: Intrinsically Disordered Proteins and Regions

Architectural functionality is a crucial characteristic of proteins that has been used to investigate and unravel their properties in the span of last decades of research. That approach enforced the understanding that proteins can be either folded, e.g. achieved their functional conformation and able to undergo biologically relevant conformational changes, or un/misfolded, where the proper conformation is not achieved/corrupted, rendering the protein dysfunctional (or even pathogenic) (Gregersen et al., 2006). In the early years of the current century a new class of proteins was described: proteins that do not have a set, default conformation, but rather exist in a constant flux of various conformational states. They were dubbed Intrinsically Disordered Proteins (IDP's)(Uversky, 2002; Uversky and Dunker, 2010), or Intrinsically Disordered Protein Regions (IDPR's)(Uversky,

2016), if only a portion of a protein displays the IDP characteristics.

What distinguishes IDP's from un/misfolded proteins is that the former retain their functionality despite not having a definable structure. Intrinsic disorder is dependent on amino acid composition of the polypeptide chain. Charged, hydrophilic amino acids are known as “disorder-promoting”, creating patterns of uniformly charged blocks that diminish the possibility of folding via electrostatic and hydrophobic forces (Uversky et al., 2000). Otherwise stably folded proteins can also contain such patches, rendering part of the structure disordered in character, a phenomenon called Intrinsically Disordered Protein Regions (IDPR's) (Oldfield and Dunker, 2014; Sanchez et al., 2017; Uversky, 2016).

Over 30% of eukaryotic proteins have been as far identified as IDP's, or containing IDPR's (Dunker et al., 2001). On the first glance it is puzzling what kind of advantage could the cell gain by creating structurally unstable components. The ability to create phase transitions as means of separating protein foci from the rest of cyto/nucleoplasm by IDP(R)'s was pointed at in recent studies (Nott et al., 2015; Protter et al., 2018), but it certainly doesn't exhaust all the possibilities. Perhaps counter intuitively, the role of unstructured proteins is strictly related to *structure*—or, rather rapid changes of thereof. By their nature, IDP(R)'s are highly dynamic, existing in a variety of states almost simultaneously. That means that a conformation required to perform a specific task exists at any given moment and can be effortlessly changed to fulfill another if necessary. Additionally, IDP's are promiscuous

binders (Oldfield and Dunker, 2014; Uversky, 2016), giving them great versatility at different time points and cellular locations.

Crucial characteristics of IDP's are their constant transitions between disordered and ordered states. An IDP(R) can temporarily assume a stable conformation in a variety of conditions: environmental (red-ox state, temperature, ion strength of the surrounding cyto/nucleoplasm), post-translational modifications, complex formation (including self-association) (Nott et al., 2015; Sanchez et al., 2017; Smith and Jelokhani-Niaraki, 2012). This unique trait gives them the ability to perform an endless list of tasks: IDP's act as scaffolds, mortars, molecular glue, molecular switches and many more (Uversky, 2016).

Structured or not, the protein exists to fulfil specific function/functions within the cell. With the advancement of biochemical techniques and cellular and animal models functions of individual proteins are unravelled almost daily. More and more, however, the question that remains unanswered is how exactly these functions are fulfilled—it is one thing to know *what* is done, quite another *how* exactly it is done. If proteins are to be viewed as tools, then one can infer a lot of how they perform their tasks from how these tools are built. In other words, a structure-to-function relationship can be established. In order to speculate on structure-to-function character of the protein, however, one has to know its structure and, crucially, what kind of changes can this structure undergo. One has to “see” proteins.

Methods of studying protein architecture

Studying the shape of objects 10,000 times smaller than the width of the human hair naturally demands advanced approaches. An array of biochemical methods has been devised to unravel protein structures and relate them to function. Most famously, crystallography can be employed to reveal positions of individual atoms within the structure with median accuracy of 2 Å (Wlodawer et al., 2008). Circular Dichroism (CD)(Greenfield, 2006) and Nuclear Magnetic Resonance (NMR)(Wuthrich, 1989) are examples of useful spectroscopic methods (NMR producing results comparable to crystallography in terms of resolution). Cryogenic Electron Microscopy (Cryo-EM) is gaining more and more popularity in studying protein architecture, since the achieved resolution can match those of crystallography (Carroni and Saibil, 2016). Small-angle X-ray Scattering (SAXS) is a very useful method that allows studying protein conformation in solution (Kikhney and Svergun, 2015). Scanning Force Microscopy (SFM) is a method relying on direct contact of the measuring apparatus with the surface of the sample that can provide topographical reconstructions of proteins with high resolution (Allison et al., 2010; Binnig et al., 1986).

Despite a wide array of methods available, practical challenges cause the structure of many proteins to remain unsolved. Most basic problem lies in the efficiency of protein purification. Save for Cryo-EM and SFM, all aforementioned methods require high (mili- to micromolar) concentrations of material in the samples (Greenfield, 2006; Kikhney and Svergun, 2015; McPherson and Gavira, 2014; Wuthrich, 1989). While this is

usually attainable for smaller proteins, proteins with higher Molecular Weight are a challenge to purify in required amounts. Size itself is a limiting factor, which excludes proteins above 50 kDa from NMR studies (Wuthrich, 1989). Flexible proteins are virtually impossible to crystallise, which all but excludes IDP's from detailed structural studies. All things considered, if one sets on studying the structure of large, flexible, difficult to purify proteins, the task is going to be heinous.

SFM is an imaging technique based on an essentially different principle than methods discussed above. It relies on direct interactions of a microscopic tip situated on a flexible lever and the sample surface. While dragging or tapping on the surface, the tip climbs up and down every object that will be lying in its way. These ascensions and descensions are registered in real-time by a laser-photodiode system and computed into a 2-dimensional topographic image. Essentially, the tip is “feeling” and “mapping” the sample surface, presenting the shape of everything it has “met” on its way with nanometer resolution (Binnig et al., 1986). Comparing to other techniques, the resolution obtained by SFM is rather low (nanometers vs Ångstroms)—it does, however overcome most difficulties. Low concentration of a sample, for instance, is not a problem, since the preparations have to be diluted anyway to avoid overcrowding and only a few nanograms of protein are enough for an image of good quality (Allison et al., 2010). Size and flexibility do not matter, as the preparation is fixed on top of the sample surface and will be visualized by the tip—big or small, long or short. The opportunities to study big, flexible proteins by SFM are matched only by Cryo-EM. The latter, however, requires costly and difficult preparation techniques, and

has to be followed by complicated data analysis. SFM is a cheap technique, where sample preparation and imaging is a matter of minutes and user- friendly data analysis software is readily available online (Sanchez and Wyman, 2015).

Perhaps the most important quality of SFM imaging that sets it apart from the rest of aforementioned techniques is that it is the most direct. Instead of reconstructions of *statistically most prevalent* form of the studied protein, each molecule appears frozen in the conformation assumed in the moment of sample preparation. This is extremely important while studying flexible, dynamic proteins, as it allows to obtain whole collections of possible structural arrangements. Similarly, unravelling conformation changes upon complex formation, binding of a co-factor or in different environmental conditions is possible with high efficiency and accuracy, and is very often just a matter of adding an extra component to the test tube.

Scope of the thesis

In this work, we aim to investigate structure-to-function properties of big, flexible DNA-associating proteins. We employ SFM imaging to reveal dynamic conformation changes that remain inaccessible when using other biochemical techniques. The first part is devoted to studying architectural rearrangements of two SMC proteins—MukB, a bacterial condensin, and RAD50/MRE11 complex, a human DNA repair protein. Structure of both, despite years of research devoted to the subject, remains enigmatic, which creates gaps in our understanding of their function. In the second part of the thesis we investigate the structure and conformation changes of the famous BRCA2 protein. Large,

difficult to purify, containing multiple IDPR's, BRCA2's structure has puzzled researchers for years. Using SFM imaging, we unravel the dynamic character of this protein and translate our findings into its function within Homologous Recombination. Additionally, we summarize our advances in developing state-of-the-art imaging techniques.

References

Allison, D.P., Mortensen, N.P., Sullivan, C.J., and Doktycz, M.J. (2010). Atomic force microscopy of biological samples. *Wiley Interdiscip Rev Nanomed Nanobiotechnol* 2, 618-634.

Binnig, G., Quate, C.F., and Gerber, C. (1986). Atomic force microscope. *Phys Rev Lett* 56, 930-933.

Carroni, M., and Saibil, H.R. (2016). Cryo electron microscopy to determine the structure of macromolecular complexes. *Methods* 95, 78-85.

Crick, F. (1970). Central dogma of molecular biology. *Nature* 227, 561-563.

Crick, F.H., Barnett, L., Brenner, S., and Watts-Tobin, R.J. (1961). General nature of the genetic code for proteins. *Nature* 192, 1227-1232.

Davies, A.A., Masson, J.Y., McIlwraith, M.J., Stasiak, A.Z., Stasiak, A., Venkitaraman, A.R., and West, S.C. (2001). Role of BRCA2 in control of the RAD51 recombination and DNA repair protein. *Mol Cell* 7, 273-282.

de Jager, M., van Noort, J., van Gent, D.C., Dekker, C., Kanaar, R., and Wyman, C. (2001). Human Rad50/Mre11 is a flexible complex that can tether DNA ends. *Mol Cell* 8, 1129-1135.

Deshpande, R.A., Lee, J.H., Arora, S., and Paull, T.T. (2016a). Nbs1 Converts the Human Mre11/Rad50 Nuclease Complex into an Endo/Exonuclease Machine Specific for Protein-DNA Adducts. *Mol Cell* 64, 593-606.

Deshpande, R.A., Williams, G.J., Limbo, O., Williams, R.S., Kuhnlein, J., Lee, J.H., Classen, S., Guenther, G., Russell, P., Tainer, J.A., *et al.* (2016b). ATP-driven Rad50 conformations regulate DNA tethering, end resection, and ATM checkpoint signaling. *EMBO J* 35, 791.

Doublie, S., Sawaya, M.R., and Ellenberger, T. (1999). An open and closed case for all polymerases. *Structure* 7, R31-35.

Doublie, S., Tabor, S., Long, A.M., Richardson, C.C., and Ellenberger, T. (1998). Crystal structure of a bacteriophage T7 DNA replication complex at 2.2 Å resolution. *Nature* 391, 251-258.

Dunker, A.K., Lawson, J.D., Brown, C.J., Williams, R.M., Romero, P., Oh, J.S., Oldfield, C.J., Campen, A.M., Ratliff, C.M., Hipps, K.W., *et al.* (2001). Intrinsically disordered protein. *J Mol Graph Model* 19, 26-59.

Ellgaard, L., McCaul, N., Chatsisvili, A., and Braakman, I. (2016). Co- and Post-Translational Protein Folding in the ER. *Traffic* 17, 615-638.

Fritsch, M., Leary, C.M., Furlow, J.D., Ahrens, H., Schuh, T.J., Mueller, G.C., and Gorski, J. (1992). A ligand-induced conformational change in the estrogen receptor is localized in the steroid binding domain. *Biochemistry* 31, 5303-5311.

Gerstein, M., Lesk, A.M., and Chothia, C. (1994). Structural mechanisms for domain movements in proteins. *Biochemistry* 33, 6739-6749.

Giglia-Mari, G., Zotter, A., and Vermeulen, W. (2011). DNA damage response. *Cold Spring Harb Perspect Biol* 3, a000745.

Grant, B.J., Gorfe, A.A., and McCammon, J.A. (2010). Large conformational changes in proteins: signaling and other functions. *Curr Opin Struct Biol* 20, 142-147.

Green, C.M., and Almouzni, G. (2002). When repair meets chromatin. First in series on chromatin dynamics. *EMBO Rep* 3, 28-33.

Greenfield, N.J. (2006). Using circular dichroism spectra to estimate protein secondary structure. *Nat Protoc* 1, 2876-2890.

Gregersen, N., Bross, P., Vang, S., and Christensen, J.H. (2006). Protein misfolding and human disease. *Annu Rev Genomics Hum Genet* 7, 103-124.

Groothuizen, F.S., Winkler, I., Cristovao, M., Fish, A., Winterwerp, H.H., Reumer, A., Marx, A.D., Hermans, N., Nicholls, R.A., Murshudov, G.N., *et al.* (2015). MutS/MutL crystal structure reveals that the MutS sliding clamp loads MutL onto DNA. *Elife* 4, e06744.

Hoeijmakers, J.H. (2009). DNA damage, aging, and cancer. *N Engl J Med* 361, 1475-1485.

Kanaar, R., Hoeijmakers, J.H., and van Gent, D.C. (1998). Molecular mechanisms of DNA double strand break repair. *Trends Cell Biol* 8, 483-489.

Kikhney, A.G., and Svergun, D.I. (2015). A practical guide to small angle X-ray scattering (SAXS) of flexible and intrinsically disordered proteins. *FEBS Lett* 589, 2570-2577.

Lesnik, E.A., and Freier, S.M. (1995). Relative thermodynamic stability of DNA, RNA, and DNA:RNA hybrid duplexes: relationship with base composition and structure. *Biochemistry* 34, 10807-10815.

McPherson, A., and Gavira, J.A. (2014). Introduction to protein crystallization. *Acta Crystallogr F Struct Biol Commun* 70, 2-20.

- Meli, M., Sustarsic, M., Craggs, T.D., Kapanidis, A.N., and Colombo, G. (2016). DNA Polymerase Conformational Dynamics and the Role of Fidelity-Confering Residues: Insights from Computational Simulations. *Front Mol Biosci* 3, 20.
- Molloy, J.E., Burns, J.E., Kendrick-Jones, J., Tregear, R.T., and White, D.C. (1995). Movement and force produced by a single myosin head. *Nature* 378, 209-212.
- Moreno-Herrero, F., de Jager, M., Dekker, N.H., Kanaar, R., Wyman, C., and Dekker, C. (2005). Mesoscale conformational changes in the DNA-repair complex Rad50/Mre11/Nbs1 upon binding DNA. *Nature* 437, 440-443.
- Nott, T.J., Petsalaki, E., Farber, P., Jervis, D., Fussner, E., Plochowietz, A., Craggs, T.D., Bazett-Jones, D.P., Pawson, T., Forman-Kay, J.D., *et al.* (2015). Phase transition of a disordered nuage protein generates environmentally responsive membraneless organelles. *Mol Cell* 57, 936-947.
- Oldfield, C.J., and Dunker, A.K. (2014). Intrinsically disordered proteins and intrinsically disordered protein regions. *Annu Rev Biochem* 83, 553-584.
- Paull, T.T., and Gellert, M. (1999). Nbs1 potentiates ATP-driven DNA unwinding and endonuclease cleavage by the Mre11/Rad50 complex. *Genes Dev* 13, 1276-1288.
- Ponomarenko, E.A., Poverennaya, E.V., Ilgisonis, E.V., Pyatnitskiy, M.A., Kopylov, A.T., Zgoda, V.G., Lisitsa, A.V., and Archakov, A.I. (2016). The Size of the Human Proteome: The Width and Depth. *Int J Anal Chem* 2016, 7436849.
- Protter, D.S.W., Rao, B.S., Van Treeck, B., Lin, Y., Mizoue, L., Rosen, M.K., and Parker, R. (2018). Intrinsically Disordered Regions Can Contribute Promiscuous Interactions to RNP Granule Assembly. *Cell Rep* 22, 1401-1412.

Sanchez, H., Paul, M.W., Grosbart, M., van Rossum-Fikkert, S.E., Lebbink, J.H.G., Kanaar, R., Houtsmuller, A.B., and Wyman, C. (2017). Architectural plasticity of human BRCA2-RAD51 complexes in DNA break repair. *Nucleic Acids Res* *45*, 4507-4518.

Sanchez, H., and Wyman, C. (2015). SFMetrics: an analysis tool for scanning force microscopy images of biomolecules. *BMC Bioinformatics* *16*, 27.

Smith, F.D., and Scott, J.D. (2013). Scaffolding proteins: not such innocent bystanders. *Curr Biol* *23*, R515-517.

Smith, M.D., and Jelokhani-Niaraki, M. (2012). pH-induced changes in intrinsically disordered proteins. *Methods Mol Biol* *896*, 223-231.

Sun, Y., Schroeder, H.W., 3rd, Beausang, J.F., Homma, K., Ikebe, M., and Goldman, Y.E. (2007). Myosin VI walks "wiggly" on actin with large and variable tilting. *Mol Cell* *28*, 954-964.

Uversky, V.N. (2002). Natively unfolded proteins: a point where biology waits for physics. *Protein Sci* *11*, 739-756.

Uversky, V.N. (2016). Dancing Protein Clouds: The Strange Biology and Chaotic Physics of Intrinsically Disordered Proteins. *J Biol Chem* *291*, 6681-6688.

Uversky, V.N., and Dunker, A.K. (2010). Understanding protein non-folding. *Biochim Biophys Acta* *1804*, 1231-1264.

Uversky, V.N., Gillespie, J.R., and Fink, A.L. (2000). Why are "natively unfolded" proteins unstructured under physiologic conditions? *Proteins* *41*, 415-427.

Watanabe, H., Takehana, K., Date, M., Shinozaki, T., and Raz, A. (1996). Tumor cell autocrine motility factor is the neuroleukin/phosphohexose isomerase polypeptide. *Cancer Res* *56*, 2960-2963.

Wlodawer, A., Minor, W., Dauter, Z., and Jaskolski, M. (2008). Protein crystallography for non-crystallographers, or how to get the best (but not more) from published macromolecular structures. *FEBS J* 275, 1-21.

Wright, W.D., Shah, S.S., and Heyer, W.D. (2018). Homologous recombination and the repair of DNA double-strand breaks. *J Biol Chem* 293, 10524-10535.

Wuthrich, K. (1989). Protein structure determination in solution by nuclear magnetic resonance spectroscopy. *Science* 243, 45-50.

Wyman, C., and Kanaar, R. (2006). DNA double-strand break repair: all's well that ends well. *Annu Rev Genet* 40, 363-383.

Wyman, C., Ristic, D., and Kanaar, R. (2004). Homologous recombination-mediated double-strand break repair. *DNA Repair (Amst)* 3, 827-833.

Xu, W., Seiter, K., Feldman, E., Ahmed, T., and Chiao, J.W. (1996). The differentiation and maturation mediator for human myeloid leukemia cells shares homology with neuroleukin or phosphoglucose isomerase. *Blood* 87, 4502-4506.

Chapter 2

The Bacterial Condensin MukB Compacts DNA by Sequestering Supercoils and Stabilizing Topologically Isolated Loops

R Kumar¹, M Grosbart², P Nurse¹, S Bahng¹, C. L. Wyman^{2,3}, K. J. Marians¹

¹Molecular Biology Program, Memorial Sloan Kettering Cancer Center, New York, New York 10065

Department of ²Molecular Genetics and ³Radiation Oncology, Erasmus University Medical Center, P. O. Box 2040, 3000CA Rotterdam, The Netherlands

Published in Journal of Biological Chemistry 2017 292, 16904-16920

Abstract

MukB is a Structural Maintenance of Chromosome-like protein required for DNA condensation. The complete condensin is a large tripartite complex of MukB, the kleisin, MukF, and an accessory protein, MukE, that may trap DNA topologically. We find instead that MukB DNA condensation is a stepwise process that proceeds first *via* the formation of negative supercoils that are sequestered by the protein followed by hinge-hinge interactions between MukB dimers that stabilize topologically isolated loops in the DNA that can be supercoiled by DNA gyrase. DNA condensation by MukB did not require and was unaffected by either ATP or MukEF, indicating that MukB itself is sufficient to condense DNA. We show that the MukB hinge region binds DNA and that this region of the protein is involved in sequestration of supercoils. Cells carrying mutations in the MukB hinge that reduce DNA condensation *in vitro* exhibit nucleoid decondensation, and dispersal of MukB *in vivo*. The mode of DNA condensation we describe is unique among SMC proteins.

INTRODUCTION

Chromosomal DNA must be condensed in an orderly manner that allows access by the enzymes that transcribe, repair, and replicate the chromosomes. DNA condensation is achieved by a combination of a number of factors that include the binding of small proteins that shape the DNA, counterion condensation, modification of DNA topology, and the action of members of the Structural Maintenance of Chromosome (SMC) family of proteins.

SMC proteins (Uhlmann, 2016) are characterized by head domains that are separated by a long coiled coil region that folds back on themselves to bring the head domains, which generally bind DNA and possess ATPase activity, together. The apex of the coiled-coil region is referred to as the hinge region. Active SMC proteins are dimeric. The coiled coil region can be 40-50 nm in length. In eukaryotes there are three versions of SMC proteins, cohesin, condensin, and the SMC5/6 complex. Bacteria either contain a true SMC protein (e.g., in *B. subtilis*), or SMC-like proteins called MukB (e.g., in *E. coli*) or MksB (e.g., in *P. aeruginosa*). SMC proteins interact with other accessory proteins: the kleisin, which bridges the head domains of the SMC proteins to form a topologically closed ring of proteins, and a third protein that interacts with the kleisin.

Mutations in condensins cause chromosome decompaction and segregation defects. In *E. coli*, MukB is encoded in an operon along with its kleisin, MukF, and an additional accessory protein MukE (Niki et al., 1991; Yamanaka et al., 1996). Mutations in *mukB*, *mukF*, and *mukE* cause nucleoid decondensation, chromosome segregation defects, and the generation of anucleate

cells (Niki et al., 1991; Yamanaka et al., 1996). MukB binds double-stranded DNA, and can alter DNA topology when bound to DNA in the presence of a topoisomerase, forming right-handed knots and negative supercoils in DNA rings (Petrushenko et al., 2006a). Interestingly, whereas MukF and MukE form a tripartite complex with MukB (Yamazoe et al., 1999), and are required for localization of MukB in the cell (Badrinarayanan et al., 2012; Ohsumi et al., 2001; She et al., 2007), they inhibit the DNA binding and topology modification activities of MukB *in vitro* (Petrushenko et al., 2006b) while stimulating the ATPase activity (Chen et al., 2008). ATP hydrolysis causes a rearrangement of the MukBEF complex, loosening the contacts between MukF and one MukB head domain in the MukB dimer (Woo et al., 2009). These observations suggest that MukF and MukE may be involved in the loading of MukB to DNA rather than in DNA condensation *per se*.

How these SMC complexes act to cohes and condense DNA has been under active investigation for some time. The eukaryotic cohesin and condensin both trap DNA helices topologically within the tripartite protein complex (Uhlmann, 2016). Recent studies also reported the same mechanism held true for the *B. subtilis* SMC protein (Wilhelm et al., 2015) and MukB (Niki and Yano, 2016). Several models have been elaborated as to how this topological trapping of DNA can lead to chromosome condensation (Uhlmann, 2016).

We have investigated the mechanism of DNA condensation by MukB. We found that unlike either the eukaryotic or *B. subtilis* condensins, MukB did not entrap DNA topologically, but progressively condensed DNA by first forming and sequestering

negative supercoils and then forming larger loops that were mediated by hinge-hinge interactions between MukB dimers. These large loops of DNA could be supercoiled by DNA gyrase. DNA condensation did not require and was unaffected by either ATP or MukEF.

RESULTS

MukB Protects Negative Supercoils and Forms Topologically Stable Loops in the DNA

We asked whether we could detect direct alteration of DNA topology by MukB in the absence of a topoisomerase. Because agarose gel electrophoresis of DNA rings is very sensitive to DNA topology (15), we used this technique as the basis of our assay. The substrate was an 11-kbp dsDNA plasmid that had been relaxed by the introduction of a single nick in one of the DNA strands. Remarkably, under conditions previously established for examining MukB binding to DNA (4, 9), protein-DNA complexes formed at low concentrations of MukB moved faster during electrophoresis than the nicked DNA (Fig. 1A). As the concentration of MukB was increased, the mobility of the protein-DNA complex became retarded compared with that of the DNA alone (Fig. 1A). We refer to the protein-DNA complexes formed at 31 and 63 nM MukB as the fast-moving complexes (FMcx) and those formed at 125 and 250 nM MukB as the slow-moving complexes (SMcx). It is important to note that these reaction mixtures contained only MukB and the relaxed DNA. There was neither topoisomerase, ATP, nor MukEF present. The fact that the FMcx had a greater mobility in the gel than that of the DNA alone suggested that the DNA must be condensed. We investigated the

nature of this condensation using a number of biochemical assays, as well as scanning force microscopy (SFM).

Because MukB stabilizes negative supercoils (4) and supercoiling of the DNA increases its electrophoretic mobility, we asked whether MukB was inducing supercoils in the nicked DNA. The MukB-DNA complex was incubated with *E. coli* DNA ligase, which uses NAD as a cofactor rather than ATP (16), to seal the nick. After deproteinization, DNA products were analyzed by electrophoresis through an agarose gel containing chloroquine (Fig. 1B). Chloroquine unwinds the DNA helix, and thus in a relaxed, closed circular DNA (the product of sealing the nick) it will induce positive supercoils, and the mobility of the DNA will increase (Fig. 1B, compare lanes 1 and 2). DNA sealed in the presence of MukB contained negative supercoils, as indicated by the ladder of topoisomers with retarded mobility compared with that of the relaxed DNA alone (Fig. 1B, compare lane 2 to lanes 4–6).

To determine whether these negative supercoils were merely stabilized by MukB or actually protected by it, we exposed the protein-DNA complexes to two different type I topoisomerases that relax negative supercoils: *E. coli* topoisomerase I, which acts on single-stranded DNA (17), and vaccinia virus topoisomerase, which acts on dsDNA (18). Both the SMCx and FMCx were examined (Fig. 1C). In neither case did the topoisomerase reduce the mobility of the protein-DNA complex, arguing that the supercoils were unavailable to be relaxed. We conclude that the supercoils could not be relaxed because the helical crossover that forms them was completely protected by binding to MukB. Controls showing that the amount of topoisomerase used was

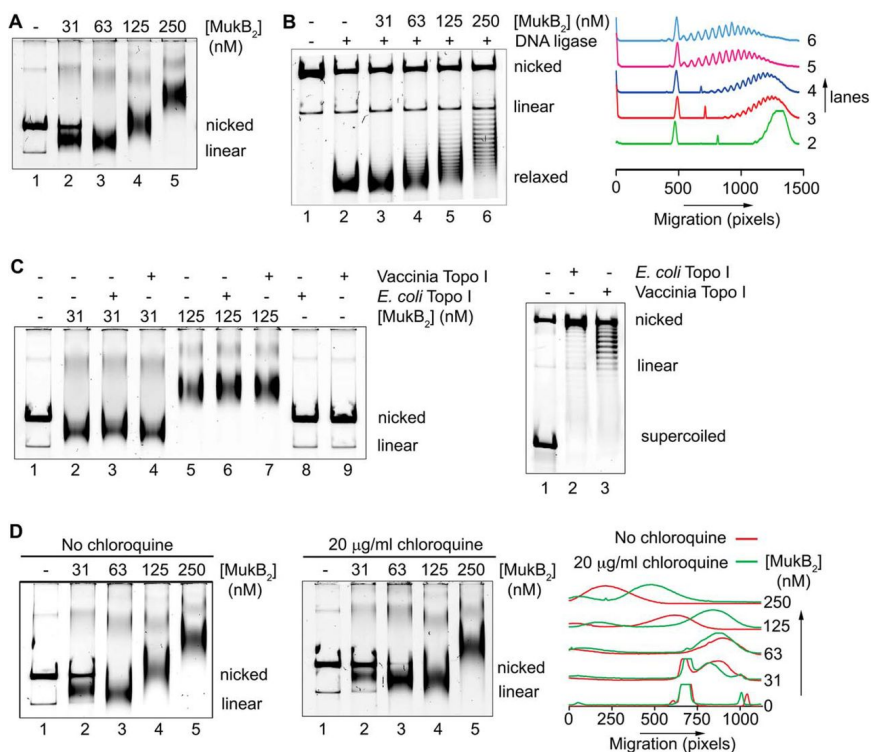


Fig 1. MukB protects negative supercoils and stabilizes loop formation in DNA that is not constrained topologically. **A**, gel mobility shift analysis of MukB binding to DNA. MukB was incubated with the nicked DNA substrate for 5 min at 37 °C and then analyzed by agarose gel electrophoresis as described under “Experimental procedures.” *Nicked*, nicked 11-kbp DNA substrate; *Linear*, linear form of the DNA substrate. **B**, MukB forms negative supercoils in the nicked DNA substrate. MukB, *E. coli* DNA ligase, and NAD were incubated with the nicked DNA substrate for 30 min at 37 °C. *Left panel*, products were deproteinized and then analyzed by electrophoresis through agarose gels containing 10 µg/ml chloroquine in both the gel and the running buffer as described under “Experimental procedures.” *Right panel*, densitometric lane traces of the gel shown in the *left panel*. **C**, MukB protects negative supercoils in the nicked DNA substrate. *Left panel*, MukB was incubated with the nicked DNA for 5 min at 37 °C. Either *Vaccinia* DNA topoisomerase or *E. coli* DNA topoisomerase I were added to 3.1 and 6.2 nM, respectively, as indicated, and the incubation continued for 30 min. The products were then analyzed by gel electrophoresis as described in **A** above. *Right panel*, concentrations of topoisomerase (*Topo*) I used was more than sufficient to completely relax the supercoiled form of the plasmid in the absence of bound MukB. **D**, MukB stabilizes topologically isolated loops in the DNA. *Left and middle panels*, MukB was incubated with the nicked DNA substrate for 5 min at 37 °C and then analyzed by electrophoresis through agarose gels either in the absence or presence of 20 µg/ml chloroquine, respectively, in the gel and running buffer. *Right panel*, densitometric lane traces of the gels shown in the *left and middle panels*.

sufficient to relax all the supercoiled DNA are shown in Fig. 1C (right panel).

Thus, it seemed likely that the FMcx was probably dominated by the formation of protected supercoils. The addition of more MukB to the reaction causes the mobility of the complex to slow. This effect could simply be because of the accretion of additional mass to the complex via the binding of more MukB. However, we wondered whether the effect could also be attributed to another distinct layer of topology imposed by the increasing amounts of bound MukB. To probe this possibility, we compared the mobility of the protein-DNA complexes in gels that either did or did not contain chloroquine (Fig. 1D). As expected, the mobility of the FMcx was reduced compared with the nicked DNA in the chloroquine-containing gel (Fig. 1D, compare lanes 1–3 in each panel). This effect is because chloroquine is a small molecule and can unwind even the MukB-protected supercoils, confirming that this FMcx is dominated by negative supercoiling.

Remarkably, the presence of chloroquine in the gel caused the SMcx to now move faster than the nicked DNA (Fig. 1D, compare lanes 4 and 5 in each panel). This inversion of DNA mobility can only be explained by the presence of relaxed loops in the SMcx that become positively supercoiled in the presence of chloroquine. In order for these loops to exist, they must be stabilized by MukB.

To confirm our interpretation, we reasoned that these loops should be stable enough to be supercoiled by the action of

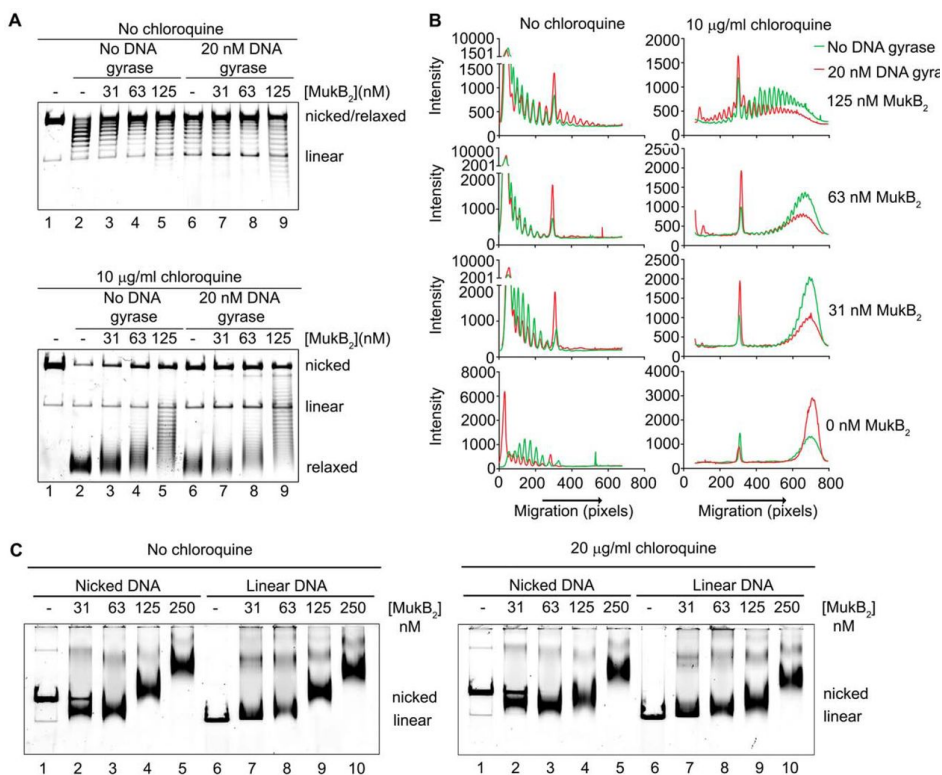


Fig 2. Loops formed by MukB in the DNA are topologically isolated. *A*, DNA loops stabilized by MukB can be supercoiled by DNA gyrase. MukB was incubated with the nicked DNA substrate, 2 mM ATP, and 20 nM DNA gyrase for 30 min at 37 °C. Novobiocin was added to 10 µM and the incubation continued for 5 min. Bacteriophage T4 DNA ligase was then added and the incubation continued for 30 min. The products were deproteinized before analysis by electrophoresis through agarose gels either in the absence (*top panel*) or presence (*bottom panel*) of 10 µg/ml chloroquine in both the gel and the running buffer. *Relaxed*, the nicked DNA sealed by DNA ligase to give a closed DNA ring that does not contain supercoils. *B*, densitometric lane traces of the gels shown in *A*. *C*, MukB stabilizes topologically isolated loops in linear DNA. MukB was incubated with either nicked DNA (*lanes 1–5*) or linear DNA (blunt-end cut, *lanes 6–10*) for 5 min at 37 °C. Protein-DNA complexes were then analyzed as in Fig. 1D above.

DNA gyrase (19). We therefore designed an experimental scheme so that gyrase would only be acting when the DNA was still nicked and thus not topologically constrained by virtue of the continuity of both DNA strands. MukB protein-DNA complexes were formed and then incubated with DNA gyrase and ATP (19). Novobiocin was then added to inactivate gyrase (20) followed by DNA ligase to seal the DNA and lock in any changes in DNA topology. After deproteinization, the DNA products were analyzed by gel electrophoresis in either the absence or presence of chloroquine (Fig. 2, A, top and bottom panels, respectively, and B). As before (Fig. 1B), the presence of MukB alone caused the accumulation of negative supercoils (Fig. 2, A and B, lanes 2–6). However, the addition of DNA gyrase to SMCxs clearly caused the accumulation of additional supercoils (Fig. 2, A and B, compare lane 6 to lane 9). Controls demonstrating the inhibition of DNA gyrase by novobiocin are shown in Fig. 8C (lanes 11 and 12).

We considered that some curvature of the DNA might be required for the action of MukB, as has been suggested for the yeast condensin (21). Accordingly, we asked whether we could detect topological loop stabilization when MukB was bound to linear DNA (Fig. 2C). This proved to be the case. MukB-linear DNA complexes formed at 125 and 250 nm MukB clearly displayed increased mobility in gels containing chloroquine compared with those that did not (Fig. 2C, compare lanes 4 and 5 and 9 and 10 in the presence and absence of chloroquine). We conclude that at higher concentrations MukB was stabilizing topologically isolated loops in dsDNA. These observations are consistent with the results of single molecule analyses of the effect of MukB binding on the

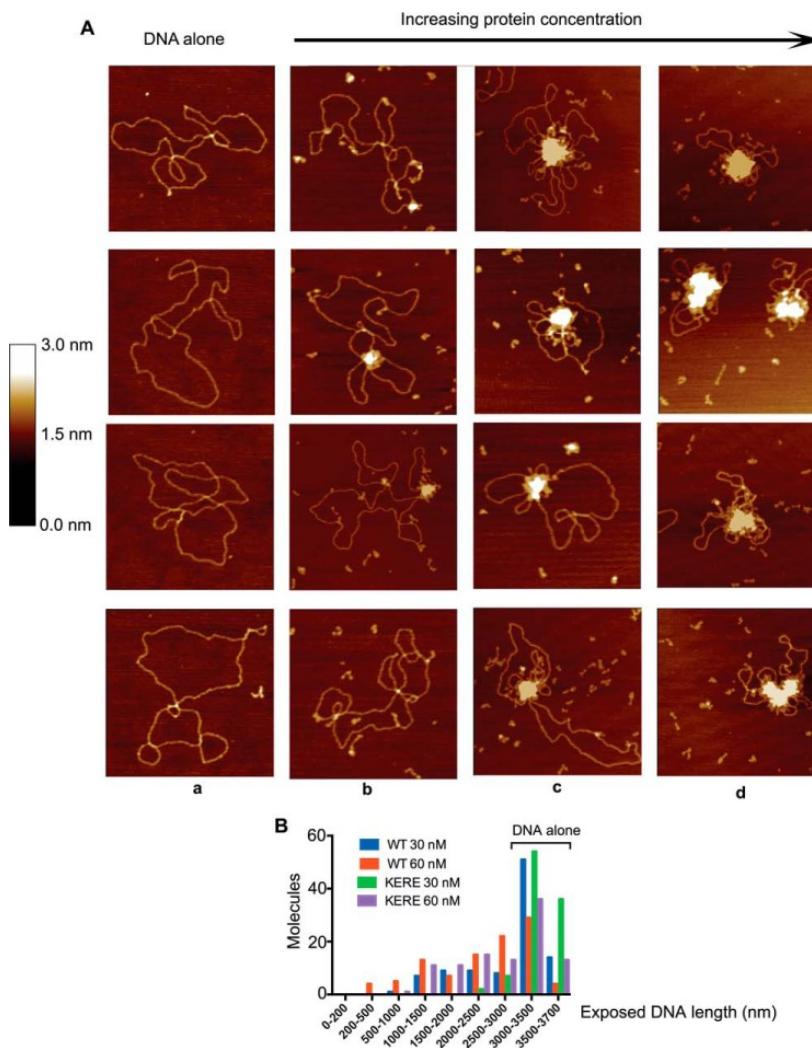


Fig 3. Progressive DNA condensation of DNA by MukB as visualized by scanning force microscopy. *A*, protein-DNA complexes formed with the nicked DNA substrate either in the absence of MukB (*column a*) or in the presence of 15 (*column b*), 30 (*column c*), or 60 nM MukB (*column d*) were imaged in the SFM as described under “Experimental procedures.” Images shown are for wild-type MukB. Images for MukB KE,RE were identical in appearance. *B*, extent of exposed DNA in the MukB-DNA complexes, measured as described under “Experimental procedures,” is presented as a function of MukB concentration. *WT*, wild type; *KERE*, MukB K761E/R765E protein variant. A total of 100 molecules was measured from three independent experiments

length of DNA that were interpreted as indicating MukB-mediated loop formation (14).

Scanning Force Microscopy of MukB-DNA Complexes: The MukB Hinge Region Binds DNA

MukB-DNA complexes were examined using SFM. As the MukB concentration increased, the protein-DNA complexes presented increasingly aggregated masses with unbound DNA extruding from the central mass (Fig. 3A). The extent of exposed DNA in the protein-DNA complex decreased as the concentration of MukB increased, suggesting that the SFM images corresponded to that of progressively condensed DNA. To assess this possibility, we used as a measure of DNA condensation the length of dsDNA not present in the aggregated mass. Unconstrained plasmid DNA was 3430 ± 174 nm long. As the concentration of MukB increased from 15 to 60 nM, the length of exposed DNA decreased (Fig. 3B), suggesting that the SFM images report on the structure of the same MukB-DNA complex that has an electrophoretic mobility greater than that of the nicked DNA alone (the FMcx). Therefore, under these conditions the DNA becomes condensed. Such aggregation of MukB-DNA complexes is consistent with cooperative binding of MukB to the DNA (14). It was not possible to examine MukB-DNA complexes formed at higher concentrations of MukB with the SFM because these appeared as highly aggregated forms.

At low concentrations of MukB, individual MukB molecules bound to the DNA could be observed (Fig. 4). Head and hinge domains could not be distinguished per se; however, we know that MukB is a V-shaped molecule with the hinge domain at the

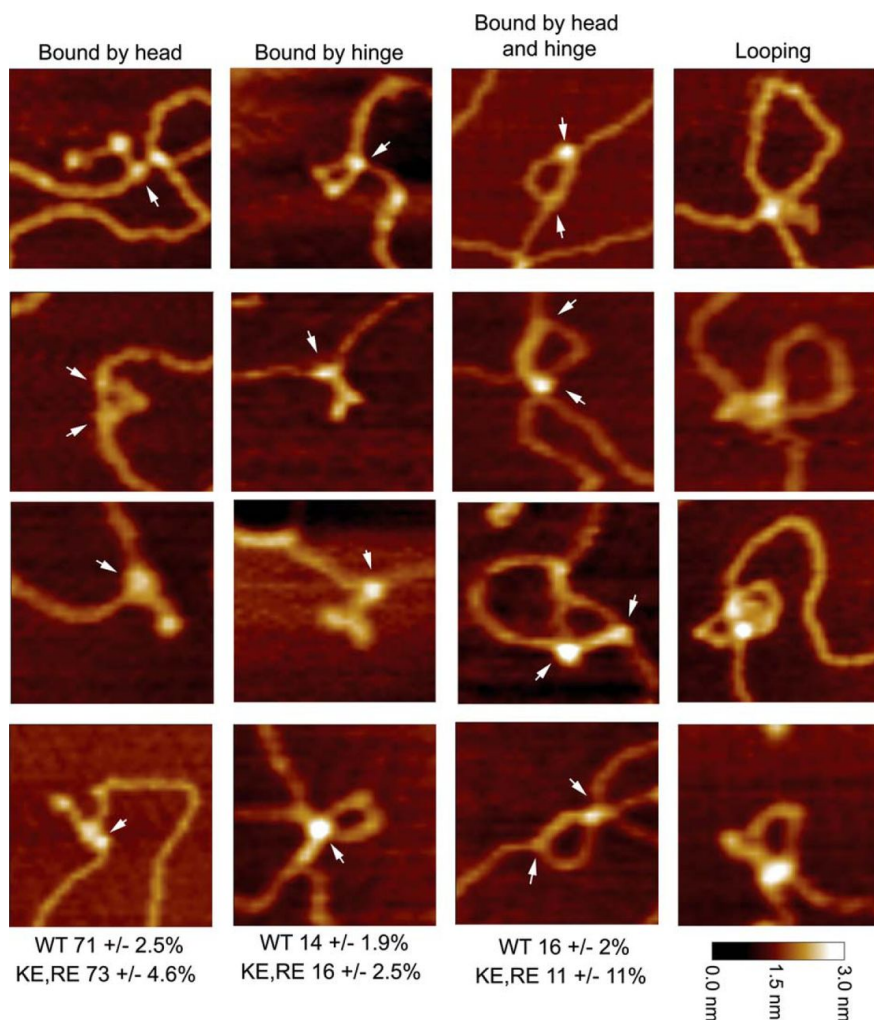


Fig 4. Modes of MukB binding to DNA as visualized by scanning force microscopy. One hundred molecules were counted from three independent experiments where MukB at either 7.5 or 15 nm was bound to the nicked DNA substrate, and the modes of binding were characterized as either MukB bound to the DNA by a head domain, MukB bound to the DNA by the hinge domain, or MukB bound to the DNA by both the hinge and head domains. The fraction of molecules bound in each mode for the wild type and the MukB KE,RE variant is given *below each column*. Images where the DNA is looped with MukB bound at the apex are also shown. Images shown are for wild-type MukB. Images for MukB KE,RE were identical in appearance. Arrows indicate points of contact between MukB and the DNA.

apex and the head domains on the opposite ends (22). Bound MukB molecules could be divided into three classes (Fig. 4): bound by the head domain; bound by the hinge domain; and bound by both the head and hinge domains. Anchoring of DNA loops by bound MukB was also observed. MukB binding via the hinge domain was unexpected because it had been reported previously that the isolated hinge domain did not bind DNA (23, 24). Binding of isolated MukB hinge to DNA was confirmed as described below.

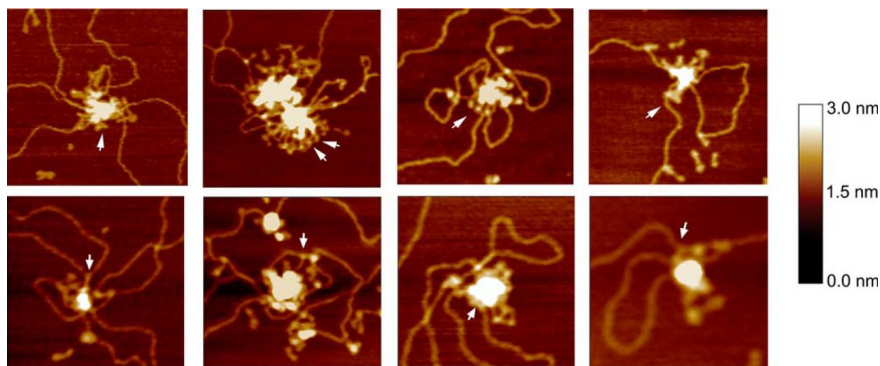


Fig 5. Close-up views of MukB-DNA aggregates. Close-up SFM views of MukB aggregates on the DNA showing individual MukB molecules protruding from the aggregated mass (arrows).

A Model for DNA Condensation by MukB

Individual MukB molecules could often be observed protruding from the aggregated mass in the SFM images (Fig. 5). In many instances, it seemed like the hinge domain was in the root of the aggregate with the head domains, occasionally bound to DNA, protruding (see arrows in Fig. 5). This possible feature of the aggregation, MukB protection of negative supercoils and formation of topologically stable loops, and hinge domain binding of DNA suggested a model for how MukB condenses DNA (Fig. 6,

A–C). We propose that MukB initially binds DNA such that a negative supercoil is protected by interactions with the DNA-binding domains on the head domains and the hinge domain (Fig. 6A). In this fashion, extraneous topoisomerases might not be able to relax the supercoil, whereas chloroquine would be able to bind the DNA and untwist it. As more MukB binds the DNA, interactions between the hinge domains of MukB dimers become more likely leading to the formation of loops (Fig. 6B). These loops are topologically isolated and can be supercoiled by DNA gyrase (Fig. 6C). Aspects of this model are similar to models proposed recently describing the action of the yeast condensin (21), MukB (25), and the *B. subtilis* SMC protein (26). Using single molecule analyses of the effect of MukB on the length of DNA with one end tethered to a glass slide and the other held in magnetic tweezers, Cui et al. (14) proposed a stepwise condensation of DNA and inferred from their results that one aspect of the condensation would be the formation of loops in the DNA.

Because our model seems counter to the topological entrapment model for SMC proteins binding DNA (1), we addressed this issue directly. First, we asked whether MukF or MukEF and ATP had an effect on MukB DNA condensation. MukEF appeared to stabilize the FMcx somewhat, but it also inhibited MukB binding overall. The mobility of the SMcx was unaffected, and ATP had little effect on either complex with or without MukEF being present (Fig. 6D). Thus, MukB alone is sufficient for DNA condensation, as has been demonstrated previously (14). Second, we used the assay developed by Uhlmann's laboratory for topological trapping (27). Here, MukB-HA bound to a mixture of

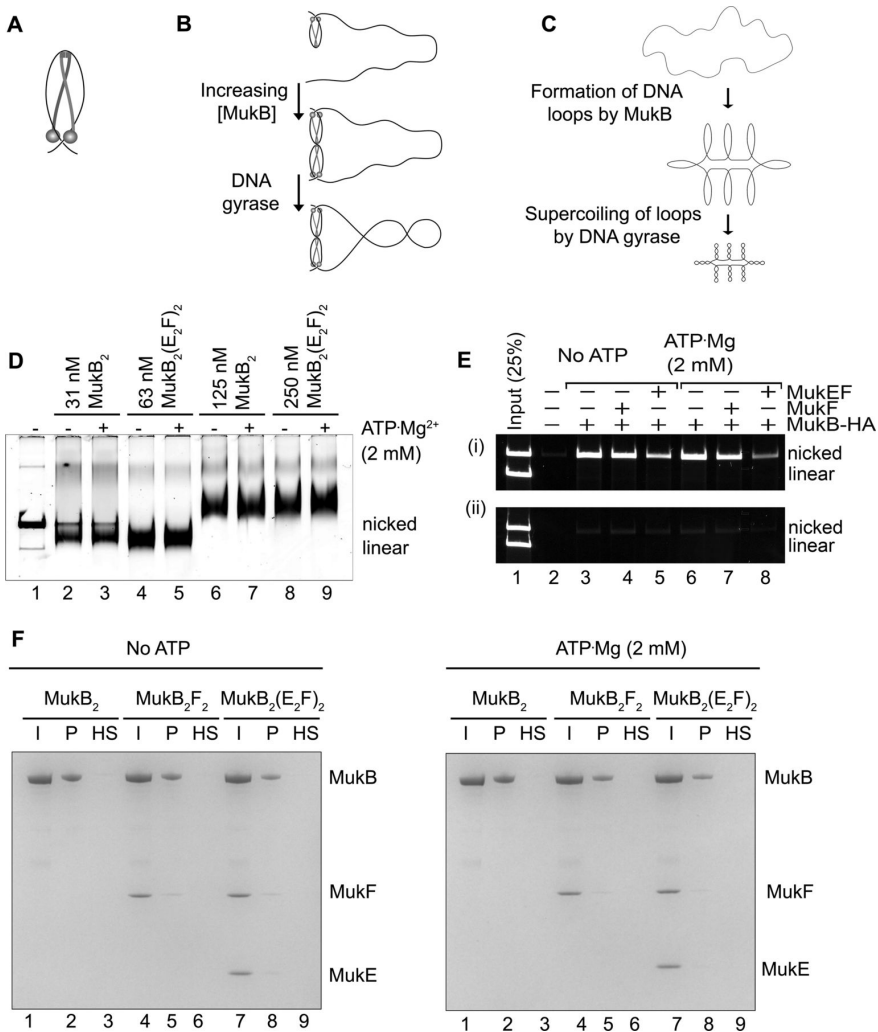


Fig 6. Model for MukB condensation of DNA. *A*, MukB protects a negative supercoil when bound to DNA. We propose that DNA binding by both the head domain and hinge domain is necessary for efficient negative supercoiling. The disposition of the DNA on the MukB hinge should not be taken literally. Whereas the amino acid residues mutated in the KE,RE variant are on the top of the hinge, we do not know precisely how DNA is bound to the hinge. One presumes that the DNA binds in such a manner that it does not interfere with MukB oligomerization via hinge–hinge interactions. *B*, MukB can form stable, topologically isolated loops in the DNA via hinge–hinge interactions between dimers. *C*, progressive DNA condensation by MukB is driven by the formation of topological loops in the DNA that can be further negatively supercoiled by DNA gyrase. *D*, effect of MukEF and ATP on MukB-DNA complexes. The indicated concentrations of MukB or the MukBEF complex in the presence or absence of ATP were incubated for 5 min at 37 °C. Protein-DNA complexes were then analyzed as described in the legend to Fig. 1A. Note that twice the concentration of MukBEF was required to observe similar DNA-binding activity compared with MukB alone. *E*, MukB does not trap DNA topologically. The indicated concentrations of MukB-HA, MukF, and MukEF, either in the presence or absence of ATP, were incubated with an equal mixture of nicked and linear DNA for 5 min at 37 °C. *Panel i*, anti-HA antibody attached to magnetic beads was then added and immediately thereafter NaCl was added to 0.5 m, the beads were washed in 0.75 m NaCl after pulldown, and the bound DNA was released by proteinase K digestion and analyzed by agarose gel electrophoresis. *Panel ii* shows the same experiment except that 0.5 m NaCl was added immediately before addition of the antibody beads. *F*, MukB does not trap DNA topologically. Tailed, nicked plasmid DNA attached to magnetic beads was incubated with the indicated concentrations of MukB, MukF, and MukEF either in the presence or absence of ATP in standard reaction buffer (low salt) for 5 min at 37 °C. The beads were then pulled down and resuspended in buffer containing 0.5 m NaCl. The beads were pulled down again, and the samples were resuspended in 1× SDS-PAGE loading dye. After heating to 95 °C, the proteins released were analyzed by SDS-PAGE. *I*, input; *P*, DNA bound to the beads after the first pulldown in low salt; *HS*, DNA bound to the beads after washing the beads in high salt.

circular and linear DNA is captured by anti-HA antibodies bound to beads in the presence of 500 mM NaCl and washed with 750 mM NaCl. Persistence of the circular, but not linear, DNA after the high salt wash argues that the DNA had been bound topologically by the protein. Our initial result appeared to support a similar conclusion (Fig. 6E, panel i). Consistent with previous studies, the presence of MukEF and ATP decreased MukB DNA binding (9, 14). However, we were concerned that this result could be artifactual, dependent on the antibodies bound to the beads providing the topological seal that prevented dissociation of the circular DNA. This proved to be the case. If the high salt was added immediately before binding of MukB by the antibodies, no DNA was retained, neither circular nor linear (Fig. 6E, panel ii). Finally, we used an additional assay. Here, we bound tailed, nicked circular DNA to a magnetic bead, then bound MukB to the DNA in the presence or absence of MukEF and ATP, pulled down the beads to isolate the bound DNA, washed the beads with 500 mM NaCl, and then pulled them down again. After high salt washing, no MukB remained bound to the DNA (Fig. 6F). We conclude that MukB does not trap DNA topologically as proposed for the eukaryotic SMC proteins, lending support to our model for MukB binding. This is consistent with the observation that MukB can stably bind linear DNA (9).

MukB K761E R765E, Mutated in the Hinge Domain, is Defective in DNA Condensation

Our model and the SFM data suggest that the MukB hinge domain should bind DNA. The structure of the hinge domains of the *B. subtilis* SMC protein and MukB is somewhat similar, although the *B. subtilis* structure showed a clear

electropositive patch that was demonstrated to bind DNA (28). We examined the surface charge distribution of the MukB hinge and found several small electropositive patches. Charge reversal amino acid substitutions were made, and one, MukB K761E/R765E (named MukB KE,RE), which is located at the top of the hinge (Fig. 7A), distinct from the amino acid residues that interact with ParC (29), was found to have a DNA condensation defect as assayed by both SFM (Fig. 3B) and gel electrophoresis (Fig. 7B). MukB KE,RE was about 50% as effective as the wild type in both assays. Although the mutant protein had a tendency to aggregate at high concentrations, both the FMcx and SMcx were observed.

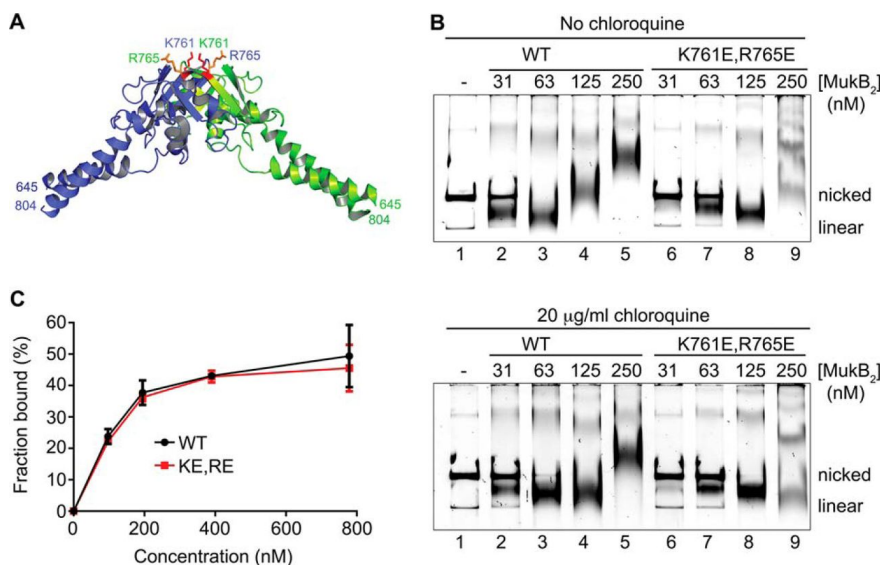


Fig 7. MukB K761E/R765E variant protein can form both the FMcx and SMcx. **A**, schematic of the structure of the hinge domain of MukB (23) with the Lys-761 and Arg-765 amino acids shown in stick representation. **B**, MukB KE,RE can form both the FMcx and SMcx protein-DNA complexes. Either wild type or MukB KE,RE was incubated with the nicked DNA for 5 min at 37 °C, and the products were analyzed by electrophoresis through agarose gels that either did or did not contain 20 µg/ml chloroquine in the gel and running buffer. **C**, wild type and MukB KE,RE binding to a 5'-end-labeled 50-bp duplex DNA is identical. Nitrocellulose filter-binding assays were performed as described under "Experimental procedures."

The defect in DNA condensation by MukB KE,RE could be a result of defective DNA binding; however, DNA-binding activity, as measured in a nitrocellulose filter-binding assay with a duplex 50-mer oligonucleotide, was identical for both the wild-type and variant proteins (Fig. 7C). This assay presumably measures the DNA-binding activity of the head domains but not the hinge domain. Binding of the hinge domain to DNA as measured by SFM was also identical to that of the wild-type protein (Fig. 3). Nevertheless, a DNA-binding defect was apparent for the variant protein as measured by the extent of supercoiling imparted to the nicked DNA after it was sealed by DNA ligase (Fig. 8, A and B, compare lanes 5 and 6 to lanes 9 and 10). Similarly, although MukB KE,RE could form loops that could be supercoiled by DNA gyrase, it was less efficient compared with the wild type (Fig. 8, C and D).

Thus, the MukB KE,RE was deficient in protecting supercoils and forming loops in the DNA. One possible explanation for these observations was that the mutations might cause an alteration of the disposition of the hinge and/or coiled-coil regions that affected the ability of the protein to condense DNA. Long-range transmission of the effect of mutations in the hinge region of the *B. subtilis* SMC protein on the ATPase activity of the protein have been reported (30). To investigate this issue, we purified the isolated hinge domain and examined its activities.

Isolated Hinge Domain Interferes with MukB-mediated DNA Condensation

Neither gel-shift assays (23, 24), nitrocellulose filter DNA-binding assays (data not shown), nor DNA pulldown assays

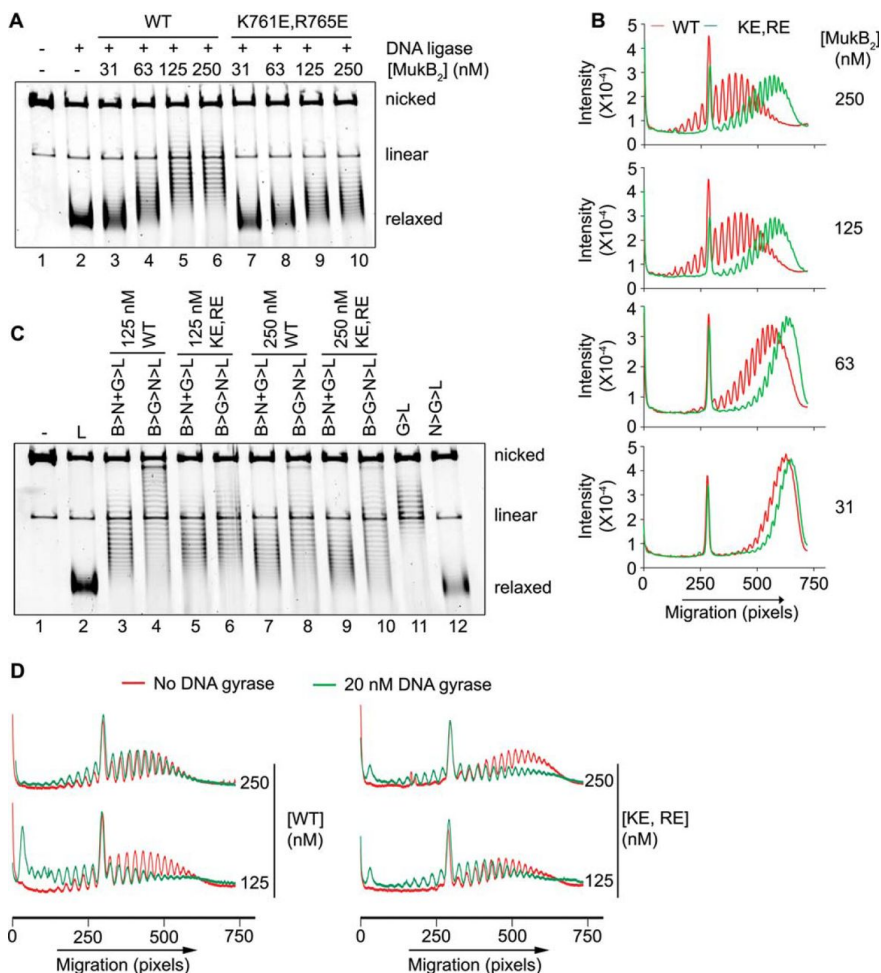


Fig 8. MukB KE,RE variant is defective in the induction of negative supercoils and stabilizing loops in the DNA. **A**, either wild type or MukB KE,RE was incubated with the nicked DNA substrate, NAD, and *E. coli* DNA ligase for 30 min at 37 °C. The products were deproteinized and then analyzed by electrophoresis through an agarose gel containing 10 µg/ml chloroquine in both the gel and the running buffer. **B**, densitometric lane traces of the gels shown in **A**. **C**, MukB KE,RE is defective in stabilizing loops in the DNA. The indicated concentrations of either wild type or MukB KE,RE were incubated with the nicked DNA substrate, DNA gyrase, novobiocin, and DNA ligase as described in the legend to Fig. 2A. The order of addition of components is outlined at the top of the gel. **B**, MukB; **N**, novobiocin; **G**, DNA gyrase; **L**, DNA ligase. Electrophoresis was with 10 µg/ml chloroquine in both the gel and the running buffer. **D**, densitometric lane traces of the gel shown in **C**.

To probe this issue further, we performed DNA cross-linking experiments with the wild-type and mutant hinge domain fragments and a 5'-32P-end-labeled 100-mer DNA duplex. Wild-type hinge fragment cross-linked to the DNA, and this cross-linking was sensitive to prior treatment of the hinge fragment with SDS (Fig. 9C, compare lanes 2 and 3), similar to the cross-linking observed with the full-length wild-type protein (Fig. 9C, compare lanes 6 and 7). However, the KE,RE hinge fragment exhibited little cross-linking to the DNA (Fig. 9C, lanes 4 and 5). Thus, the hinge domain fragment clearly binds DNA.

Fig 10. Hinge domain fragments interfere with MukB-mediated loop formation in the DNA. *A, left panel*, MukB was incubated with the nicked DNA substrate for 5 min at 37 °C. Either wild-type or KE,RE hinge domain fragment was then added, and the incubation continued for an additional 30 min. Reaction products were then analyzed by agarose gel electrophoresis. *Middle panel*, densitometric lane traces of *lanes 1–4* of the gel shown in *A*. *Right panel*, densitometric lane traces of *lanes 1* and *5–7* of the gel shown in *A*. *B, left panel*, as in *A*, except that the concentration of wild-type hinge domain fragment in the reaction was varied. *Right panel*, densitometric lane traces of the gel shown in the *left panel*. *C*, hinge domain fragment does not interfere with MukB-induced supercoiling in the slow-moving protein-DNA complex. MukB and wild-type hinge domain fragment were incubated with the nicked DNA substrate for 5 min at 37 °C. NAD and *E. coli* DNA ligase were then added, and the products were analyzed as in the legend to Fig. 1*B*.

On the one hand, the assay presented in Fig. 9 apparently confirmed the SFM data suggesting that the hinge region bound DNA; however, on the other hand, the absence of supercoiling by the KE,RE hinge is inconsistent with its apparent binding as assayed by SFM. It is possible that DNA binding of the KE,RE hinge is stabilized in the full-length variant protein; however, given that the defects of MukB KE,RE-less negative supercoiling imparted to the DNA—and the KE,RE hinge fragment correspond, we conclude that the defect in supercoiling contributes to the observed defect in DNA condensation by the variant protein.

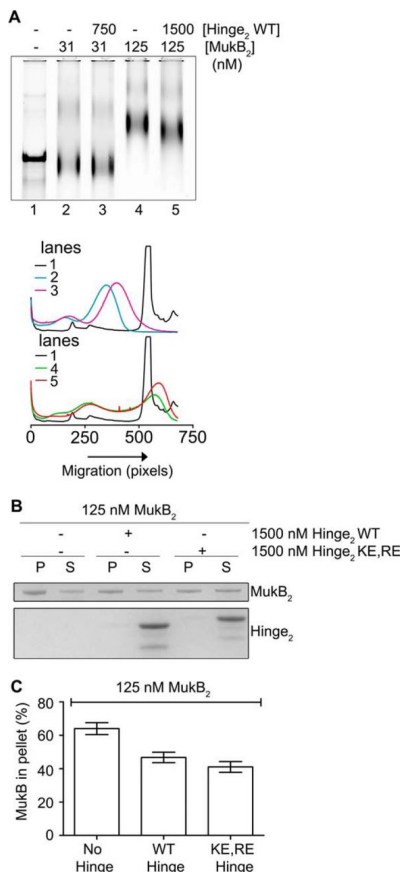


Fig 11. Hinge domain fragment destabilizes bound MukB. A, tailed, nicked DNA substrate is equivalent to the nicked DNA substrate. *Top panel*, MukB in the presence or absence of the wild-type hinge domain fragment was incubated with the biotinylated, tailed, nicked DNA substrate for 5 min at 37 °C and then analyzed by agarose gel electrophoresis. *Bottom panel*, densitometric tracings of the lanes of the gel shown in the *top panel*. B and C, hinge domain fragment destabilizes bound MukB. MukB was incubated on a rotator with tailed, nicked DNA substrate that had been bound to magnetic beads for 5 min at 37 °C. Either wild-type or KE,RE hinge domain fragment was then added, and the incubation continued for 30 min. The beads were then pulled down on a magnet, and the supernatant was removed. The protein present in the pellet (P) and supernatant (S) fractions were then assayed by SDS-PAGE. B, example of the SDS gel. C, distribution of MukB in the pellet and supernatant. The mean and standard deviation is shown for three independent experiments.

To assess whether interfering with hinge–hinge interactions also affected DNA condensation, we added wild-type and KE,RE hinge fragment to the DNA condensation gel assay. The presence of the hinge domains had little effect on the FMcx (Fig. 10A, lanes 2–4); however, they increased the mobility of the SMcx (Fig. 10A, lanes 5–7). This effect was confirmed by the results of a titration of the hinge domain, showing that at lower concentrations of the hinge domain than those used in Fig. 10A, intermediate effects on the mobility of the SMcx were observed (Fig. 10B). The hinge domains had no effect on the mobility of the DNA (Fig. 10A, compare lanes 8 and 9 to lane 1), so the observed increase in the mobility of the SMcx was not a result of additional supercoiling of the DNA. This was confirmed by examining the extent of supercoiling induced when the nicked DNA was sealed in the presence of MukB either in the presence or absence of the wild-type hinge fragment (Fig. 10C). Instead, this observation suggested that the presence of the hinge fragment was destabilizing MukB binding to the DNA and that destabilization did not require hinge fragment DNA binding.

To examine this question, we used a version of the nicked DNA substrate with a 5′-biotinylated single-stranded DNA tail so that protein-DNA complexes could be isolated using streptavidin-coated magnetic beads. We confirmed that this DNA behaved similarly to the nicked DNA substrate, forming both the FMcx and SMcx, and that hinge fragment affected the mobility of the SMcx (Fig. 11A).

Pulldown assays showed that either the wild-type or KE,RE hinge fragment reduced the amount of MukB in the SMcx by one-

quarter and one-third, respectively (Fig. 11, B and C). Similar amounts of MukB were destabilized from the FMcx (data not shown). We conclude that the hinge fragment was disrupting interactions between MukB dimers in the protein-DNA complexes that led to their destabilization. This effect is likely independent of the ability of the hinge domain to bind DNA. Because the mobility of the slow-moving complex is dominated by loop formation, this destabilization results in an observable change in mobility.

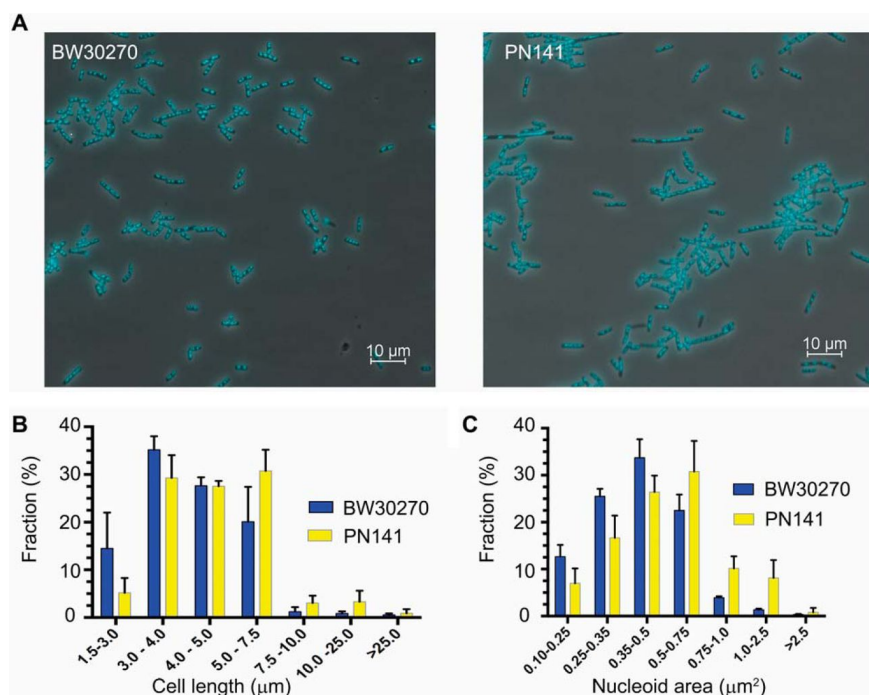


Fig 12. *mukBK761E/R765E* mutant cells are longer, and their nucleoids are larger than wild type. **A**, fields of DAPI-stained wild-type (BW30270) and mutant PN141 (BW30270*mukBK761E/R765E*) *E. coli* cells. **B**, the mutant strain is somewhat longer than the wild type. Cell length distribution of BW30270 and PN141 grown to early log phase in LB medium. A total of 1641 and 1717 cells of the wild type and PN141, respectively, were measured from three independent experiments. **C**, mutant strain has a somewhat larger average nucleoid area than the wild type. A total of 1630 and 1610 nucleoids in wild-type and PN141 cells, respectively, was measured from three independent experiments. The very diffuse nucleoids present in the filamented cells of PN141 were excluded from the latter measurement.

Wild-type and KE,RE hinge fragments consistently had different mobilities in SDS-PAGE (Fig. 11C). Proteolysis followed by mass spectrometry showed identical coverage of peptides on both the C and N termini of the protein fragments, indicating that the difference in mobility was not because of proteolytic clipping. Thus the K761E/R765E mutations may also affect the flexibility of the hinge. This change in hinge flexibility could also be responsible for some of the defects of the variant MukB KE,RE protein.

The Hinge Mutations Disrupt DNA Condensation and MukB Localization *in vivo*

To determine the effect of the hinge mutations on MukB-mediated condensation of DNA *in vivo*, we engineered an *E. coli* strain where the wild-type mukB allele had been replaced with the mukBK761E/R765E allele. PN141 cells carrying the mukBK761E/R765E allele and wild-type BW2070 were grown in LB medium, stained with DAPI (without fixation), and examined (Fig. 12). Mutant cells displayed slightly increased cell length (Fig. 12B) and nucleoid area (Fig. 12C), with median cell length increasing by 13% (from 4.0 to 4.53 μm) and median nucleoid area increasing by 25% (from 0.4 to 0.5 μm^2), suggesting a delay in cell division and nucleoid decondensation. However, because these differences were small and not statistically significant, we examined nucleoid compaction in the mutant strain by velocity sedimentation of isolated nucleoids directly (Fig. 13). Spermidine nucleoids were isolated from wild-type and PN141 cells grown to mid-log phase in MOPS medium and sedimented through 10–40% sucrose gradients. Nucleoids isolated from PN141 sedimented significantly more slowly than those isolated from the wild-type strain (Fig. 13).

Thus, we conclude that the KE,RE MukB hinge mutation causes nucleoid decondensation.

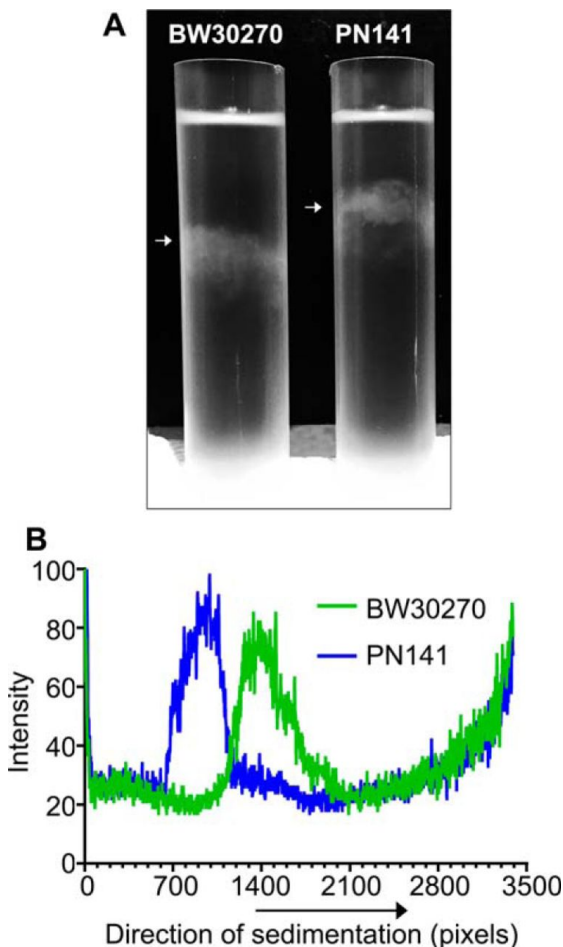


Fig 13. Nucleoids in the *mukBK761E/R765E* mutant strain are decondensed. Wild-type (BW30270) and hinge mutant (PN141, BW30270*mukBK761E/R765E*) *E. coli* cells were grown to mid-log phase in MOPS medium. Cells were harvested, and spermidine nucleoids were prepared and sedimented through sucrose gradients as described under “Experimental procedures.” A, photograph of typical sucrose gradients showing the positions of the sedimented nucleoids (white arrows). B, densitometric trace of OD₂₆₀ absorbing material in each gradient.

Discussion

Chromosomes must be condensed efficiently and in an orderly fashion so as to both fit in their specialized compartments and to allow for proper chromosome replication, gene expression, and segregation. We applied several novel biochemical assays supported by molecular imaging and studies *in vivo* to examine how the bacterial SMC-like protein MukB condenses DNA.

MukB forms and stabilizes two distinct types of topological domains to condense DNA

Of the two types of MukB-DNA complexes formed, the FMcx was dominated by the formation of negative supercoils, as shown by decreased mobility in chloroquine-containing gels. Surprisingly, although the decrease in mobility of the SMcx compared with that of the FMcx could be accounted for by the addition of more MukB, we found that under these conditions relaxed DNA loops were being formed in the DNA that could be supercoiled by DNA gyrase. Thus, MukB was sufficient to establish two distinct types of topological domains in a nominally unconstrained DNA in the absence of either ATP or its accessory proteins.

SFM imaging of these protein-DNA complexes demonstrated progressive condensation of the DNA as MukB concentrations increased, as had been observed previously (14). Moreover, these analyses pointed to the likelihood that the MukB hinge region both bound DNA and seeded aggregation of MukB dimers. Accordingly, we developed a model for MukB DNA condensation that incorporated these features whereby

condensation of the DNA occurs by the protection of negative supercoils by individual MukB dimers locally and the formation of loops in the DNA anchored by MukB dimers that interact via their hinge domains. These loops are themselves topologically isolated and can thus be supercoiled further.

MukB DNA condensation

A current model for how SMC proteins bind DNA invokes topological entrapment of the DNA by the tripartite complex formed by the dimer of the SMC proteins, the kleisin, and additional accessory proteins (1). Yeast cohesin has been shown to become trapped around circular but not linear small and large minichromosomes if all three interfaces in the tripartite complex are cross-linked (32, 33). In an assay where the cohesin was loaded onto DNA, the protein immunoprecipitated, the precipitate washed in high salt, and the DNA contents analyzed, only circular DNA survived the process (27). Similar types of assays have led to similar conclusions about the manner of DNA binding for the yeast condensin (34), the Smc5/6 complex (35), the *B. subtilis* SMC protein (12), and MukB (13).

There are, however, counterarguments to the topological trapping model. In single molecule experiments, Sun et al. (21) measuring compaction of DNA by the yeast cohesin against applied force concluded that DNA condensation proceeded via supercoiling-dependent loop formation that did not require ATP. Eng et al. (36) demonstrated interallelic complementation between two inactive (non-DNA binding) yeast cohesin complexes and argued that cohesion of sister chromatids was elaborated via protein-protein interactions between two distinct cohesin

complexes, one bound to each of the sisters. Recently, in single molecule analysis of the manner by which the *Schizosaccharomyces pombe* cohesin diffused along a DNA molecule, Stigler et al. (37) found that although the protein could be pushed along by motor proteins, it could not bypass nucleosomes, and its diffusion was hindered by bound EcoRI and dCas9. These authors concluded that although the binding of the cohesin to DNA was topological, the central pore of the complex was not as large as predicted, with the protein assuming a more rod-like structure.

When tested directly in two different assays, we found no evidence that MukB, with or without MukF, MukE, or ATP, trapped DNA topologically. Furthermore, significant curvature of the DNA was not required for MukB DNA condensation. Our model for DNA condensation is more in line with the “handcuff” model proposed by Koshland and co-workers (36), where protein–protein interactions between MukB molecules are responsible for tethering distal parts of the chromosome to form stable loops in the DNA. Binding of MukB to the chromosome and the linking of distal loops are likely to elaborate a structure with a central core with loops extending from the core (Fig. 6C). Previous models describing a similar chromosomal arrangement have suggested MukB stabilization of loops in supercoiled DNA (14, 25). We point out, however, that MukB loading to the DNA is likely continuous with DNA replication, during which the two sister chromosomes are not topologically closed and thus cannot be supercoiled. Our model therefore suggests a manner by which overall chromosomal supercoiling can be maintained. Topological isolation of the loops formed by MukB allows supercoiling by DNA gyrase of the nascent

sisters, thus ensuring continuous readout of the information imprinted in a supercoiled chromosome. Our model also is consistent with the order of genetic markers along the long axis of the cell and the occupancy of individual replichores in opposite cell halves (38, 39). Indeed, the latter organizing principle is dependent on active MukB (40). Our model is also consistent with ChIP data that show MukB distributed more or less evenly across the genome (41). In addition, a structure of this type is consistent with proposed nucleoid structures whether it be a constrained ring polymer (42, 43) or not (44, 45).

Role of the MukB hinge domain in DNA condensation

We have shown that the MukB hinge domain does bind DNA, an activity previously unappreciated (23, 24). This activity of the hinge clearly contributes to DNA condensation directly in that condensed MukB-DNA complexes formed with the wild-type protein were supercoiled to a greater extent than those formed with the KE,RE variant. The discovery of the hinge DNA-binding activity brings the action of MukB more in line with that of other SMC proteins where the hinge is known to bind DNA (1).

However, we believe that the most important activity of the hinge domain is in its requirement for DNA loop formation by MukB. Both the wild-type and KE,RE hinge fragment interfered with DNA loop formation by causing the dissociation of MukB. This interference was clearly not because the hinge fragments either competed or interfered with binding of the MukB head domains to DNA. Neither version of the hinge domain was recovered in the pellet of the DNA-binding pulldown assays, and the KE,RE hinge fragment may not bind DNA at all. Thus, the destabilization likely

comes from disruption of hinge–hinge interactions between MukB dimers that themselves are stabilizing to the MukB on the DNA. DNA binding of the hinge region in the full-length MukB KE,RE protein may result in additional stabilizing interactions between the protein and the DNA.

The picture developed thus differs from other models of aggregate formation by SMC proteins that most often show aggregation via head–head domain interactions fueled by binding of ATP. ATP was not required nor did it affect the DNA condensation that we observed, implying that even when head domains are in the open configuration (i.e. not joined together by the binding of ATP), the protein can condense DNA effectively.

DNA condensation, MukEF, and ATP

Our observations that neither ATP nor MukEF were required to observe MukB-mediated condensation of DNA are consistent with previous observations reporting: (i) MukB did not require ATP to bind plasmid DNA and alter its topology (4); (ii) preformed MukBEF complex is defective in binding plasmid DNA (9); and (iii) adding MukEF to a MukB-DNA complex displaced MukB from the DNA (9). As mentioned above, MukE and MukF are required for and are present in MukB foci in vivo (6–8). MukB focus formation in vivo also requires its ATP-binding and hydrolysis activities (8). Strains carrying the MukB hinge mutations exhibited nucleoid decondensation. Thus, our model for MukB DNA condensation is reflective of what occurs in vivo. One possibility is that the effects of MukEF and ATP relate more to turnover of MukB on the DNA than to DNA condensation per se. We report

the effect of the MukB-topoisomerase IV interaction on MukB-mediated DNA condensation in the accompanying article (31).

Materials and methods

DNAs and proteins

The nicked DNA substrate was prepared as described in Bahng et al. (48). Biotinylated, tailed, and nicked DNA substrate was prepared as described in Yeeles and Mariani (46) using the oligonucleotide 5'-TTTTTTTTTTTTCCCCCTCGAGCCCCTTTTTTTTTTTTGGGGCTCGAGGGG GTTTTTGTTTTCCCAGTCACGAC-3' to prime synthesis of the complementary strand of M13mp18 viral DNA by bacteriophage T4 DNA polymerase. MukB was prepared as described previously (47). MukB K761E/R765E was prepared as for the wild type from *E. coli* strain BL21(DE3)pET11d-mukBK761E/R765E.

Hinge domain fragments were prepared as follows. A DNA fragment (synthesized by Gene Script) encoding amino acids 645–804 of MukB and provided cloned into pUC57 was cloned into the NdeI and XhoI restriction sites of the pET28a vector (Novagen) to create the overexpression vectors for both the WT and K761E, R765E variant of the hinge domain. *E. coli* BL21(DE3) Rosetta cells (Novagen) carrying these plasmids were grown in LB medium at 37

°C to an A600 = 0.4. isopropyl 1-thio-β-d-galactopyranoside (0.4 mm) was added and growth continued for 3 h. After harvesting, cells were lysed by sonication in buffer A (50 mm Tris-HCl (pH 8.0), 5% glycerol, 500 mm NaCl, and 2 mm DTT) with 10 mm imidazole-HCl followed by centrifugation at 20,000 × g for 30 min. The supernatant was mixed with Ni-NTA resin (Invitrogen) pre-equilibrated with buffer A containing 10 mm imidazole-HCl. The resin was washed with buffer A containing 75 mm imidazole HCl, and the hinge fragments were eluted with the same buffer containing 250 mm imidazole HCl. Fractions containing the hinge fragments were pooled and dialyzed against buffer B (50 mm Tris-HCl (pH 8.0), 150 mm NaCl, 2.5 mm CaCl₂, 2 mm DTT, and 20 mm imidazole HCl). One unit of thrombin per mg of protein was added followed by incubation for 8–10 h at 4 °C. PMSF was then added to a final concentration of 1 mm, and the NaCl concentration was increased to 500 mm. The sample was mixed with Ni-NTA resin pre-equilibrated with buffer A containing 20 mm imidazole HCl for 1 h, and the flow-through was collected, dialyzed against buffer C (50 mm Tris-HCl (pH 8.0), 5% glycerol, 10 mm DTT, 100 mm NaCl, and 1 mm EDTA), and loaded on to a Q-Sepharose column (GE Healthcare) equilibrated with the same buffer. The column was washed with buffer C, and protein was eluted with a linear gradient of 100–800 mm NaCl in buffer C. Fractions were pooled

based on SDS-PAGE analysis, dialyzed first against buffer C, and then storage buffer (50 mm HEPES-KOH (pH 7.5), 150 mm NaCl, 0.1 mm EDTA, 2 mm DTT. and 40% glycerol). Purified hinge fragments were stored at -80°C .

Agarose gel assays

All biochemical gel assays were repeated at least three times.

DNA condensation

Reaction mixtures (20 μl) containing 50 mm HEPES-KOH (pH 7.5), 20 mm KCl, 10 mm DTT, 0.5 mm $\text{Mg}(\text{OAc})_2$, 7.5% glycerol, 0.34 nm nicked or ScaI -cut (linearized) pCG09, and either wild-type or the KE,RE variant MukB at the concentrations indicated in the figures were incubated for 5 min at 37°C . Reaction products were analyzed by electrophoresis through 0.8% agarose gels at 23 V for 18 h at 4°C using 50 mm Tris-HCl (pH 7.8 at 23°C), 40 mm NaOAc, 1 mm EDTA, and 1.5 mm $\text{Mg}(\text{OAc})_2$ as the electrophoresis buffer. Chloroquine di-phosphate (20 $\mu\text{g}/\text{ml}$) was added to the gel and the electrophoresis buffer as indicated in the figures. Gels were stained with SYBR gold (Invitrogen) and scanned using a Typhoon FLA 9500 scanner. Profiles of gel lanes were taken using Multi Gauge Software (Fujifilm).

Negative supercoil induction

To assay the introduction of negative supercoils, 26 μ m NAD and E. coli DNA ligase (New England Biolabs, 0.4 units) were included in the DNA condensation reaction mixtures that were incubated for 30 min at 37 °C. Reactions were stopped by the addition of EDTA and KCl to final concentrations of 50 and 300 mM, respectively, followed by incubation for 5 min at room temperature. Proteinase K and SDS were then added to 0.2 mg/ml and 0.2%, respectively, and the incubation was continued for 30 min at 37 °C. DNA products were resolved by electrophoresis through 0.8% agarose gels at 30 V for 15 h at room temperature using 50 mM Tris-HCl (pH 7.8 at 23 °C), 40 mM NaOAc, 1 mM EDTA, and 10 μ g/ml chloroquine di-phosphate as the running buffer. Gels were stained and visualized as above.

DNA loop supercoiling by DNA gyrase

In DNA loop supercoiling assays, 2 mM ATP, an additional 2 mM Mg(OAc)₂, and 20 nM DNA gyrase were included in DNA condensation reaction mixtures that were incubated for 30 min at 37 °C. Novobiocin was then added to 10 μ M to inhibit DNA gyrase, and the incubation was continued at room temperature for 5 min. Bacteriophage T4 DNA ligase (New England Biolabs, 0.2 units) was

then added, and the incubation continued for 30 min at 37 °C. The reactions were terminated and analyzed as above for negative supercoil induction.

MukB DNA-binding assays

Nitrocellulose filter binding

As described in Bahng et al. (48), data shown are from three independent experiments.

Magnetic bead pulldown

The biotinylated, tailed, nicked DNA was attached to M280 beads using the Kilobase Binder Kit (Invitrogen). Reaction mixtures were as for the DNA condensation assays except that beads carrying 50 ng of DNA were used and 0.05% Nonidet P-40 was also included. Reactions were incubated for 30 min at 37 °C on a rotator. DNA beads were then pulled down with a magnet. The supernatant was removed, and the beads were resuspended in the same volume of reaction buffer without protein and DNA. SDS sample loading buffer was added, and the reaction mixtures were heated at 95 °C for 5 min. The reaction products were resolved by 4–12% gradient SDS-PAGE (Invitrogen) and visualized by Coomassie staining. Intensity of the bands was quantified using

MultiGauge Software (Fujifilm). Data shown are from three independent experiments.

DNA entrapment assayed by MukB-HA immunoprecipitation

DNA entrapment was assayed by MukB-HA immunoprecipitation as per Murayama and Uhlmann (27). Reaction mixtures (20 μ l) containing 50 mM HEPES-KOH (pH 8.0), 20 mM KCl, 10 mM DTT, 7.5% glycerol, 0.05% Nonidet P-40, 100 ng each of NtBspQ1-nicked pET11b (nicked) and HpaI-cut pET11b (linear) DNA, the indicated concentrations of MukB-HA, MukF, and MukE, and either 0.5 mM Mg(OAc)₂ or 2.5 mM Mg(OAc)₂ and 2 mM ATP were incubated for 5 min at 37 °C. In the experiment shown in Fig. 6E, panel i, anti-HA antibody conjugated to magnetic beads (Pierce) that was sufficient to bind all of the MukB-HA was added; the reaction was mixed quickly by vortex genie for 2 s, and 500 μ l of overnight binding buffer (50 mM HEPES-KOH (pH 8.0), 500 mM NaCl, 7.5% glycerol, 0.05% Nonidet P-40, and either 0.5 mM Mg(OAc)₂ or 2.5 mM Mg(OAc)₂ and 2 mM ATP, as appropriate) was then added, and the incubation continued overnight on a rotator at 4 °C. In the experiment shown in Fig. 6E, panel ii, the order of addition of the overnight binding buffer and the anti-HA antibody beads was reversed. The magnetic beads were then pulled down and washed by resuspension followed by

re-pelleting twice in 400 μ l and once in 200 μ l of 50 mM HEPES-KOH (pH 8.0), 750 mM NaCl, 0.05% Nonidet P-40, and either 0.5 mM Mg(OAc)₂ or 2.5 mM Mg(OAc)₂ and 2 mM ATP as appropriate. The beads were then resuspended once in 200 μ l of 50 mM HEPES-KOH (pH 8.0), 100 mM NaCl, and 0.05% Nonidet P-40, and pelleted. The beads were then resuspended in 15 μ l of 50 mM HEPES-KOH (pH 8.0), 1 mM EDTA, 50 mM NaCl, 1% SDS, and 1 mg/ml proteinase K, and incubated at 50 °C for 30 min. The beads were pelleted, and DNA present in the supernatants was analyzed by electrophoresis through 0.9% agarose gels at 60 V for 150 min at room temperature using 50 mM Tris-HCl (pH 7.8 at 23 °C), 40 mM NaOAc, 1 mM EDTA, as the running buffer. Gels were stained and visualized as above. Shown is a typical gel from three independent experiments.

DNA entrapment assayed by DNA pulldown

Reaction mixtures (20 μ l) as for the DNA condensation assay, but also containing 0.05% Nonidet P-40 and tailed, nicked DNA (50 ng) immobilized on M280 streptavidin beads as above, as well as the indicated concentrations of MukB, MukF, and MukE either in the presence or absence of ATP, were incubated for 5 min at 37 °C. The beads were pulled down and washed twice by resuspension and pulldown in reaction buffer (200 μ l) containing

500 mM NaCl. The final washed pellet was resuspended in 25 μ l of SDS-PAGE sample buffer and heated at 95 °C for 5 min, and the protein present was analyzed by SDS-PAGE. Shown are typical gels from three independent experiments.

DNA binding by cross-linking

Standard MukB DNA-binding reactions containing 1500 nm wild-type MukB hinge domain fragment, KE,RE hinge domain fragment, or full-length MukB (125 nm) and 1 nm 5'-32P duplex 100-mer (5'-GAGTGC GGATCCTGACGACATGCCGTGCTACGACTGACGACATGCGAAGTACTGCAGCCCGGGTACGACGATGGATCGCATGCAGTGGATCCACGCATCA-3') were incubated for 5 min at 37 °C. HCHO/MeOH mix (33–10%) was added to a final concentration of 0.7% HCHO and the incubation continued on ice for 30 min. Tris-HCl (pH 7.5) and SDS-polyacrylamide gel loading buffer were then added to 0.5 M and 1 \times , respectively, and the reaction products were resolved by electrophoresis through a 4–12% gradient BisTris polyacrylamide gel using MES-SDS (Invitrogen) as the electrophoresis buffer. Gels were dried, exposed to a phosphorimager screen, and scanned using a Fujifilm FLA 7000 scanner. Images were processed using ImageQuant TL (GE Healthcare).

Scanning force microscopy

DNA condensation by MukB and MukB K761E/R765E

Either wild type or MukB K761E/R765E at 15, 30, and 60 nm were incubated in standard MukB DNA-binding reactions as above. The samples were then diluted 5-fold with deposition buffer (10 mM HEPES-NaOH (pH 8), 20 mM MgCl₂), and 20 μ l were deposited on freshly cleaved mica. After 1 min of incubation at room temperature, the mica was washed with Milli-Q water and dried with filtered air. Samples were imaged at room temperature and humidity with a Nanoscope IV (Digital Instruments) operating in tapping mode. Type NHC-W silicon tips of the resonance frequency at 310–372 kHz were obtained from Nanosensors (Veeco Instruments, Europe). Images were collected at $2 \times 2 \mu\text{m}$, with a standard resolution of 512 lines and 512 rows, and processed only by flattening to remove background slope. Detailed methods for SFM sample preparation can be found in Ref. 49. The degree of DNA condensation was quantified using the SFMetricsV4c1 analysis tool (50). The length of DNA protruding from MukB multimers was manually traced. As a point of reference, the total length of 50 unambiguous individual circular pCG09 molecules was measured in the same manner, giving an average length of 3430 ± 174 nm. The length of DNA loops

protruding from a MukB multimer complex was summed. Data shown are from three independent experiments.

DNA binding of MukB and MukB K761E/R765E

Either wild type or MukB K761E/R765E at 7.5 and 15 nm were incubated in standard MukB DNA-binding reactions as above. Samples were then diluted and imaged as above. Images were quantified by visual inspection where individual MukB molecules were identified as bound to a DNA strand with the head domain (appearing as one large or two small globular domains), with the hinge domain (appearing as a singular small globular domain), with both head and hinge domains, or as a “looping” mode where the head domain was bound to two DNA regions bringing them together, forming a DNA loop. All DNA-bound MukB molecules in an image were counted and categorized. The frequency of the different forms was expressed as percentage of total molecules counted. Data shown are from three independent experiments.

E. coli strains

BW30270 is an *rph*⁺ derivative of MG1655 (*E. coli* Genetic Stock Center). PN141 (BW30270mukBK761E/R765E) was engineered using CRISPR-Cas9 technology as described (51). A

guide RNA plasmid was constructed by inserting an EcoRI- and SpeI-digested DNA fragment (GeneScript) 5'-GCTCGTCCACTAGTGGAAGGCTTGACCGATTGCCGTTTTAGAGCTAGA AATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGG CACCGAGTCGGTGCTTTTTTTGAATTCTCAGTGCC-3', where the underlined sequence represents the mukB N20 sequence, into similarly digested pTargetF to create pTargetF-mukB2172. An editing template PCR fragment, 5'-TGAGCGAGAGGCGATTGTTGAACGCGATGAAGTGGGCGCGCGCAAAA ACGCCGTCGATGAAGAGATCGAACGTTTAAGCCAGCCTGGCGGCTCTG AAGATCAGCGTCTGAACGCGCTGGCGGAGCGTTTTGGTGGTGTGCTG CTGTCAGAAATTTATGACGACGTTAGCCTGGAAGATGCGCCGTACTTC TCAGCGCTGTATGGCCCGTCACGCCACGCCATCGTGGTGCCAGATCTG TCACAGGTA ACTGAACACCTGGAAGGCTTGACCGATTGCCAGAAGAT CTCTATCTGATCGAAGGAGATCCGCAGTCATTCGATGACAGCGTGTT CAGCGTTGATGAGCTGGAAAAAGCGGTAGTGGTGGAAATCGCCGATGA ACAGTGGCGTTATTCACGTTTCCCGGAAGTGCCGCTGTTTGGTCGTGCT GCGCGTGAAAGCCGTATTGAAAGCCTCCATGCCGAGCGTGAAGTGCTT TCCGAACGCTTCGCCACGCTCTCCTTTGATGTACAGAAA ACTCAGCGTC TGCATCAGGCGTTCAGCCGCTTTATCGGCAGTCATCTGGCGGTTGCGTT TGAGTCTGACCCGGAAGCAGAAATCCGTCAACTGAACAGCCGTCGCGT CGAACTGGAGCGGGCGTTAAGTAATCATGAAAATGATAACCAGCAGC AGCGTATTCAGTTTGAGCAGGCGAAAGAGGG-3', where the bold

nucleotide represents the change in the photospacer adjacent motif, and the underlined nucleotide triplets are the amino acid substitutions generated by overlap PCR using the following primers for mukB: 5'-TGAGCGAGAGGCGATTGTTG-3'; 5'-GATCAGATAGAGATCTTCTGGGCAATCGGTCAAGCCTTCCAGGTGTTCA GTTACCTGTGA; TCACAGGTA ACTGAACACCTGGAAGGCTTGACCGATTGCCCAGAAGATCTCTATCTGATC; CTGTACATCAAAGGAGAGCGTGGC-3'; 5'-GCCACGCTCTCCTTTGATGTACAG-3'; and 5'-CCCTCTTTCGCCTGCTCAAA-3'. BW30270 was transformed with plasmid pCas, supplying Cas9. Competent cells were prepared from BW30270(pCas9) grown in the presence of 1 mM l-arabinose to induce Cas9 expression. These cells were transformed with pTargetF-mukB2172 (100 ng) and editing template DNA (400 ng). After loss of both plasmids, surviving cells were screened by DNA sequencing of mukB. Whole-genome sequencing was then performed to ensure that there were no off-target mutations introduced during the strain construction.

Isolation of spermidine nucleoids

BW30270 and PN141 were grown at 37 °C in MOPS EZ Rich Defined Medium with 2% glucose to an OD600 = 0.5. Cells (25 ml) were pelleted by centrifugation, lysed, and sedimented through

10–40% sucrose gradients as described by Murphy and Zimmerman (52) with the exception that the gradients were centrifuged for 20 min.

Cell growth and fluorescent microscopy

For DAPI staining, strains were grown at 37 °C in LB medium with 0.5% glucose, 20 µg/ml thymine, and 20 µg/ml thiamine to OD₆₀₀ = 0.2. Five milliliters of cell culture were incubated with 1 µg/ml DAPI for 10 min at 37 °C. Cells were pelleted, washed with 5 ml of M9 medium, pelleted, and resuspended in 0.5 ml of M9 medium, and 100 µl of this suspension were spread on freshly-coated poly-l-lysine slides for 2 min at room temperature. The slides were drained and washed with 1× PBS, air-dried, covered with 2 drops fluorescent mounting medium and a coverslip, and observed in a Nikon Eclipse Ti microscope using a 100 × CFI Plan Apo DM 1.4 numerical aperture objective. Images were captured with an Andor Neo sCMOS camera. Cell length and nucleoid area measurements were made using the measurement tools in version 4.4 of Nikon Elements Software.

References

- 1.Uhlmann, F. (2016) SMC complexes: from DNA to chromosomes. *Nat. Rev. Mol. Cell Biol.* 17, 399–412
- 2.Niki, H., Jaffé, A., Imamura, R., Ogura, T., and Hiraga, S. (1991) The new gene *mukB* codes for a 177-kd protein with coiled-coil domains involved in chromosome partitioning of *E. coli*. *EMBO J.* 10, 183–193
- 3.Yamanaka, K., Ogura, T., Niki, H., and Hiraga, S. (1996) Identification of two new genes, *mukE* and *mukF*, involved in chromosome partitioning in *Escherichia coli*. *Mol. Gen. Genet.* 250, 241–251
- 4.Petrushenko, Z. M., Lai, C. H., Rai, R., and Rybenkov, V. V. (2006) DNA reshaping by MukB. Right-handed knotting, left-handed supercoiling. *J. Biol. Chem.* 281, 4606–4615
- 5.Yamazoe, M., Onogi, T., Sunako, Y., Niki, H., Yamanaka, K., Ichimura, T., and Hiraga, S. (1999) Complex formation of MukB, MukE and MukF proteins involved in chromosome partitioning in *Escherichia coli*. *EMBO J.* 18, 5873–5884
- 6.Ohsumi, K., Yamazoe, M., and Hiraga, S. (2001) Different localization of SeqA-bound nascent DNA clusters and MukF-MukE-MukB complex in *Escherichia coli* cells. *Mol. Microbiol.* 40, 835–845
- 7.She, W., Wang, Q., Mordukhova, E. A., and Rybenkov, V. V. (2007) MukEF is required for stable association of MukB with the chromosome. *J. Bacteriol.* 189, 7062–7068
- 8.Badrinarayanan, A., Reyes-Lamothe, R., Uphoff, S., Leake, M. C., and Sherratt, D. J. (2012) In vivo architecture and action of

bacterial structural maintenance of chromosome proteins. *Science* 338, 528–531

9. Petrushenko, Z. M., Lai, C. H., and Rybenkov, V. V. (2006) Antagonistic interactions of kleisins and DNA with bacterial condensin MukB. *J. Biol. Chem.* 281, 34208–34217

10. Chen, N., Zinchenko, A. A., Yoshikawa, Y., Araki, S., Adachi, S., Yamazoe, M., Hiraga, S., and Yoshikawa, K. (2008) ATP-induced shrinkage of DNA with MukB protein and the MukBEF complex of *Escherichia coli*. *J. Bacteriol.* 190, 3731–3737

11. Woo, J. S., Lim, J. H., Shin, H. C., Suh, M. K., Ku, B., Lee, K. H., Joo, K., Robinson, H., Lee, J., Park, S. Y., Ha, N. C., and Oh, B. H. (2009) Structural studies of a bacterial condensin complex reveal ATP-dependent disruption of intersubunit interactions. *Cell* 136, 85–96

12. Wilhelm, L., Burmann, F., Minnen, A., Shin, H. C., Toseland, C. P., Oh, B. H., and Gruber, S. (2015) SMC condensin entraps chromosomal DNA by an ATP hydrolysis dependent loading mechanism in *Bacillus subtilis*. *Elife* 4, e06659

13. Niki, H., and Yano, K. (2016) In vitro topological loading of bacterial condensin MukB on DNA, preferentially single-stranded DNA rather than double-stranded DNA. *Sci. Rep.* 6, 29469

14. Cui, Y., Petrushenko, Z. M., and Rybenkov, V. V. (2008) MukB acts as a macromolecular clamp in DNA condensation. *Nat. Struct. Mol. Biol.* 15, 411–418

15. Michieletto, D., Marenduzzo, D., and Orlandini, E. (2015) Topological patterns in two-dimensional gel electrophoresis of DNA knots. *Proc. Natl. Acad. Sci. U.S.A.* 112, E5471–E5477

16. Lehman, I. R. (1974) DNA ligase: structure, mechanism, and function. *Science* 186, 790–797
17. Liu, L. F., Depew, R. E., and Wang, J. C. (1976) Knotted single-stranded DNA rings: a novel topological isomer of circular single-stranded DNA formed by treatment with *Escherichia coli* omega protein. *J. Mol. Biol.* 106, 439–452
18. Shuman, S., and Prescott, J. (1990) Specific DNA cleavage and binding by vaccinia virus DNA topoisomerase I. *J. Biol. Chem.* 265, 17826–17836
19. Gellert, M., Mizuuchi, K., O'Dea, M. H., and Nash, H. A. (1976) DNA gyrase: an enzyme that introduces superhelical turns into DNA. *Proc. Natl. Acad. Sci. U.S.A.* 73, 3872–3876
20. Gellert, M., O'Dea, M. H., Itoh, T., and Tomizawa, J. (1976) Novobiocin and coumermycin inhibit DNA supercoiling catalyzed by DNA gyrase. *Proc. Natl. Acad. Sci. U.S.A.* 73, 4474–4478
21. Sun, M., Nishino, T., and Marko, J. F. (2013) The SMC1-SMC3 cohesin heterodimer structures DNA through supercoiling-dependent loop formation. *Nucleic Acids Res.* 41, 6149–6160
22. Matoba, K., Yamazoe, M., Mayanagi, K., Morikawa, K., and Hiraga, S. (2005) Comparison of MukB homodimer versus MukBEF complex molecular architectures by electron microscopy reveals a higher-order multimerization. *Biochem. Biophys. Res. Commun.* 333, 694–702
23. Ku, B., Lim, J. H., Shin, H. C., Shin, S. Y., and Oh, B. H. (2010) Crystal structure of the MukB hinge domain with coiled-coil stretches and its functional implications. *Proteins* 78, 1483–1490
24. Li, Y., Schoeffler, A. J., Berger, J. M., and Oakley, M. G. (2010) The crystal structure of the hinge domain of the *Escherichia coli*

structural maintenance of chromosomes protein MukB. *J. Mol. Biol.* 395, 11–19

25. Rybenkov, V. V., Herrera, V., Petrushenko, Z. M., and Zhao, H. (2014) MukBEF, a chromosomal organizer. *J. Mol. Microbiol. Biotechnol.* 24, 371–383

26. Kim, H., and Loparo, J. J. (2016) Multistep assembly of DNA condensation clusters by SMC. *Nat. Commun.* 7, 10200

27. Murayama, Y., and Uhlmann, F. (2014) Biochemical reconstitution of topological DNA binding by the cohesin ring. *Nature* 505, 367–371

28. Hirano, M., and Hirano, T. (2002) Hinge-mediated dimerization of SMC protein is essential for its dynamic interaction with DNA. *EMBO J.* 21, 5733–5744

29. Hayama, R., Bahng, S., Karasu, M. E., and Mariani, K. J. (2013) The MukB-ParC interaction affects the intramolecular, not intermolecular, activities of topoisomerase IV. *J. Biol. Chem.* 288, 7653–7661

30. Soh, Y. M., Bürmann, F., Shin, H. C., Oda, T., Jin, K. S., Toseland, C. P., Kim, C., Lee, H., Kim, S. J., Kong, M. S., Durand-Diebold, M. L., Kim, Y. G., Kim, H. M., Lee, N. K., Sato, M., et al. (2015) Molecular basis for SMC rod formation and its dissolution upon DNA binding. *Mol. Cell* 57, 290–303

31. Kumar, R., Nurse, P., Bahng, S., Lee, C. M., and Mariani, K. J. (2017) The MukB-topoisomerase IV interaction is required for proper chromosome compaction. *J. Biol. Chem.* 292, 16921–16932

32. Haering, C. H., Farcas, A. M., Arumugam, P., Metson, J., and Nasmyth, K. (2008) The cohesin ring concatenates sister DNA molecules. *Nature* 454, 297–301

- 33.Farcas, A. M., Uluocak, P., Helmhart, W., and Nasmyth, K. (2011) Cohesin's concatenation of sister DNAs maintains their intertwining. *Mol. Cell* 44, 97–107
- 34.Cuylen, S., Metz, J., and Haering, C. H. (2011) Condensin structures chromosomal DNA through topological links. *Nat. Struct. Mol. Biol.* 18, 894–901
- 35.Kanno, T., Berta, D. G., and Sjögren, C. (2015) The Smc5/6 complex is an ATP-dependent intermolecular DNA linker. *Cell Rep.* 12, 1471–1482
- 36.Eng, T., Guacci, V., and Koshland, D. (2015) Interallelic complementation provides functional evidence for cohesin-cohesin interactions on DNA. *Mol. Biol. Cell* 26, 4224–4235
- 37.Stigler, J., Çamdere, G. Ö., Koshland, D. E., and Greene, E. C. (2016) Single-molecule imaging reveals a collapsed conformational state for DNA-bound cohesin. *Cell Rep.* 15, 988–998
- 38.Wang, X., Liu, X., Possoz, C., and Sherratt, D. J. (2006) The two *Escherichia coli* chromosome arms locate to separate cell halves. *Genes Dev.* 20, 1727–1731
- 39.Nielsen, H. J., Ottesen, J. R., Youngren, B., Austin, S. J., and Hansen, F. G. (2006) The *Escherichia coli* chromosome is organized with the left and right chromosome arms in separate cell halves. *Mol. Microbiol.* 62, 331–338
- 40.Danilova, O., Reyes-Lamothe, R., Pinskaya, M., Sherratt, D., and Possoz, C. (2007) MukB colocalizes with the *oriC* region and is required for organization of the two *Escherichia coli* chromosome arms into separate cell halves. *Mol. Microbiol.* 65, 1485–1492
- 41.Nolivos, S., Upton, A. L., Badrinarayanan, A., Müller, J., Zawadzka, K., Wiktor, J., Gill, A., Arciszewska, L., Nicolas, E., and

- Sherratt, D. (2016) MatP regulates the coordinated action of topoisomerase IV and MukBEF in chromosome segregation. *Nat. Commun.* 7, 10466
42. Jun, S., and Mulder, B. (2006) Entropy-driven spatial organization of highly confined polymers: lessons for the bacterial chromosome. *Proc. Natl. Acad. Sci. U.S.A.* 103, 12388–12393
43. Youngren, B., Nielsen, H. J., Jun, S., and Austin, S. (2014) The multifork *Escherichia coli* chromosome is a self-duplicating and self-segregating thermodynamic ring polymer. *Genes Dev.* 28, 71–84
44. Fisher, J. K., Bourniquel, A., Witz, G., Weiner, B., Prentiss, M., and Kleckner, N. (2013) Four-dimensional imaging of *E. coli* nucleoid organization and dynamics in living cells. *Cell* 153, 882–895
45. Hadizadeh Yazdi, N., Guet, C. C., Johnson, R. C., and Marko, J. F. (2012) Variation of the folding and dynamics of the *Escherichia coli* chromosome with growth conditions. *Mol. Microbiol.* 86, 1318–1333
46. Yeeles, J. T., and Marians, K. J. (2011) The *Escherichia coli* replisome is inherently DNA damage tolerant. *Science* 334, 235–238
47. Hayama, R., and Marians, K. J. (2010) Physical and functional interaction between the condensin MukB and the decatenase topoisomerase IV in *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* 107, 18826–18831
48. Bahng, S., Hayama, R., and Marians, K. J. (2016) MukB-mediated catenation of DNA is ATP and MukEF independent. *J. Biol. Chem.* 291, 23999–24008

- 49.Ristic, D., Sanchez, H., and Wyman, C. (2011) Sample preparation for SFM imaging of DNA, proteins, and DNA-protein complexes. *Methods Mol. Biol.* 783, 213–231
- 50.Sánchez, H., and Wyman, C. (2015) SFMetrics: an analysis tool for scanning force microscopy images of biomolecules. *BMC Bioinformatics* 16, 27
- 51.Jiang, Y., Chen, B., Duan, C., Sun, B., Yang, J., and Yang, S. (2015) Multigene editing in the *Escherichia coli* genome via the CRISPR-Cas9 system. *Appl. Environ. Microbiol.* 81, 2506–2514
- 52.Murphy, L. D., and Zimmerman, S. B. (1997) Isolation and characterization of spermidine nucleoids from *Escherichia coli*. *J. Struct. Biol.* 119, 321–335

Chapter 3

Investigating the structure-to-function relationships of MRE11-RAD50 complex

M. Grosbart¹, H. Sanchez¹, S.A. van Rossum¹, R. Kanaar,^{1,2} and C. Wyman^{1,2}.

Department of ¹Molecular Genetics and ²Radiation Oncology Erasmus University Medical Center, P. O. Box 2040, 3000CA Rotterdam, The Netherlands

Abstract

The MRE11-RAD50 complex (MR) is an evolutionarily conserved protein crucial for a myriad of biological processes, including DNA repair. Both MRE11 and RAD50 are proteins essential for cell survival. The functions of the MR complex, even within one pathway, are extremely versatile. In homologous recombination MR takes part in DNA damage signaling, DNA processing and DNA tethering. Understanding how MR can perform functionally different activities was one of the goals of our research. The MR complex consists of large proteins with notably flexible parts limiting the structural tools that can be applied. We employ Scanning Force Microscopy (SFM) to investigate the architecture and conformation changes of the MR complex upon interaction with ligands. In order to determine the impact of ATP binding and hydrolysis on the structure of MR we created protein variants with point mutations in the RAD50 ATPase interface. Our results show, that both ATP binding and hydrolysis are important for protein oligomerization and formation of DNA bound multimers needed for tethering. Images of the MR complex with different DNA substrates revealed a new conformation state that possibly corresponds to ssDNA processing and damage signaling functions. Overall, our data points to the importance of dramatic conformation changes of MR complex and contributes to understanding how exchange of the various protein-protein interfaces modulate function of this protein in DNA break repair.

Introduction

The MRE11-RAD50 complex functions at several steps in DNA break repair including: break recognition, DNA end-processing, DNA tethering and cell-cycle checkpoint signaling (Wyman and Kanaar, 2006). These various functions require the specific arrangement and dynamic rearrangement of component parts in the MRE11-RAD50 complex. Based on its structure, the RAD50 component is a member of the structural maintenance of chromosomes (SMC) family. SMC proteins share distinct features, most notably they consist of a pair of proteins (homo or hetero dimers) that are long (30-50 nm) intra-molecular coiled-coil with a bipartite ATPase active sites at one end (Kinoshita et al., 2009). The termini of the coiled coils are distinct dimer interfaces or otherwise participate in interactions with partners. This creates the possibility for either an open V shape if one of the interfaces is joined or a closed protein loop if both interfaces are joined. The ABC type ATPase active sites are created at the interface of monomers at one end of the elongated complex. The MRE11 component is a homodimer, each subunit bound to one RAD50 at the domain including the ATPase site. We refer to the combined MRE11 dimer connected to two RAD50 ATPase domains as the globular domain of the complex. MRE11 is a weak nuclease on single-stranded (ss) as well as double-stranded (ds) DNA, with both endo- and exonuclease activity. These enzymatic activities function in DNA processing, as well as DNA end recognition and ATM signaling (Paull and Deshpande, 2014; Williams et al., 2010)

Functional changes in MR architecture are logically linked to dynamics at the interfaces of the components. X-ray

crystallography studies of isolated parts of the complex have defined interfaces and revealed dramatic structural variations associated with DNA or ATP binding (Lammens et al., 2011; Lim et al., 2011; Mockel et al., 2012; Rojowska et al., 2014). The MRE11 dimer associates with two RAD50s at one end of the elongated structure. DNA binds across the MRE11 dimer and some DNA arrangements are substrates for its nuclease activity (Paull and Gellert, 1999; Trujillo and Sung, 2001). The presence of flexible linkers or structurally undefined amino acid stretches in RAD50 between the interface with MRE11 and the ATPase domains allow for the bipartite ATPase active sites to be separated up to 10nm (Lammens et al., 2011). Binding ATP creates a compact RAD50 dimer oriented next to MRE11 obstructing the MRE11 DNA binding and nuclease active sites. ATP binding by RAD50 creates another DNA binding site, located between the protruding coiled coils across the RAD50 globular domain dimer (Lammens et al., 2011). ATP binding causes a 30 degree rotation of the RAD50 globular domains relative to each other with corresponding effects on orientation of the protruding coiled coils (Williams et al., 2011). The RAD50 coiled-coil apex, some 50 nm away from the globular ATPase active site, comprises another dimer interface, the Zn hook (Hopfner et al., 2002). Two Zn hooks can coordinate a zinc ion to create an interface joining the ends of the coiled coils either within one complex (intramolecular) or between different MR complexes (intermolecular). Flexibility in the coiled-coils leads to variable, constrained and possibly controlled, orientation of the Zn hooks. Intramolecular Zn hook association is dynamic (de Jager et al., 2001; Moreno-Herrero et al., 2005). Intramolecular joined Zn hooks occur with bent coiled coils creating a closed protein loop

(de Jager et al., 2002; Moreno-Herrero et al., 2005; van Noort et al., 2003). A recent X-ray crystallographic study of the eukaryotic RAD50 reveals an additional interface along the coiled-coils independent of the Zn hook binding. The coiled-coils in this conformation are arranged in an anti-parallel (end-on-end) fashion with stretches of amino acids forming interfaces that stabilize the MR arms in a rod-like structure (Park et al., 2017).

We apply SFM imaging to reveal how the various interactions interfaces of MR connect and manifest as conformational arrangements in the complete complex. For instance this has shown that DNA bound at the globular domains results in parallel oriented coiled-coils (Moreno-Herrero et al., 2005). This exemplifies how binding of ligands at one site of a large molecular complex dramatically influence structure and function of a distant but connected site of the same complex. The DNA bound orientation of RAD50 Zn hooks was proposed to favor an intermolecular interaction that would serve to tether DNA ends in need or coordinated repair. This tethering was assumed to occur via ZN hooks joined end to end as observed in crystallographic structures of an archaeal RAD50 (He et al., 2012). Considering the newly identified eukaryotic Zn hook orientation and coiled-coil interface (Park et al., 2017), the DNA bound parallel arrangement could either promote or prevent intra complex interactions depending on rotational orientation of the coiled-coils. This is because the coiled coil orientation is constrained by connection to the globular domains at the base, an aspect that is lacking when only parts of the proteins are isolated for analysis. SFM imaging identified MR oligomerization upon DNA binding and suggests this is needed for DNA tethering (de Jager et al.,

2001). Several potential interactions, including DNA binding (to either MRE11 or RAD50), ATP binding, and coiled-coil orientation simultaneously contribute to biologically relevant oligomerization and function in DNA tethering. We created RAD50 variants with altered ATPase active sites to analyze the importance of ATP binding and hydrolysis for MR complex architecture, oligomerization and DNA tethering.

The effect of ATP on the architecture of MR complexes formed with WT RAD50 or ATPase deficient RAD50 revealed that: (i) ATP stimulated MR oligomerization, (ii) MR oligomerization was stabilized by both ATP binding and hydrolysis (iii). Higher resolution imaging of MR bound to short ssDNA revealed a new conformation with an elongated bi-partite globular domain. Together these results show that the MR complex is dynamic and flexible, that biological function is modulated by ligand binding and communicated through architectural arrangements connecting distant parts of the complex. We propose links between the observed conformational transitions and specific biological functions of MR.

Results

Scanning Force Microscopy (SFM) allows observing full-length human MR complex in conditions compatible with biochemical activity. This tool has been successfully used to characterize architectural arrangements and changes in the MR complex with nanometer resolution (Allison et al., 2010). Specifically the orientation and arrangement of the RAD50 coiled coils can easily be determined by simple inspection of the SFM images. To analyze the consequence of interactions with ATP we

introduced point mutations into ATP-binding pocket of RAD50. We generated a MR^{S1202R} in which RAD50 is defective in ATP binding and thus also hydrolysis and MR^{E229Q} in which RAD50 can bind ATP but has a reduced hydrolyzing capacity. These RAD50 variants were co expressed and purified as MRE11-RAD50 complexes in the same manner as the wild type protein (Lee and Paull, 2006). The effect of these mutation in RAD50 on ATP binding and hydrolysis by MR complex were as expected. Both MR^{S1202R} and MR^{E229Q} had DNA binding activity similar to MR wild type as assayed by EMSA using dsDNA oligo substrates (data not shown).

Influence of ATP binding and hydrolysis on MR architecture

Human MR protein has a distinct architecture consisting of a globular domain, including an MRE11 dimer and two RAD50 ATPase domains, from which the extended coiled coils of RAD50 protrude. Several aspects of the coiled coils are distinguishable in SFM image. The two RAD50 coiled coils of an MR complex can either be straight and parallel or bend with their apex Zinc hooks not joined or joined to each other, resulting in a closed protein loop. MR complexes often form oligomers interacting via their globular domains with protruding coiled-coils (de Jager et al., 2001; Moreno-Herrero et al., 2005). These different forms reflect the variable association of the multiple protein-protein interfaces in the complex. With SFM imaging we can observe and categorize all MR complexes in a population. The distribution of forms represents a snapshot view of the variety of conformation in solution. As observed previously the most prevalent forms observed consist of MRE11-RAD50 globular domain with protruding RAD50 coiled coils, either separated or linked Zn hook

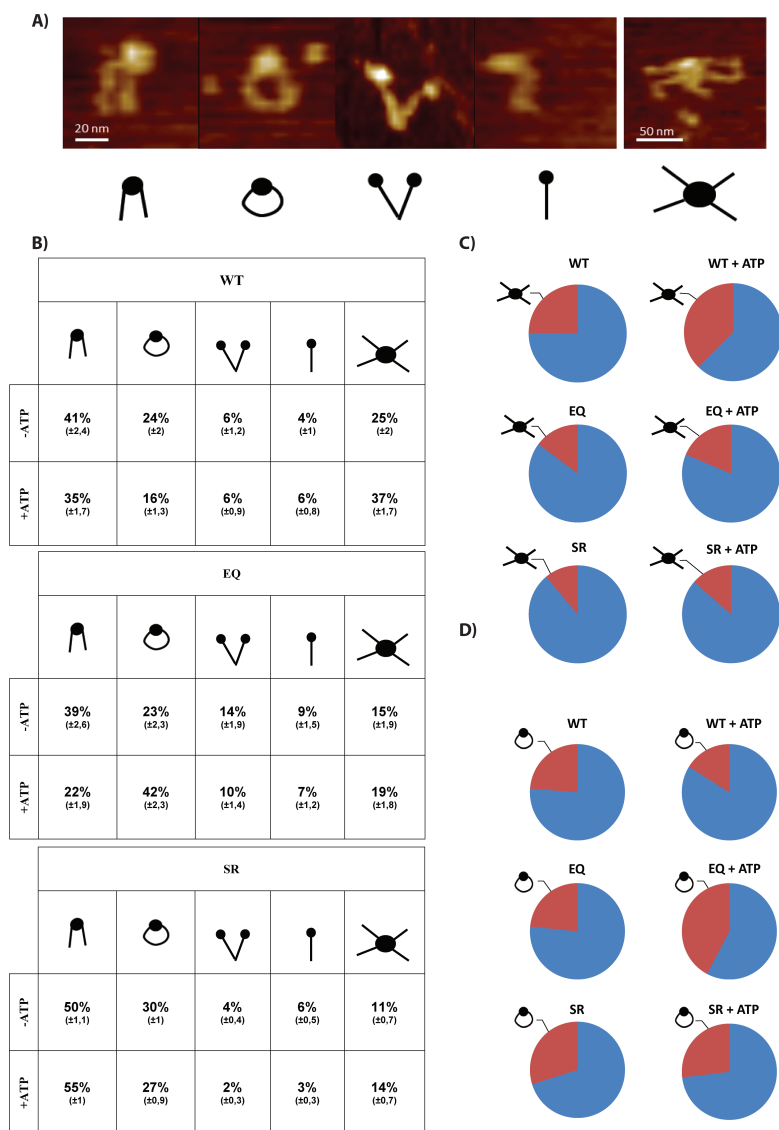


Fig 1. The effect of ATP on MR architecture. A) MR conformation states as seen in SFM. B) Frequency of individual forms in MR WT, MR^{E229Q} and MR^{S1202R} samples, with and without ATP C) Impact of ATP on oligomerization of MR WT, MR^{E229Q} and MR^{S1202R} D) Impact of ATP on coiled-coil apexes interaction of MR WT, MR^{E229Q} and MR^{S1202R}. All data was collected with $n \geq 1000$, error is expressed as standard error of proportions.

apexes (de Jager et al., 2001) (Fig 1A, left panel) Oligomers, 2-4 MR (based on the number of protruding coiled-coils) interacting via the globular domains, are also typically observed (de Jager et al., 2001) (Fig 1A, right panel). Additionally we observed a small fraction of molecules referred to as “half-complexes” or “monomers”, consisting of one MRE11 and one RAD50, and another fraction as complexes with joined coiled-coil apexes but not connected at the globular domain referred to as “reversed dimers”. The distribution of these forms in WT MR is consistent among preps and over years of analysis (Kinoshita et al., 2009). MR complexes with ATPase deficient RAD50 exhibited fewer oligomers than WT, 11 and 15% for MR^{E229Q} and MR^{S1202R}, respectively, vs 25% for the WT (Fig 1B). This indicates the importance of the ATPase active site interface for conformations that oligomerize even in the absence of ATP. The MR^{E229Q} had more “reverse dimers” (14% vs 6% or 4%, Fig 1B) and monomers (9% vs 4 or 6%, Fig 1B) compared to WT and MR^{S1202R} respectively, indicating weaker interface or more often disassociated state of the globular domain interfaces, likely those between RAD50 and MRE11. Conversely MR^{S1202R} had more globular domain joined forms with either open and joined Zn hooks (50%+30% = 80% vs 65% or 62%, Fig 1B) compared to WT or MR^{E229Q} respectively. This suggests a stronger interface or more often associated state of the globular domains.

Based on X-ray crystallography the ATPase domains of RAD50 can have vary different arrangements with respect to each other and to MRE11 in ATP bound and ATP free forms (Hopfner et al., 2000; Lammens et al., 2011; Lim et al., 2011). However the crystalized forms do not include full-length coiled-coils that are

expected both to respond to and to constrain movement of the globular domains. We used SFM imaging to determine the influence of ATP binding on the arrangement of full-length complete MR complex. Addition of ATP to WT MR resulted in a shift in the distribution of different forms, notably oligomers were more prevalent (37% vs 25%, Fig 1B, C).

Addition of ATP did not increase oligomerization for either the MR^{E229Q} or MR^{S1202R} MR complexes (fig 1B, C), even though the MR^{E229Q} binds ATP. This implies that the conformational change in MR promoting oligomerization includes ATP binding and subsequent hydrolysis, possibly requiring accompanying dynamic transitions. ATP addition to MR^{E229Q} increased the dimeric forms with joined Zn hooks (42% vs 23%, fig 1B, D) while this changed little or even decreased for the wild type complex. This indicates that the ATP bound form of RAD50, presumably the arrangement defined in various crystallography experiments with the compact RAD50 dimer associated with the MRE11 dimer (Seifert et al., 2016), orients the coiled coils so that intramolecular joining of the Zn hooks is promoted. The transition upon ATP hydrolysis (active in the wild type) apparently re-orientes the coiled coils to disfavor intracomplex Zn hook interaction (fig 1B, D). As expected the MR^{S1202R} variant does not change with added ATP (Fig 1 B, C and D).

Influence of ATP binding and hydrolysis on DNA tethering

DNA tethering is a well-described function of MR, in which the protein keeps damaged DNA ends in close proximity for further processing and repair. We defined this activity based on

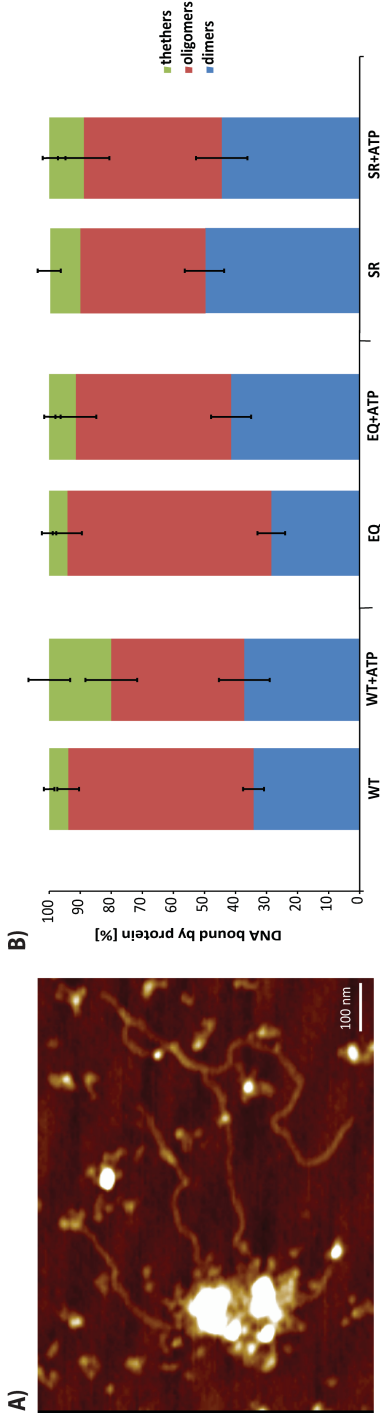


Fig 2. The effect of ATP on tethering capacity of MR A) Tethering of long, ds DNA strands by MR multimers B) The effect of ATP on DNA tethering by MR WT, MR^{E229Q} and MR^{S1202R}. Tethering capacity is expressed as number of DNA strands bound by tethering complexes (green bars) vs DNA bound by MR oligomers and dimers (red and blue bars, respectively). All data was collected with $n \geq 200$, error is expressed as standard error of proportions.

SFM images where we observe DNA molecules brought together by MRE11-RAD50 protein complexes apparently interacting via their coiled-coil apexes. Two distinct aspects of MR change upon DNA binding: the coiled-coils in individual complexes become parallel and oligomeric clusters accumulate, particularly on long dsDNA (Fig 2A)

We have suggested that tethering requires both the formation of large oligomeric protein complexes on DNA (de Jager et al., 2001; de Jager et al., 2002) and parallel orientation of the coiled coils. The parallel coiled-coils presumably constrain orientation of the apexes and prevent intramolecular interaction (Moreno-Herrero et al., 2005). Here we analyzed the effect of ATP binding and hydrolysis on oligomerization and tethering by DNA bound MR. We compared the frequency of DNA-tethering induced by MR, MR^{S1202R} and MR^{E229Q}. As was proposed before (de Jager et al., 2002) ATP increased the amount of tethered DNA by WT MR (Fig 2B). In contrast, there was no increase in DNA tethering by either MR^{S1202R} or MR^{E229Q} mutant after incubation with ATP (Fig 2B). The MR^{E229Q} complex did form oligomers on DNA similar to wild type. However, there was no ATP stimulation of DNA tethering for MR^{E229Q}. This observation is consistent with the effect of ATP stabilizing intramolecular coiled-coil interaction of MR^{E229Q} (Fig 1B, D) as this precludes the inter-complex ZN hook interactions involved in tethering (Moreno-Herrero et al., 2005). These results indicate that both ATP binding and hydrolysis are needed to stimulate DNA tethering.

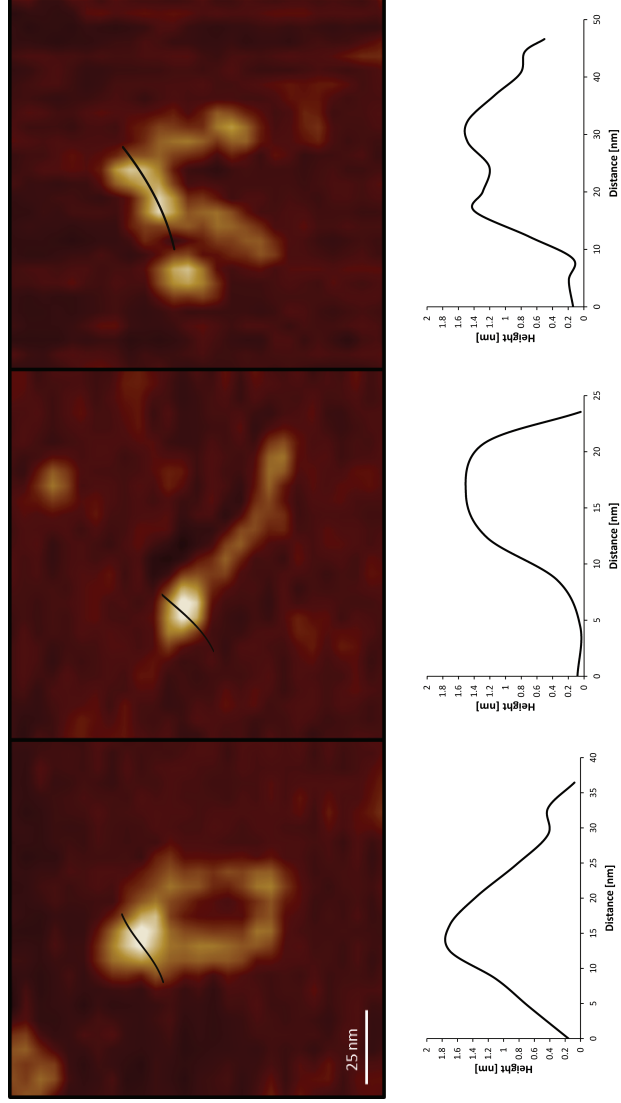


Fig 3. Examples and characteristics of MR dimers' structural arrangements corresponding to binding short DNA substrates. Upper panel: left, the most prevalent form without DNA present with compact globular domain and disengaged coiled-coils; center, the most prevalent form after dsDNA binding, with parallel coiled-coils and compact globular domain; right, the most prevalent form after ssDNA binding, with elongated globular domain and disengaged coiled-coils. The strikethrough indicates how the height profile of the globular domain is obtained. Lower panel: Height profiles of the globular domains of the molecules presented above

Influence of DNA binding on MR architecture

It is difficult to analyze the influence of ATP and DNA on the arrangement of individual MR complexes because both of these ligands induce formation of oligomers in which individual complexes cannot be unambiguously identified in SFM images. To circumvent this complication we used short DNA molecules (90 nt or bp), that cannot accumulate oligomers to observe individual MR complexes where the conformation of the coiled-coils is taken as evidence of DNA binding (Moreno-Herrero et al., 2005). Here also we observed that in the absence of DNA only 7.5% (n=121, Fig 4) of the MR complexes had parallel coiled-coils (figure Fig 3, upper panel). When MR was allowed to bind short dsDNA the number of molecules with parallel coiled-coils tripled, reaching 23%(n=249, Fig 4).

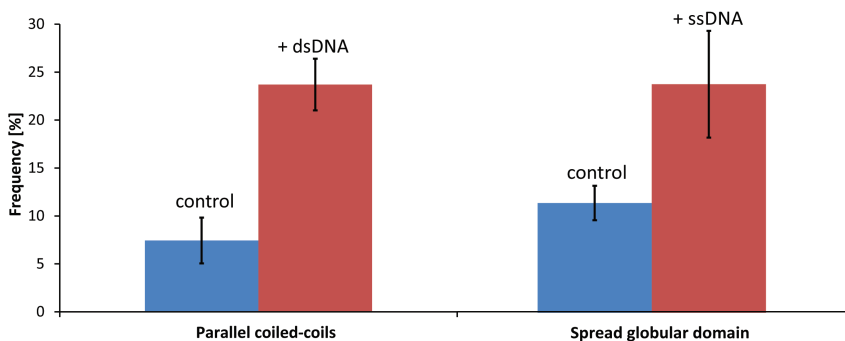


Fig 4. The impact of short DNA fragments on the arrangement of MR. On the left, the frequency of molecules with parallel coiled-coils in the absence and presence of dsDNA (blue and red bars, respectively). On the right, the frequency of molecules with elongated globular domain in the absence and presence of ssDNA (blue and red bars, respectively). All data was collected with $n \geq 100$, error is expressed as standard error of proportions

The ATPase mutants of MR also assume the parallel coiled-coil conformation when bound to DNA. The MRE11-RAD50 complexes were incubated with short dsDNA, in conditions identical to EMSA assays previously used to show MR-DNA binding (data not shown), and analyzed by SFM imaging. Both WT and mutant proteins assumed the anticipated parallel coiled-coil orientation when incubated with ds 90-mer (data not shown).

A strikingly different conformational change was observed after incubating MR with 90 nt ssDNA. A quarter of MR complexes assumed a distinctive, elongated shape of the globular domain, with apparently separated lobes (Fig 3, upper panel, Fig 4). The width of the globular domain cross-section was on average (34 nm $n=233$ $sd=4\text{ nm}$) 10 nm longer than that of MR bound to ds DNA or in the absence of DNA (Fig 3, lower panel). Intriguingly, the coiled coils of the complexes were almost always disengaged in these cases. This conformation of MR is also observed in the presence of an NBS1 peptide (Kim et al., 2017), which stabilizes the MRE11 dimer and promotes MRE11 DNA binding. This conformation of MRE11 with ssDNA bound across the dimer interface would require the RAD50 ATPase domains to be disengaged, not occluding the DNA binding face of MRE11. This would explain the elongated globular domain appearance we observe which could be the same as the forms crystalized in the absence of ATP with two RAD50 ATPase domains separated on either side of an MRE11 dimer (Lammens et al., 2011). The change

observed as a consequence of ssDNA binding is striking and will require follow-up experiments to characterize further.

Discussion

The MR complex performs multiple functions during the initial stages of homologous recombination DNA repair (Wyman and Kanaar, 2006). The array of tasks that the complex performs *in vitro* are mechanistically distinct: DNA processing (Deshpande et al., 2016a; Paull and Gellert, 1999), DNA damage signaling (Deshpande et al., 2016b) and structural function to organize broken DNA (de Jager et al., 2001; Moreno-Herrero et al., 2005). But, as much as we know about *what* the complex does, there are many open questions into *how* it performs these many functions. Based on our work, we propose that MR owes this functional versatility to inherent structural dynamics linked by conformation switches and interface exchanges triggered by interaction with its ligands.

Over the years, crystallography has defined a number of conformations of the globular region of MR. Notably, this has identified dramatic changes within the globular domain upon binding ATP (Lammens et al., 2011) or various DNA substrates (Rojowska et al., 2014; Seifert et al., 2016). The complete complex cannot be crystalized, due to the length and flexibility of the RAD50 coiled coils. Although lower in resolution, SFM imaging can easily be applied to reveal the architectural arrangements of the complete MR complex. As a single-molecule method, SFM can

also categorize and classify the distribution of different conformations (de Jager et al., 2001; Kinoshita et al., 2009).

SFM images of MR reveal a multitude of conformations, regarding both the globular domain and coiled coils (Fig 1A). The variability in complex arrangement indicate a dynamic and flexible character of the complex, assuming different forms by associations and disassociations of its various interfaces. The most prevalent form of MR alone is RAD50 and MRE11 associated closely within the globular domain and the apexes of the RAD50 coiled coils disengaged (Fig 1B). Even in the absence of ligands, an abundance of other forms is observed, which points to possible rapid transition between short-lived conformations.

Allosteric regulation, where ligand binding causes a conformational, is a crucial process governing protein function,. The MR complex binds various ligands. ATP binding triggers rotation of the globular head components, causing the RAD50 molecules to interact tightly to form a functional ATPase interface (Hopfner et al., 2000; Lim et al., 2011). Although not visible in crystallographic reconstructions, this rotation and rearrangement has to have an impact on the arrangement of the coiled-coils. Addition of ATP increases oligomerization of MR via the globular domains (de Jager et al., 2002). The impact of ATP on MR, however, could be attributed to binding of the nucleotide or to subsequent hydrolysis. To test the impact of ATP binding and hydrolysis separately, we created enzymatically defective variants of the protein: MR^{S1202R} (non-binding, non-hydrolyzing) and MR^{E229Q} (binding, reduced hydrolysis). We confirmed increased

oligomerization of MR when exposed to ATP while MR^{E229Q} failed to exhibit any stimulation of oligomer formation. In the presence of the nucleotide the non-hydrolyzing variant assumed the conformation with both globular domain and coiled coil apexes tightly interacting (Fig 1B, C and D). This structural arrangement likely precludes some functions and its prevalence when ATP is bound but not hydrolyzed was intriguing from two reasons. First, it became obvious that even though separated by 50 nm distance, the orientation of the globular head components have a clear impact on the tips of the coiled coils. More importantly, it was previously believed that slow hydrolysis of ATP serves to cycle between open and closed RAD50 Walker domains (Lafrance-Vanasse et al., 2015). Our results strongly suggest that ATP processing involves at least two steps, binding of ATP and subsequent hydrolysis, essential for exposing interfaces responsible for oligomerization. If ATP hydrolysis is hindered the complexes appear to be locked in an intermittent, seemingly non-functional, state with locked coiled coils apexes.

When imaged bound to kilobasepair long dsDNA, MR appears as multimeric clusters capturing multiple DNA molecules. This DNA tethering presumably functions to hold DNA strands in close proximity for subsequent repair (de Jager et al., 2001; Moreno-Herrero et al., 2005). Tethering increases after addition of ATP (de Jager et al., 2002; Deshpande et al., 2016b), which seems quite intuitive, since ATP stimulates MR oligomerization in general. Here, MR, MR^{S1202R} and MR^{E229Q} all exhibited some tethering activity, which was only stimulated by ATP in the case of WT protein. MR^{E229Q} failure to show increased DNA tethering even

though it can bind ATP corresponds logically to the lack of ATP-stimulated oligomerization discussed above. Taking into consideration that all protein variants bind DNA with comparable efficiency, we propose, that both ATP binding and hydrolysis is needed to prompt conformation change exposing interfaces involved in oligomerization and subsequent building up of tethering complexes.

DNA binding has complex allosteric effects on MR. Crystals structures of the MRE11 homologues from other species have identified binding sites for both single- and double stranded DNA and revealed allosteric alterations within the MRE11 dimer (Lammens et al., 2011; Lim et al., 2011; Williams et al., 2011; Williams et al., 2008). Our SFM results show that DNA binding has consequences for the complex as a whole, promoting parallel orientation of the coiled coils (Moreno-Herrero et al., 2005) and promoting formation of tethering complexes (de Jager et al., 2001; de Jager et al., 2002). We have now identified a new, DNA induced conformation change: a separation of globular domain components upon ssDNA binding (Fig 3 and 4). The change is dramatic with the globular domain wider by 10 nm on average, but not disengaged (as can be seen in a small subpopulation of the complexes which seem “broken in half”, Fig 1A).

Similarly, in the absence of ATP, the globular domain of a bacterial MR homolog is elongated, as studied by crystallography and SAXS (Lammens et al., 2011). In contrast, archaeal Me11-Rad50 complex is dynamic and freely samples between “open” (Rad50 Walker domains disengaged) and “closed” (Rad50 Walker

domains tightly engaged) conformation states (Rambo and Tainer, 2011; Williams et al., 2011). It is clear from our data, showing the variety and distribution of arrangements, of both MRE11-RAD50 globular domains and RAD50 coiled-coils, that human MR has dynamic propensities. The complex with an elongated globular domain, which we observe upon ssDNA binding, corresponds with the proposed role for that MR in the ATP-unbound state (with RAD50 Walker domains disengaged) in DNA processing and damage signaling (Lafrance-Vanasse et al., 2015; Lee et al., 2013).

The elongated globular domain associated with binding ssDNA places this form at early stages of repair: break recognition, ssDNA processing, and/or ATM signaling. Binding of NBS1 to MRE11 produces a similar effect on MR architecture (Kim et al., 2017), linking this conformation to signaling functions. In the absence of NBS1 the ATM activation does not take place, but ATM is still recruited to the damage site by MRE11 (Lee and Paull, 2005). It is possible, that this conformation (observed both with ssDNA and NBS1 (Kim et al., 2017)) exposes interfaces for interaction with the ATM kinase. Presumably other MR complexes bound to ATP would trigger a conformation change preferable for dsDNA interaction and tethering.

Materials and Methods

Protein Expression and Purification

MR preparations were produced by infection of Sf9 cells in liquid culture with baculovirus expressing human histidine tagged RAD50 and untagged MRE11 at an MOI of ~10. The purification procedure

was based on a method described previously (Lee and Paull, 2006). Briefly, infected cells were collected and flash frozen in liquid nitrogen. They were subsequently thawed on ice and resuspended in 100 ml of ice-cold buffer (50 mM Tris-HCl pH 8.0; 10% glycerol, 500 mM NaCl, 10 mM β -mercaptoethanol) and two tablets of complete, EDTA-free protease inhibitor cocktail tablets (Roche). The cells were disrupted by 20 strokes of a type B pestle in a Dounce homogenizer and centrifuged for 1 hour at 80 000 *g* at 4°C. The soluble fraction was loaded onto 5 ml of Ni-NTA agarose beads (Qiagen) equilibrated with buffer (50 mM Tris-HCl pH 8.0; 10% glycerol, 500 mM NaCl, 1 mM DTT). The column was washed with 10 column volumes of Buffer A (50 mM Tris-HCl pH 8.0; 10% glycerol, 300 mM NaCl, 1 mM DTT) containing 10 mM and then 40 mM imidazole. Proteins were eluted by Buffer A with 130 mM and then 200 mM imidazole. MR fractions were pooled, diluted by two to lower the salt concentration to 150 mM NaCl and loaded onto a 1ml HitrapQ column (GE Healthcare) equilibrated with Buffer A (50 mM Tris-HCl pH 8.0; 10% glycerol, 100 mM NaCl, 1mM DTT). After washing the column with 5 column volumes, the proteins were eluted with a salt block of 600 mM NaCl. Relevant fractions were pooled and loaded on a Superose 6 size exclusion column (GE Healthcare) equilibrated with Buffer A (200 mM NaCl). Relevant fractions were pooled and loaded onto a MonoQ column equilibrated with Buffer A. The column was washed with 5 column volumes of Buffer A and the proteins were eluted with a salt gradient from 100 mM to 600 mM NaCl. Protein concentration was determined, fractions were aliquoted, flash frozen in liquid nitrogen and stored at -80°C. Immuno-blotting of RAD50 and MRE11 was performed with antibodies MS-RAD10 (GeneTex), and

PC388 (Oncogene) respectively, on nitrocellulose transfer membrane (Whatman) using standard immuno-blotting techniques. Protein concentrations were determined by the Bradford method, using BSA as a standard.

DNA Substrates

PuC19 plasmid was obtained through transformation of DH5 α competent *E.coli* cell. Purified DNA was linearized with *SmaI*. The isolated product was checked via electrophoresis and SFM inspection.

ssSK3 and ssSK3-Cy5 were obtained from Integrated DNA Technologies. dsSK3-Cy5 was produced by annealing with untagged complimentary SK3 oligo in buffer containing 50 mM Tris-HCl, pH 8.0; 100 mM NaCl; 10 mM MgCl₂ and 1 mM DTT placed in water bath at 95 C for 10 min which was then left to slowly cool overnight.

DNA₆₆ for gel retardation assays was obtained by annealing 5' - AF532 - AGA AAC TGG GCA TGT GGA GAC AGA GAA GAC TCT TGG GTT TCT GAT AGG CAC TGA CTC TCT CTG CCT-3' and its complement 5' - AGG CAG AGA GAG TCA GTG CCT ATC AGA AAC CCA AGA GTC TTC TCT GTC TCC ACA TGC CCA GTT TCT - 3' (both synthesized by Eurogenetec) in buffer containing 50 mM Tris-HCl, pH 8.0; 100 mM NaCl; 10 mM MgCl₂ and 1 mM DTT placed in water bath at 95°C for 10 min which was then left to slowly cool overnight.

ATPase and ATP binding assays

ATPase assays were performed using radiolabelled [γ - 32 P]ATP as described (de Jager et al., 2002) 60 nM MR was mixed with radiolabelled [γ - 32 P]ATP supplemented with non-labelled ATP to make a final concentration of 1.5 μ M in 10 μ l binding buffer (50 mM Tris-HCl, pH 8.0, 5% glycerol, 2% PEG-6000, 100 mM NaCl₂, 5 mM MgCl₂ and 1 mM DTT). The reaction was incubated for 90 minutes at 37°C. Samples were taken every 15-30 minutes and stopped by adding EDTA to a final concentration of 125 mM. Samples were analyzed by thin layer chromatography (Merck TLC plates) run in 0.7M K₂HPO₄/0.4M Boric acid, exposed to phosphor image plate (Molecular Diagnostics) for 1 hour and imaged using an image analyzer (Typhoon 9200).

For ATP binding assay, 60 nM MR was mixed with radiolabelled [γ - 32 P]ATP complimented with non-labelled ATP up to final concentration of 1.5 μ M in 10 μ l in 10 μ l binding buffer (50 mM Tris-HCl pH 8.0, 5% glycerol, 2% PEG-6000, 100 mM NaCl₂, 5 mM MgCl₂ and 1mM DTT). Samples were incubated 5 minutes at room temperature and loaded onto a 4% non-denaturing polyacrylamide gel, which was then run at 150 V for 4 hours. Subsequently, the gel was dried onto a nitrocellulose membrane (Whatman). The membrane was exposed for 1 hour to phosphor image plate (Molecular Diagnostics) and visualized using an image analyzer (Typhoon 9200).

DNA binding assays

1 nM of ds or ss-66-AF532 was mixed with MR at the indicated concentrations and incubated for 5 min at 25°C in 20 μ l of binding buffer (25 mM Tris-HCl, pH 7.5, 5% glycerol, 2% PEG-6000, 100

mM NaCl₂, 5 mM MgCl₂ and 1 mM DTT). The reaction products were separated on a 4.8% non-denaturing polyacrylamide gel run in 0.5 x TB buffer. The labelled DNA was visualized by scanning the gel with at 532 nm using an image analyzer (Typhoon 9200). The emission signal was recorded using a 555BP 20 nm filter.

SFM analysis

For structural analysis of MR complex, protein preparations were diluted with storage buffer (300 mM NaCl, 10% glycerol, 50 mM Tris-HCl-HCl, pH8) to a concentration of 8 nM in final volume of 20 µl. Samples of were diluted 5 times with deposition buffer (20 mM MgCl₂, 10 mM Hepes, pH8) and deposited on freshly cleaved mica. After 1 min incubation at room temperature the mica was washed with distilled water and dried with filtered air. Samples were imaged at room temperature and humidity by tapping mode SFM using Nanoscope IV (Digital Instruments). Silicon tips were obtained from Digital Instruments (Nanoprobes). Images were collected as 2.5 x 2.5 µm scans and processed only by flattening to remove background slope. To determine the influence of nucleotide binding on MR architecture 1 mM ATP was added to the protein samples along with the storage buffer to a final volume of 20 µl. The results were quantified by counting MR monomers (seen as globular domain with one 50 nm coiled coil), dimers (globular domain with two 50 nm coiled coils) and oligomers (large globular bodies with more than two 50 nm coiled coils). Each population is presented as the percentage of total amount of protein molecules present on the images. Error is expressed as Standard Error for Proportions:

$$\sigma_p = \sqrt{p * (100 - p) / N}$$

For analysis of MR bound to the short DNA fragments, 1 nM protein was incubated with 100 nM ss/ds SK3 oligo in storage buffer for 10 min at 25°C in the total volume of 20 µl. The whole reaction was then deposited on freshly cleave mica. After 1 min incubation at room temperature the mica was washed with distilled water and dried with filtered air. Images were collected as 2.5 x 2.5 µm scans and processed only by flattening to remove background slope. The results were quantified as described above. To characterize and quantify the ss/dsDNA-MR complexes, the cross-section of globular domain of each MR dimer was measured using SFMetrics_V4c1 software tools (Sanchez and Wyman, 2015).

To characterize interactions of MR with long DNA fragments, linearized PUC19 was incubated with protein as described above. The molar ratios of protein vs DNA (fragments) were 15:36 nM, the concentration of was ATP 1 mM, if present. The samples were diluted 100x in deposition buffer, deposited and imaged as described above. The results were quantified by counting single MR molecules (dimers or monomers) bound to DNA, oligomers bound to DNA ends and oligomers bound internally to the DNA strand. Furthermore, the amount of tethered DNA molecules was quantified by counting the DNA strands coming out of the tether. The DNA was measured using SFMetrics Vc1 software tools (Sanchez and Wyman, 2015). Fragments having the ≥0.5 length of linearized pUC19 were counted as 1 whereas shorter pieces were counted as 0.5 DNA strand. All the results are presented as the

percentage of total molecules of DNA present on the image, with Standard Error of Proportions.

References

- Allison, D.P., Mortensen, N.P., Sullivan, C.J., and Doktycz, M.J. (2010). Atomic force microscopy of biological samples. *Wiley Interdiscip Rev Nanomed Nanobiotechnol* 2, 618-634.
- Binnig, G., Quate, C.F., and Gerber, C. (1986). Atomic force microscope. *Phys Rev Lett* 56, 930-933.
- Carroni, M., and Saibil, H.R. (2016). Cryo electron microscopy to determine the structure of macromolecular complexes. *Methods* 95, 78-85.
- Crick, F. (1970). Central dogma of molecular biology. *Nature* 227, 561-563.
- Crick, F.H., Barnett, L., Brenner, S., and Watts-Tobin, R.J. (1961). General nature of the genetic code for proteins. *Nature* 192, 1227-1232.
- Davies, A.A., Masson, J.Y., McIlwraith, M.J., Stasiak, A.Z., Stasiak, A., Venkitaraman, A.R., and West, S.C. (2001). Role of BRCA2 in control of the RAD51 recombination and DNA repair protein. *Mol Cell* 7, 273-282.
- de Jager, M., van Noort, J., van Gent, D.C., Dekker, C., Kanaar, R., and Wyman, C. (2001). Human Rad50/Mre11 is a flexible complex that can tether DNA ends. *Mol Cell* 8, 1129-1135.
- de Jager, M., Wyman, C., van Gent, D.C., and Kanaar, R. (2002). DNA end-binding specificity of human Rad50/Mre11 is influenced by ATP. *Nucleic Acids Res* 30, 4425-4431.
- Deshpande, R.A., Lee, J.H., Arora, S., and Paull, T.T. (2016a). Nbs1 Converts the Human Mre11/Rad50 Nuclease Complex into an Endo/Exonuclease Machine Specific for Protein-DNA Adducts. *Mol Cell* 64, 593-606.

Deshpande, R.A., Williams, G.J., Limbo, O., Williams, R.S., Kuhnlein, J., Lee, J.H., Classen, S., Guenther, G., Russell, P., Tainer, J.A., *et al.* (2016b). ATP-driven Rad50 conformations regulate DNA tethering, end resection, and ATM checkpoint signaling. *EMBO J* 35, 791.

Doublie, S., Sawaya, M.R., and Ellenberger, T. (1999). An open and closed case for all polymerases. *Structure* 7, R31-35.

Doublie, S., Tabor, S., Long, A.M., Richardson, C.C., and Ellenberger, T. (1998). Crystal structure of a bacteriophage T7 DNA replication complex at 2.2 Å resolution. *Nature* 391, 251-258.

Dunker, A.K., Lawson, J.D., Brown, C.J., Williams, R.M., Romero, P., Oh, J.S., Oldfield, C.J., Campen, A.M., Ratliff, C.M., Hipps, K.W., *et al.* (2001). Intrinsically disordered protein. *J Mol Graph Model* 19, 26-59.

Ellgaard, L., McCaul, N., Chatsisvili, A., and Braakman, I. (2016). Co- and Post-Translational Protein Folding in the ER. *Traffic* 17, 615-638.

Fritsch, M., Leary, C.M., Furlow, J.D., Ahrens, H., Schuh, T.J., Mueller, G.C., and Gorski, J. (1992). A ligand-induced conformational change in the estrogen receptor is localized in the steroid binding domain. *Biochemistry* 31, 5303-5311.

Gerstein, M., Lesk, A.M., and Chothia, C. (1994). Structural mechanisms for domain movements in proteins. *Biochemistry* 33, 6739-6749.

Giglia-Mari, G., Zotter, A., and Vermeulen, W. (2011). DNA damage response. *Cold Spring Harb Perspect Biol* 3, a000745.

Grant, B.J., Gorfe, A.A., and McCammon, J.A. (2010). Large conformational changes in proteins: signaling and other functions. *Curr Opin Struct Biol* 20, 142-147.

Green, C.M., and Almouzni, G. (2002). When repair meets chromatin. First in series on chromatin dynamics. *EMBO Rep* 3, 28-33.

Greenfield, N.J. (2006). Using circular dichroism spectra to estimate protein secondary structure. *Nat Protoc* 1, 2876-2890.

Gregersen, N., Bross, P., Vang, S., and Christensen, J.H. (2006). Protein misfolding and human disease. *Annu Rev Genomics Hum Genet* 7, 103-124.

Groothuizen, F.S., Winkler, I., Cristovao, M., Fish, A., Winterwerp, H.H., Reumer, A., Marx, A.D., Hermans, N., Nicholls, R.A., Murshudov, G.N., *et al.* (2015). MutS/MutL crystal structure reveals that the MutS sliding clamp loads MutL onto DNA. *Elife* 4, e06744.

He, J., Shi, L.Z., Truong, L.N., Lu, C.S., Razavian, N., Li, Y., Negrete, A., Shiloach, J., Berns, M.W., and Wu, X. (2012). Rad50 zinc hook is important for the Mre11 complex to bind chromosomal DNA double-stranded breaks and initiate various DNA damage responses. *J Biol Chem* 287, 31747-31756.

Hoeijmakers, J.H. (2009). DNA damage, aging, and cancer. *N Engl J Med* 361, 1475-1485.

Hopfner, K.P., Craig, L., Moncalian, G., Zinkel, R.A., Usui, T., Owen, B.A., Karcher, A., Henderson, B., Bodmer, J.L., McMurray, C.T., *et al.* (2002). The Rad50 zinc-hook is a structure joining Mre11 complexes in DNA recombination and repair. *Nature* 418, 562-566.

Hopfner, K.P., Karcher, A., Shin, D.S., Craig, L., Arthur, L.M., Carney, J.P., and Tainer, J.A. (2000). Structural biology of Rad50 ATPase: ATP-driven conformational control in DNA double-strand break repair and the ABC-ATPase superfamily. *Cell* 101, 789-800.

Kanaar, R., Hoeijmakers, J.H., and van Gent, D.C. (1998). Molecular mechanisms of DNA double strand break repair. *Trends Cell Biol* 8, 483-489.

Kikhney, A.G., and Svergun, D.I. (2015). A practical guide to small angle X-ray scattering (SAXS) of flexible and intrinsically disordered proteins. *FEBS Lett* 589, 2570-2577.

- Kim, J.H., Grosbart, M., Anand, R., Wyman, C., Cejka, P., and Petrini, J.H. (2017). The Mre11-Nbs1 Interface Is Essential for Viability and Tumor Suppression. *Cell Rep* *18*, 496-507.
- Kinoshita, E., van der Linden, E., Sanchez, H., and Wyman, C. (2009). RAD50, an SMC family member with multiple roles in DNA break repair: how does ATP affect function? *Chromosome Res* *17*, 277-288.
- Lafrance-Vanasse, J., Williams, G.J., and Tainer, J.A. (2015). Envisioning the dynamics and flexibility of Mre11-Rad50-Nbs1 complex to decipher its roles in DNA replication and repair. *Prog Biophys Mol Biol* *117*, 182-193.
- Lammens, K., Bemeleit, D.J., Mockel, C., Clausing, E., Schele, A., Hartung, S., Schiller, C.B., Lucas, M., Angermuller, C., Soding, J., *et al.* (2011). The Mre11:Rad50 structure shows an ATP-dependent molecular clamp in DNA double-strand break repair. *Cell* *145*, 54-66.
- Lee, J.H., Mand, M.R., Deshpande, R.A., Kinoshita, E., Yang, S.H., Wyman, C., and Paull, T.T. (2013). Ataxia telangiectasia-mutated (ATM) kinase activity is regulated by ATP-driven conformational changes in the Mre11/Rad50/Nbs1 (MRN) complex. *J Biol Chem* *288*, 12840-12851.
- Lee, J.H., and Paull, T.T. (2005). ATM activation by DNA double-strand breaks through the Mre11-Rad50-Nbs1 complex. *Science* *308*, 551-554.
- Lee, J.H., and Paull, T.T. (2006). Purification and biochemical characterization of ataxia-telangiectasia mutated and Mre11/Rad50/Nbs1. *Methods Enzymol* *408*, 529-539.
- Lesnik, E.A., and Freier, S.M. (1995). Relative thermodynamic stability of DNA, RNA, and DNA:RNA hybrid duplexes: relationship with base composition and structure. *Biochemistry* *34*, 10807-10815.

Lim, H.S., Kim, J.S., Park, Y.B., Gwon, G.H., and Cho, Y. (2011). Crystal structure of the Mre11-Rad50-ATPgammaS complex: understanding the interplay between Mre11 and Rad50. *Genes Dev* 25, 1091-1104.

McPherson, A., and Gavira, J.A. (2014). Introduction to protein crystallization. *Acta Crystallogr F Struct Biol Commun* 70, 2-20.

Meli, M., Sustarsic, M., Craggs, T.D., Kapanidis, A.N., and Colombo, G. (2016). DNA Polymerase Conformational Dynamics and the Role of Fidelity-Confering Residues: Insights from Computational Simulations. *Front Mol Biosci* 3, 20.

Mockel, C., Lammens, K., Schele, A., and Hopfner, K.P. (2012). ATP driven structural changes of the bacterial Mre11:Rad50 catalytic head complex. *Nucleic Acids Res* 40, 914-927.

Molloy, J.E., Burns, J.E., Kendrick-Jones, J., Tregear, R.T., and White, D.C. (1995). Movement and force produced by a single myosin head. *Nature* 378, 209-212.

Moreno-Herrero, F., de Jager, M., Dekker, N.H., Kanaar, R., Wyman, C., and Dekker, C. (2005). Mesoscale conformational changes in the DNA-repair complex Rad50/Mre11/Nbs1 upon binding DNA. *Nature* 437, 440-443.

Nott, T.J., Petsalaki, E., Farber, P., Jervis, D., Fussner, E., Plochowietz, A., Craggs, T.D., Bazett-Jones, D.P., Pawson, T., Forman-Kay, J.D., *et al.* (2015). Phase transition of a disordered nuage protein generates environmentally responsive membraneless organelles. *Mol Cell* 57, 936-947.

Oldfield, C.J., and Dunker, A.K. (2014). Intrinsically disordered proteins and intrinsically disordered protein regions. *Annu Rev Biochem* 83, 553-584.

Park, Y.B., Hohl, M., Padjasek, M., Jeong, E., Jin, K.S., Krezel, A., Petrini, J.H., and Cho, Y. (2017). Eukaryotic Rad50 functions as a rod-shaped dimer. *Nat Struct Mol Biol* 24, 248-257.

- Paull, T.T., and Deshpande, R.A. (2014). The Mre11/Rad50/Nbs1 complex: recent insights into catalytic activities and ATP-driven conformational changes. *Exp Cell Res* 329, 139-147.
- Paull, T.T., and Gellert, M. (1999). Nbs1 potentiates ATP-driven DNA unwinding and endonuclease cleavage by the Mre11/Rad50 complex. *Genes Dev* 13, 1276-1288.
- Ponomarenko, E.A., Poverennaya, E.V., Ilgisonis, E.V., Pyatnitskiy, M.A., Kopylov, A.T., Zgoda, V.G., Lisitsa, A.V., and Archakov, A.I. (2016). The Size of the Human Proteome: The Width and Depth. *Int J Anal Chem* 2016, 7436849.
- Protter, D.S.W., Rao, B.S., Van Treeck, B., Lin, Y., Mizoue, L., Rosen, M.K., and Parker, R. (2018). Intrinsically Disordered Regions Can Contribute Promiscuous Interactions to RNP Granule Assembly. *Cell Rep* 22, 1401-1412.
- Rambo, R.P., and Tainer, J.A. (2011). Characterizing flexible and intrinsically unstructured biological macromolecules by SAS using the Porod-Debye law. *Biopolymers* 95, 559-571.
- Rojowska, A., Lammens, K., Seifert, F.U., Drenth, C., Feldmann, H., and Hopfner, K.P. (2014). Structure of the Rad50 DNA double-strand break repair protein in complex with DNA. *EMBO J* 33, 2847-2859.
- Sanchez, H., Paul, M.W., Grosbart, M., van Rossum-Fikkert, S.E., Lebbink, J.H.G., Kanaar, R., Houtsmuller, A.B., and Wyman, C. (2017). Architectural plasticity of human BRCA2-RAD51 complexes in DNA break repair. *Nucleic Acids Res* 45, 4507-4518.
- Sanchez, H., and Wyman, C. (2015). SFMetrics: an analysis tool for scanning force microscopy images of biomolecules. *BMC Bioinformatics* 16, 27.
- Seifert, F.U., Lammens, K., Stoehr, G., Kessler, B., and Hopfner, K.P. (2016). Structural mechanism of ATP-dependent DNA binding and DNA end bridging by eukaryotic Rad50. *EMBO J* 35, 759-772.

- Smith, F.D., and Scott, J.D. (2013). Scaffolding proteins: not such innocent bystanders. *Curr Biol* *23*, R515-517.
- Smith, M.D., and Jelokhani-Niaraki, M. (2012). pH-induced changes in intrinsically disordered proteins. *Methods Mol Biol* *896*, 223-231.
- Sun, Y., Schroeder, H.W., 3rd, Beausang, J.F., Homma, K., Ikebe, M., and Goldman, Y.E. (2007). Myosin VI walks "wiggly" on actin with large and variable tilting. *Mol Cell* *28*, 954-964.
- Trujillo, K.M., and Sung, P. (2001). DNA structure-specific nuclease activities in the *Saccharomyces cerevisiae* Rad50*Mre11 complex. *J Biol Chem* *276*, 35458-35464.
- Uversky, V.N. (2002). Natively unfolded proteins: a point where biology waits for physics. *Protein Sci* *11*, 739-756.
- Uversky, V.N. (2016). Dancing Protein Clouds: The Strange Biology and Chaotic Physics of Intrinsically Disordered Proteins. *J Biol Chem* *291*, 6681-6688.
- Uversky, V.N., and Dunker, A.K. (2010). Understanding protein non-folding. *Biochim Biophys Acta* *1804*, 1231-1264.
- Uversky, V.N., Gillespie, J.R., and Fink, A.L. (2000). Why are "natively unfolded" proteins unstructured under physiologic conditions? *Proteins* *41*, 415-427.
- van Noort, J., van Der Heijden, T., de Jager, M., Wyman, C., Kanaar, R., and Dekker, C. (2003). The coiled-coil of the human Rad50 DNA repair protein contains specific segments of increased flexibility. *Proc Natl Acad Sci U S A* *100*, 7581-7586.
- Watanabe, H., Takehana, K., Date, M., Shinozaki, T., and Raz, A. (1996). Tumor cell autocrine motility factor is the neuroleukin/phosphohexose isomerase polypeptide. *Cancer Res* *56*, 2960-2963.
- Williams, G.J., Lees-Miller, S.P., and Tainer, J.A. (2010). Mre11-Rad50-Nbs1 conformations and the control of sensing, signaling,

and effector responses at DNA double-strand breaks. *DNA Repair (Amst)* 9, 1299-1306.

Williams, G.J., Williams, R.S., Williams, J.S., Moncalian, G., Arvai, A.S., Limbo, O., Guenther, G., SilDas, S., Hammel, M., Russell, P., *et al.* (2011). ABC ATPase signature helices in Rad50 link nucleotide state to Mre11 interface for DNA repair. *Nat Struct Mol Biol* 18, 423-431.

Williams, R.S., Moncalian, G., Williams, J.S., Yamada, Y., Limbo, O., Shin, D.S., Grocock, L.M., Cahill, D., Hitomi, C., Guenther, G., *et al.* (2008). Mre11 dimers coordinate DNA end bridging and nuclease processing in double-strand-break repair. *Cell* 135, 97-109.

Wlodawer, A., Minor, W., Dauter, Z., and Jaskolski, M. (2008). Protein crystallography for non-crystallographers, or how to get the best (but not more) from published macromolecular structures. *FEBS J* 275, 1-21.

Wright, W.D., Shah, S.S., and Heyer, W.D. (2018). Homologous recombination and the repair of DNA double-strand breaks. *J Biol Chem* 293, 10524-10535.

Wuthrich, K. (1989). Protein structure determination in solution by nuclear magnetic resonance spectroscopy. *Science* 243, 45-50.

Wyman, C., and Kanaar, R. (2006). DNA double-strand break repair: all's well that ends well. *Annu Rev Genet* 40, 363-383.

Wyman, C., Ristic, D., and Kanaar, R. (2004). Homologous recombination-mediated double-strand break repair. *DNA Repair (Amst)* 3, 827-833.

Xu, W., Seiter, K., Feldman, E., Ahmed, T., and Chiao, J.W. (1996). The differentiation and maturation mediator for human myeloid leukemia cells shares homology with neuroleukin or phosphoglucose isomerase. *Blood* 87, 4502-4506.

Chapter 4

Imaging of DNA and Protein by SFM and combined SFM-TIRF Microscopy

M. Grosbart¹, D. Ristic¹, H. Sánchez¹, C. Wyman^{1,2}

Department of ¹Molecular Genetics and ²Radiation Oncology Erasmus University Medical Center, P. O. Box 2040, 3000CA Rotterdam, The Netherlands

Published in Single Molecule Analysis: Methods and Protocols, Second Edition, ed. Peterman E.J.G., Springer Science + Business Media, New York, NY USA.

ABSTRACT

Direct imaging is invaluable for understanding the mechanism of complex genome transactions where proteins work together to organize, transcribe, replicate and repair DNA. Scanning (or atomic) force microscopy is an ideal tool for this, providing 3D information on molecular structure at nm resolution from defined components. This is a convenient and practical addition to *in vitro* studies as readily obtainable amounts of purified proteins and DNA are required. The images reveal structural details on the size and location of DNA bound proteins as well as protein induced arrangement of the DNA, which are directly correlated in the same complexes. In addition, even from static images, the different forms observed and their relative distributions can be used to deduce the variety and stability of different complexes that are necessarily involved in dynamic processes. Recently available instruments that combine fluorescence with topographic imaging allow the identification of specific molecular components in complex assemblies, which broadens the applications and increases the information obtained from direct imaging of molecular complexes. We describe here basic methods for preparing samples of proteins, DNA and complexes of the two for topographic imaging and quantitative analysis. We also describe special considerations for combined fluorescence and topographic imaging of molecular complexes.

Introduction

Proper expression and maintenance of genomic DNA is executed with precision and control by the coordinated action of proteins arranged in specific assemblies on DNA. Understanding how these proteins work together, to package, transcribe, replicate and repair DNA, requires knowing how they are arranged into functional assemblies. Direct images of protein-DNA complexes are an essential tool to achieve this understanding. They reveal a wealth of inherently correlated information on structures, their variation and distributions. Scanning force microscopy (SFM), also commonly known as atomic force microscopy (AFM), is an excellent and practical method for direct imaging of proteins, DNA and complexes of the two at the single molecule/complex level. Molecules and complexes are individually analyzed providing information on the variety of arrangements possible and their frequency in a mixture. Importantly, this type of single molecule structural analysis allows coherent description of features that would otherwise be lost in the averaging of bulk analysis. In addition, direct observation allows correlation of multiple structural features of individual molecular complexes. Sample preparation is relatively simple and requires components biomolecules in easily obtainable amounts and purity. SFM imaging has literally provided a new view of the molecular machinery responsible for DNA processing and by this new insight into molecular mechanisms of vital processes such as DNA packaging, repair, replication and transcription (1, 2)

Mechanistic information is obtained by quantitative analysis of image data. Typically, it is necessary to devise an appropriate scheme to divide complexes or structures observed into relevant categories and determine the distribution of these categories in different conditions. For instance, the percentage of DNA bound by a protein at a specific binding site versus at non-specific sites would be determined as a function of conditions such as the addition of a nucleotide cofactor or another protein. SFM images are also ideal for revealing mechanistically important protein-induced distortions in DNA such as changes in DNA bending, contour length and flexibility. Protein induced distortions of DNA can be determined at specific binding sites and at non-specific sites for comparison, usually from the same sample (3). DNA substrates are constructed with specific sequence or structural features at defined locations, such as a recognition sequence, a single modified or damaged base, nicks, gaps, various lengths of single and double stranded DNA and complex DNA junctions such as those recognized by replication or recombination proteins. In all cases the DNA strands not including the specific feature are by definition non-specific binding sites and serve as unavoidable internal control DNA. Proteins and their functional assemblies often involve multiple DNA sites and strands. These functional assemblies, for example DNA looped between protein bound at two sites or proteins associating to join or connect multiple DNA molecules, are sometimes hard to define by indirect means but obvious by simply looking at images.

Biomolecules are typically deposited onto an atomically flat mica surface and imaged in air. The samples are dried of bulk water but not desiccated and likely retain their native structure (4). The volume of the particles observed can be used to estimate size and multimeric state. DNA can be deposited on mica by equilibration on the surface so that it is not kinetically trapped. In this way, the arrangement of the DNA on the surface accurately reflects its properties in solution, such as contour length, flexibility and presence of bent segments. Protein-DNA complexes prepared in appropriate biochemical conditions are deposited onto mica in a similar fashion. Changes in the DNA induced by bound-proteins can be accurately measured. Some proteins have features that are distinct in SFM images, such as long coiled coil regions or multiple globular lobes, but most proteins appear as similar globular objects. Here we describe basic methods for obtaining topographic images of DNA, proteins and their complexes that can be used for a variety of quantitative structural analyses.

Many functionally important protein-DNA complexes include multiple component proteins. The identity of components in multi-protein complexes can be estimated based on their volume and known molecular weights. However, it is not always possible to determine molecular composition and stoichiometry unambiguously by volume alone. It is therefore necessary to label or tag specific proteins to identify different, possibly similarly sized, components in complex assemblies. Combining fluorescence and topographic imaging allows specific identification of fluorescently labeled proteins and greatly

expands the application of SFM to analysis of increasingly complex molecular assemblies. Several instruments that combine SFM topographic imaging with fluorescent imaging are currently commercially available. There are specific challenges for using these instruments for analysis of protein-DNA complexes that we address here. It is necessary to deposit the molecules of interest onto an atomically flat surface such as mica for topography but this substrate must also be sufficiently optically transparent to allow fluorescence. In addition, appropriate marker objects are essential to achieve nm accuracy in aligning optical and topographic images when the objects of interest are smaller than optical resolution. We briefly describe methods for sample preparation and image alignment that allow properly correlated SFM and fluorescence imaging.

Materials

Instrumentation

1. Scanning probe microscope: These instructions are guided toward eventual imaging by intermittent contact mode in air. We have a Digital Instruments MultiMode Scanning Probe Microscopes. The methods for sample preparation and guidelines for data acquisition are applicable to any similar instrument and imaging mode. (see Note 1)
2. Computers (PC, Mac) that meet requirements of image analysis software. Software like SFMetrics (5), ImageJ (<http://rsbweb.nih.gov/ij/>), Image SXM (extended version of NIH image by Steve Barrett, Surface Science Research Centre,

Univ. of Liverpool, Liverpool, U.K.), WSxM (Nanotec Electronica S.L.), or similar image analysis software is required for quantifications of various features of visualized molecules.

General supplies

1. Glass pasture pipettes
2. Standard facial tissues
3. Lens cleaning tissues (Whatman 105)
4. Forceps (DZM, Italy)
5. A source of filtrated air or N₂
6. For preparation of all solutions, MiliQ filtered de-ionised water (Resistivity $\leq 16 \text{ M}\Omega \text{ cm}$, TOC 1-5ppb)
7. Standard protein deposition buffer consisting of: 20 mM HEPES-KOH pH 7-8, 50-100 mM KCl (or NaCl), 1 mM DTT (chemical supplied by Sigma-Aldrich).
8. Standard DNA deposition buffer consisting of: 5-10 mM HEPES-KOH pH7.5, 5-10 mM MgCl₂ (chemical supplied by Sigma-Aldrich).

Sample substrates

1. Metal discs such as those usually supplied with scanning probe microscope
2. Mica sheets (Muscovite Mica V-5 Quality, Thickness=0.15mm - 0.21mm, Electron Microscopy Sciences)

3. Punch & die set (Precision brand) for cutting mica into discs. The diameter of mica discs is 1-2mm smaller than diameter of metal discs
4. Superglue to attach mica discs to metal discs
5. Invisible tape (19 mm width, Magic from 3M) to cleave mica

DNA preparation

General molecular biology reagents and instruments for preparation of DNA (purification from bacteria or PCR amplification) are needed. General knowledge on methods for DNA preparation, such as those found in various editions of Molecular Cloning: A Laboratory Manual from CSHL Press, is assumed.

Combining Fluorescence and SFM

Instrumentation

For identifying specific molecules and nano-objects in SFM topography images we use a combined SFM-fluorescent microscope set up which consists of: An inverted fluorescence microscope equipped with high numerical aperture (1.45) objectives with a minimum magnification of 60x (Nikon TE2000); signal detections with a Cascade II:512B EMCCD camera (Princeton Instruments); running MetaMorph software (Molecular Devices) or custom made Labview (National Instruments) software for microscope operation and image acquisition; and a coupled NanoWizard®II scanner (JPK

instruments). Similar instruments and equivalent components would perform as well.

Sample substrates for combined fluorescence and SFM

1. Mica sheets (muscovite V-1 quality, from Electron Microscopy Science)
2. Glass cover slips, round, 24 mm diameter, thickness 00 (from Menzel-glazer)
3. Optical adhesive NOA88 (Norland products)
4. Hand held UV lamp
5. Sodium tetrahydridoborate 0.25% w/v solution in water (see Note 2)
6. Invisible tape (19 mm width, Magic from 3M)

Fluorescent markers and labels

1. FluoSpheres® carboxylate-modified microspheres (0.04 μm diameter, yellow-green fluorescent (505/515), orange fluorescent (540/560) and red fluorescent (580/605) from Invitrogen)
2. TransFluoroSpheres® (0.04 μm diameter, dual green and red fluorescent (488/645) from Invitrogen)

Methods

Considerations for preparing proteins used in SFM imaging

Purified proteins should be stored and used without addition of stabilizing proteins such as BSA. The protein purity requirements differ with application and the nature of possible contaminants. For example, if the protein of interest binds to DNA, contaminating proteins not bound to DNA can be ignored. Also, DNA binding proteins should be without any trace of DNA. In general, $\geq 80\%$ protein purity, estimated by coomassie blue staining of the purified protein displayed by gel electrophoresis, is sufficient for SFM analysis.

Considerations for preparing DNA used in SFM imaging

The length of DNA to be used depends on the eventual data desired and the specific experimental question. In general DNA should be at least 500 bp or longer so that it is obviously DNA by appearance based on relative width, height and length. We commonly use DNA in the range of 500 bp to 3000 bp. Linear DNA is generally more useful. Circular DNA tends to fold over itself on the surface making it harder to analyze. In addition linear DNA allows determining the location of a specific sequence or feature of interest by its relative position from an end. DNA should be clean, free of proteins or other material that will deposit onto mica and complicate the imaging. Kits/columns for DNA purification (Quiagen, GE Healthcare, Sigma-Aldrich) usually produce DNA of sufficient purity for SFM, though some problems with residual column material or buffer components occasionally

occur. The cleanest DNA is obtained by the following methods depending on the source: from solutions such as PCR reactions or Enzyme digestion purify DNA by phenol:chloroform extraction followed by chloroform extraction; from gel slices purify DNA by electroelution; DNA from plasmid or phage preparations by cesium gradient purification. These DNA purification and isolation methods often reduce final yield. Therefore the trade off between purity and yield has to be considered when choosing purification methods for individual applications. Ethanol precipitation tends to result in contaminating material when imaging, often assumed to be excess salt. For that reason ethanol precipitation is avoided or preformed with care to wash excess salt from pellets before re-suspension.

Preparation of mica substrates

1. Attach mica to metal by applying a very small drop of superglue onto the metal discs placing the mica disc over the glue and gently pressing. Avoid glue spreading beyond the mica as this will interfere with cleaving the mica.
2. Freshly cleaved mica is prepared by applying scotch tape (Magic, 3M) to the mica glued to metal and peeling off the top mica layers. The peeled off layer stuck to the tape is inspected to see if it is a complete circle. If a complete layer was not cleaved off the procedure is repeated until a complete layer, smooth unbroken circle, is removed. The mica is usually cleaved only a few minutes to a half hour before use to assure a clean surface.

Immobilizing molecules on mica: General considerations

In order to visualize molecules in scanning force microscopy they have to be immobilized on a surface. However, molecules have to be free from the surface to enable their dynamic interactions and to prevent steric hindrance that might affect molecular interactions. Thus the immobilization of molecules on the surface has to be carefully controlled to enable imaging while minimizing disturbing the relevant molecular interactions. Molecules can be deposited on a surface through specific or non-specific interaction. We most often take advantage of relatively non-specific electrostatic adsorption, which depends on charge of the surface and molecules, and is sufficient to provide controlled attachment of DNA, proteins and their complexes. Specific interaction such as streptavidin and biotin or digoxigenin and anti-digoxigenin, provide much stronger attachment of molecules to the surface with defined molecule orientation. However, surface modification also increases roughness and interferes with imaging.

The most commonly used surface for deposition of biomolecules is muscovite mica. Mica can be cleaved at crystal planes that produce large atomically flat surfaces. This uniform flat surface allows detection of biomolecules that are only a few nanometers high. The mica surface is negatively charged and the heterogeneous charged domains on most proteins result in sufficient deposition without additional treatment of either surface or protein. HEPES buffers, in biologically relevant pH range (pH7-8), are preferred for protein deposition. Tris-HCl

buffers tend to deposit on mica and interfere with SFM imaging. Other common protein storage or reactions buffer components, such as KCl, MgCl_2 , DTT, ATP, low concentration of detergents like NP40, and glycerol, do not interfere with protein deposition or eventual imaging in our experience. The composition of buffers used for deposition of proteins is less strict than for deposition of DNA because proteins deposit effectively onto mica in a wider range of salt and pH conditions. Thus it is a good first step to deposit proteins in buffers that are optimal for maintaining protein structure and/or activity. If the protein of interest is already biochemically characterized, the optimal buffer for deposition for SFM would be the same (or very similar) as the optimal buffer for protein activity. In case of an uncharacterized protein, SFM deposition can be done in a standard deposition buffer consisting of: 20 mM HEPES pH 7-8, 50-100 mM KCl (or NaCl), 1 mM DTT. If the protein appears aggregated upon deposition increasing salt concentration and/or including some NP40 (0.05%) often helps to prevent undesired protein-protein interaction.

Adsorption of negatively-charged DNA on negatively charged mica surface requires the presence of divalent cations. Those interested in the effects of different divalent cations on DNA deposition on mica are referred to published studies on this topic (5, 6, 7, 8)

The mica surface can be modified for more efficient adsorption of DNA. This has advantages and disadvantages. The treatment of mica with either 3-aminopropyltriethoxysilane

(APTES) or poly-L-lysine results in a positively charged surface. DNA binding to such surfaces does not require the addition of divalent cations. The resulting DNA is strongly attached to the modified surface and can be imaged by SFM both in air and in buffer. However, on these modified surfaces DNA does not equilibrate but become kinetically trapped (9).

Spermidine can be added to the sample to increase the affinity of nucleic acids to mica. The presence of spermidine allows DNA binding to mica without divalent cations and/or in the presence of higher concentrations of monovalent salt. By altering the concentration of spermidine DNA deposition can be controlled to be either by equilibration or by kinetic trapping (Figure 1). A final concentration of 0.05 mM spermidine is sufficient to provide DNA deposition by equilibration in a sample with 150mMKCl even without Mg^{2+} .

Deposition by kinetic trapping results in DNA conformations on mica that are strongly influenced by interaction with the surface. This complicates and in some cases precludes determining experimentally important changes in DNA conformation, such as measuring DNA bends or distortions induced by protein binding or changes in DNA conformation induced by binding small molecules prior to deposition. In order to measure and analyze protein induced changes in DNA we do not use treated mica (9). In addition any mica treatment results in a rougher surface that can complicate imaging.

The persistence length of DNA can be determined from SFM images and is used as a test of deposition by equilibration on

the surface. Persistence length reflects intrinsic flexibility of a polymer and is a well-characterized feature of DNA. Persistence length is obtained by measurements of the contour length and the end-to-end distance of deposited DNA as described in (9).

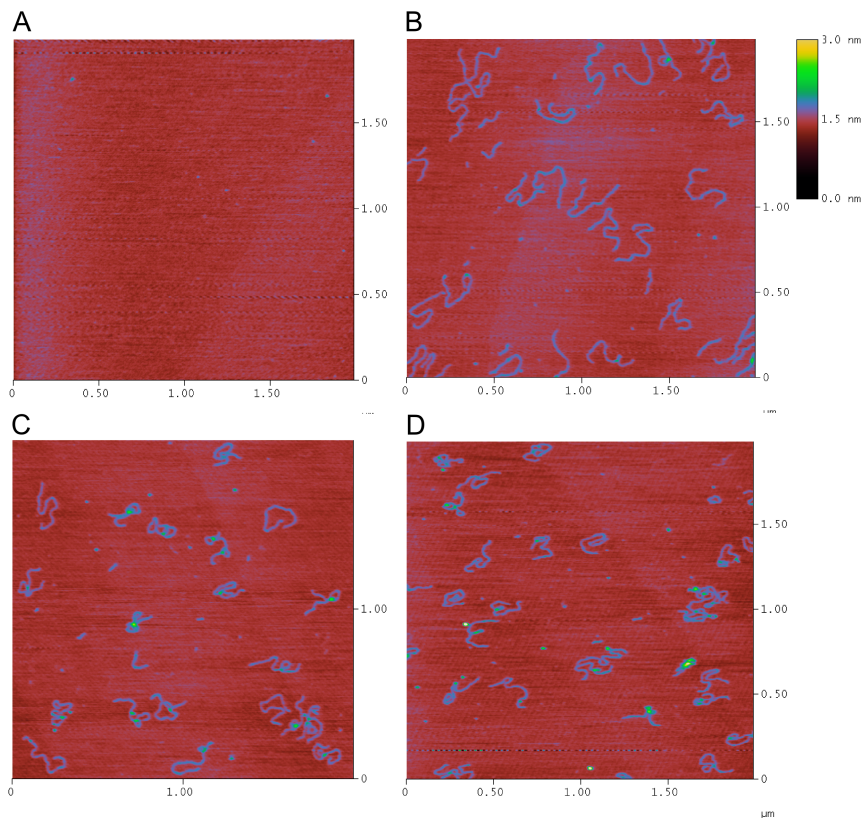


Fig 1. Example of different concentration of spermidine on DNA deposition DNA (linear double stranded 1.8kb) at 0.5 ng/μl in 10 mM Hepes and 150 mM KCl was used for deposition. (a) Addition of 0.01 mM spermidine was not enough to provide binding of DNA for mica. (b) Addition of 0.05 mM spermidine enabled deposition of DNA by equilibration. Increasing concentration of spermidine to 0.1mM (c) or 1mM (d) caused kinetic trapping of DNA on the surface. The image dimensions are labeled, height is indicated by the grey scale shown to the right.

A persistence length of ~ 50 nm is characteristic of DNA and thus indicates deposition by equilibration.

We have described that proteins adhere to mica in a wider variety of buffer conditions than DNA. Because of this, also protein bound DNA can adhere to mica due to the behavior of the proteins. Thus a protein-bound DNA will often adhere to mica in deposition conditions that are effective for the protein alone without the need to add Mg^{2+} or reduce monovalent cation concentration. It should be noted however that free DNA, not bound by proteins, will behave differently and may not adhere to mica in conditions where protein-DNA complexes. Thus deposition conditions should be chosen carefully depending on the goal of the experiment. For instance if it is necessary to quantify the amount of protein bound and protein free DNA, controls that show equal deposition of these need to be included.

The best conditions for protein-DNA complex formation should be determined by standard quantitative biochemical assays prior to initiating SFM imaging experiments. This will form the starting point for conditions to use in SFM imaging. Binding reactions for SFM typically require more concentrated DNA than many biochemical assays and there are limitations to the excess of protein that can be used. As a general rule the solution deposited onto mica should have DNA at 1-10 ng/microliter and proteins should not be more than 50 fold molar excess to their binding sites. The upper limit to excess protein is due to problems observing DNA if the surface becomes covered by protein or even

more dramatically if protein saturates the surface and prevents DNA binding. Thus protein to DNA ratios and protein concentrations optimal for biochemical assays may not be optimal for SFM imaging.

Deposition of protein for SFM imaging

1. Prepare a solution of protein to be deposited, concentration of about 0.5 μM in an appropriate buffer is a good starting point. Optimal protein concentration for deposition will differ for each protein depending, for instance, on purity, size and oligomeric state. Usually, protein concentrations of less than 0.5 μM (final in deposition buffer) provide reasonable coverage of the mica surface for further analysis.
2. Place a drop, 5-30 μl depending on the size of the mica surface, of protein solution onto the freshly cleaved mica surface and let it sit to deposit for $\sim 30\text{sec}$.
3. Rinse the mica surface with MiliQ filtered de-ionised water, about one pasture pipette full as shown in Figure 2.

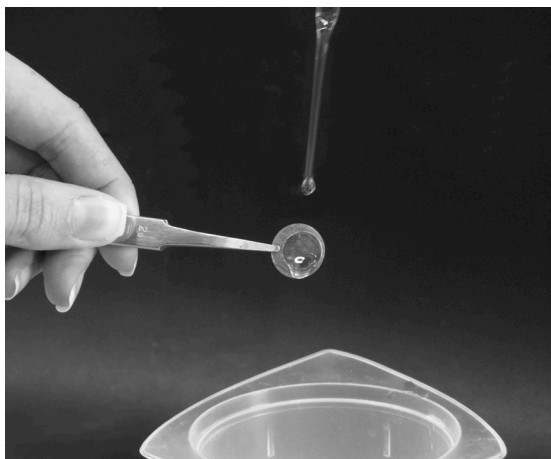


Fig 2. Washing mica surfaces with water. The sample substrate is held by forceps on the edge of the metal disc, not over the mica. Water or buffer is washed over the surface from a pasture pipette

4. Excess of water is blotted by touching a piece of facial tissue to the edge of the mica.
5. The mica surface is then dried in a stream of filtrated air (or N₂).
6. Observe the sample with the SFM, typically scanning fields of 1 μm X 1 μm with a Z scale of 5 nm or less will give a good impression of the protein coverage.
7. Assess the protein coverage and modify deposition if needed. If deposition is too crowded (Figure 3A), additional dilution of 5 to 10 fold usually results in surface coverage where proteins are nicely separated on the surface and can be analyzed (Figure 3B).

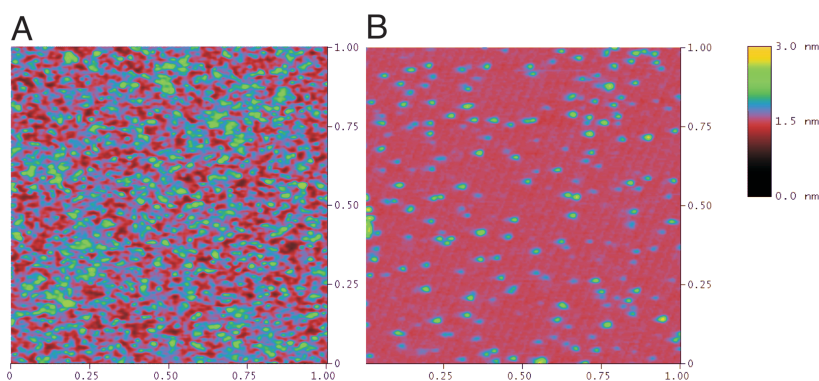


Fig 3. Examples of excess or appropriate protein coverage. The 37 kD human RAD51 protein was deposited as described from a buffer containing 25 mM HEPES-KOH pH 8.0 and 100 mM KCl and imaged in tapping mode in air with a Digital Instruments multimode NanoScope. (A) Too much protein covering the mica surface prevents identification and analysis of individual molecules. Here a 2 μM solution was used for deposition onto mica. (B) Adequate protein deposition showing many single molecules well separated and few overlapping unresolved molecule pairs. Here a 0.2 μM solution was used for deposition onto mica. The image dimensions are shown, height is indicated by color according the scale shown on the right.

Deposition of DNA for SFM imaging

1. Prepare a solution of DNA to deposit, DNA 0.5 -10 ng/ μl in deposition buffer of 5-10 mM HEPES pH7.5, 5-10 mM MgCl_2 . Since the presence of monovalent cations dramatically reduces Mg^{2+} promoted adsorption of DNA to mica, the concentration of monovalent salt in the solution being deposited should be less than or equal to the concentration of Mg^{2+} .
2. Place a drop, 5-30 μl depending on the size of the mica surface, of the solution containing DNA onto the freshly cleaved mica surface and let it sit to deposit for 30 sec to 1 min.
3. Wash the mica surface with MiliQ filtered de-ionised water, about one pasture pipette full, as shown in Figure 2.
4. Excess of water is blotted by touching a piece of facial tissue to the edge of the mica.
5. The mica surface is then dried in a stream of filtrated air (or N_2).
6. Observe the sample with the SFM, typically scanning fields of 2 μm x 2 μm with a Z scale of 2 nm will give a good impression of the DNA coverage.
7. Assess the DNA coverage and modify deposition if needed. Excess DNA coverage is shown in Figure 4A and appropriate coverage of DNA on mica is shown in Figure 4B. If the DNA on the surface is too crowded then try 2-10 fold dilution of the solution for deposition. If there is not enough DNA on the

surface (only one, a few or no molecules in a 2 μm x 2 μm scan) one of several variations in the deposition will usually help: increase the time for the DNA solution sitting on mica to 1-2 min, increase the DNA concentration, increase the Mg^{2+} concentration, or decrease the monovalent cation concentration.

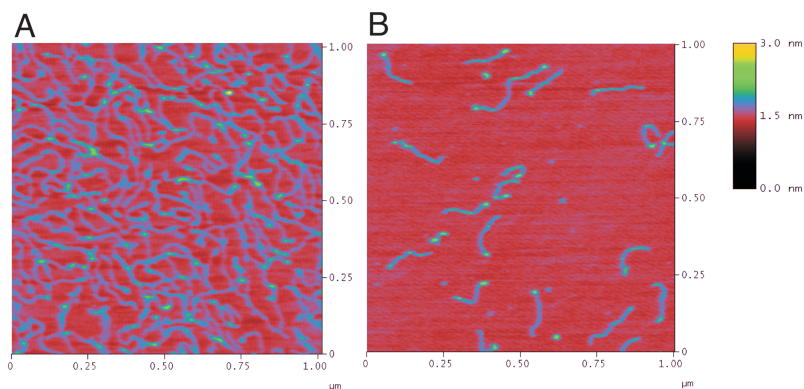


Fig 4. Examples of excess and appropriate DNA coverage. DNA was deposited onto mica as described and imaged in tapping mode in air with a Digital Instruments multimode NanoScope. (A) Too much DNA covering mica prevents identification of individual molecules and precludes any meaningful analysis. Here a solution of 800 bp linear double-stranded DNA at 10 ng/ μl in 10 mM HEPES, 10 mM MgCl was used for deposition onto mica. (B) Appropriate DNA density with clearly separated and non-overlapping DNA molecules. Depending on the length of the DNA there should be 5-20 molecules in a field. Here a solution of 0.5 ng / μl DNA in 10 mM MgCl was used for deposition onto mica. This resulted in about 15 isolated, identifiable and analyzable DNA linear DNA molecules. This linear DNA consists of a 500 bp double-stranded and 300 nt single stranded segment. The single-stranded DNA appears as a small knob at one end of the linear molecule. The image dimensions are shown, height is indicated by color according the scale shown on the right.

Deposition of Protein-DNA complexes for SFM imaging

1. Prepare a protein-DNA binding reaction for deposition. This typically should include DNA at 1-10 ng/ μ l (see Note 3).
2. Place a drop of the protein-DNA complex sample onto mica, as described above for protein or DNA alone, and allow to sit for 30 sec to 1 min.
3. Unbound material is washed off with a small volume of binding buffer as shown in Figure 2.
4. Excess buffer is removed by touching a tissue the edge of the mica surface, but the surface is ***not dried***.
5. The surface is then covered with 30-40 μ l of buffer containing 10 mM HEPES-KOH (pH7.5), 10 mM $MgCl_2$ in order for DNA to attach to mica.
6. Almost immediately or after about 5 sec the mica is washed with water, about 1 pasture pipette full, as shown in Figure 2.
7. Excess liquid is then blotted by touching a tissue to the edge of the mica, and the surface is dried in a stream of filtered air as described above.

Deposition of Protein-DNA complexes and free DNA in binding reactions for SFM imaging

1. Prepare a binding reaction and dilute with deposition buffer (10 mM HEPES, 10 mM $MgCl_2$) so that the concentration of DNA is 1-10 ng/ μ l. Since the presence of monovalent cations

dramatically reduces Mg^{2+} promoted adsorption of DNA to mica, the concentration of monovalent salt in the solution being deposited should be less than or equal to the concentration of Mg^{2+} (see Note 4). This will allow free DNA as well as protein-bound DNA to adhere to mica. The initial binding reaction may need to be adjusted so that after dilution the solution that will be deposited still contains DNA at 1-10 ng/ μl (see Note 5). If the concentration of monovalent salt in the solution being deposited has to stay high, addition of 0.05mM spermidine (final concentration) will allow both free DNA and protein-bound DNA to bind to mica.

2. Place a drop of the protein-DNA complex sample onto mica, as described above for protein or DNA alone, and allow to sit for 30 sec to 1 min.

3. Unbound material is washed off with water, about 1 pasture pipette full, as shown in Figure 2.

4. Excess water is blotted by touching a piece of facial tissue to the edge of the mica.

5. The mica surface is then dried in a stream of filtrated air (or N_2).

6. Observe the sample with the SFM, typically scanning fields of 2 μm x 2 μm or 4 μm x 4 μm with a Z scale of 2-5 nm will give a good impression of the sample. Density of DNA on the surface, either with or without bound protein, similar to that shown in Figure 4B is sufficient for most analysis (see Note 6).

Guidelines for collecting images sets for analysis

Once the stoichiometry of DNA and protein is optimized for imaging and good coverage of mica is achieved, for proteins alone, DNA alone or protein-DNA complexes, a collection of images needs to be obtained for eventual analysis.

1. Issues of actual microscope operation and image acquisition are not discussed or described here (refer to specific instrument operating manual). The operation of a SFM and data acquisition differ depending on the instrument and are beyond the scope of this article. The samples we have described are usually imaged in our lab using intermittent contact or “tapping” mode in air. We use standard silicon tapping tips for a variety of suppliers with equivalent success. It is important that the tips have a confirmed end radius of curvature of about 10 nm or less (see Note 7).

2. For most applications scan sizes ranging from 1 μm x 1 μm up to 4 μm x 4 μm are most useful. For instance when analyzing proteins, scans of 1 μm x 1 μm usually provide sufficient resolution and sufficient data per image. For DNA-protein complexes that individually cover more surface scans of 2 μm x 2 μm or 4 μm x 4 μm are better (see Note 8). In any case all images to be used in the same analysis need to be the same size and resolution.

3. Images should be collected from non-overlapping fields without selection for areas of interest. If scanning is stable and interference free, images can be collected without changing

scanning parameters, the autoscan function of the microscope software can be used to collect an unbiased series.

4. Most analysis requires a significant number of molecules or complexes to measure. For example if 100 or more DNA-protein complexes have to be analyzed and the reaction results in 1/5 of the DNA molecules being bound by protein and deposition results in 10 DNA molecules of about 1 kpb long in a 2 μm x 2 μm field, then a minimum of 50 such images need to be collected. It will take considerable time to collect these 50 images. In the optimistic case that everything works well this is a half-day of collecting images. Considering that images are needed for control samples, planning is a very important to perform efficient imaging experiments.

Guidelines for Standard Image analysis

A number of features of protein, DNA and protein-DNA complexes can be measured from SFM images (1). The size of proteins alone or bound to DNA can be accurately determined from SFM images. Also, protein-induced changes in DNA, such as wrapping, elongating and bending of can be quantitatively described.

DNA length can be measured from SFM images imported into specialized software such as IMAGE SXM, a customized version of NIH Image modified to automatically import image data from a variety of commercial scanning probe microscopes, or SFMetrics a recently develop tool based on MATLAB that can be run in

different platforms and adds several features like skeleton length to analyze irregular objects applying user-adjusted thresholds (5).

1. We determine DNA contours by manually tracing in an appropriate image analysis program. In the case of DNA–protein complexes, contour length is traced as the shortest possible DNA path through the bound protein. Custom software that can automate DNA length measurements has been developed in several labs but is not currently commercially available. The length of DNA \pm protein will indicate whether bound protein alters DNA by wrapping or stretching (see for example (10)).

2. The volume of proteins (not bound to DNA) can be determined using a semi-automated method developed by Glenn C. Ratcliff and Dorothy A. Erie (11). Using this approach, large protein populations can be analyzed in a rapid and accurate manner.

3. The volume of DNA-bound protein complexes has to be determined by manual tracing. Area and average height of complex are measured, and a background volume of the same traced area at an adjacent position including DNA is subtracted (12, 13).

4. DNA bending is a feature of many DNA binding proteins, see for example (3, 14). Using SFM, DNA bending can be directly evaluated. A comparison of different methods to determine DNA bend angles is presented in (15).

Preparation of mica substrates for combined SFM-Fluorescence

1. Cut the mica disc slightly smaller than the size of the cover slips.
2. Using a scalpel, split the individual mica discs into 2 or 3 thinner layers
3. Cleave the layered mica disc with tape until almost transparent. The color at this stage should still be slightly brownish. Be sure that one side of the mica surface is clearly flat (by eye) with no irregularities (as described in section 3.3). The flat side of the mica is to be glued to the cover slip.
4. Place the cover slip to be used onto a lens cleaning tissue (Whatman 105).
5. Put a small drop of the optical glue in the end of a yellow tip pipette.
6. Place the mica disc over the glue with the flat side facing the glue.
7. Tape the mica and the cover slip and attach to the lens cleaning tissue as shown in Figure 4. Press down on the middle of the mica with your thumb for a homogeneous distribution of the glue under the surface.
8. Cure the UV glue by placing the UV light around 4 cm over the cover slips and switch on the 350nm light for 3 minutes (see Note 9). It is most convenient to keep the glass-glued mica

taped to the lens tissue (Figure 5). A piece of lens tissue including the taped down cover slip is cut to about the size of a microscope slide. Accidentally breaking the cover slip is avoided by handling the tissue. The mica-cover slips can be stored for future use for as long as necessary.

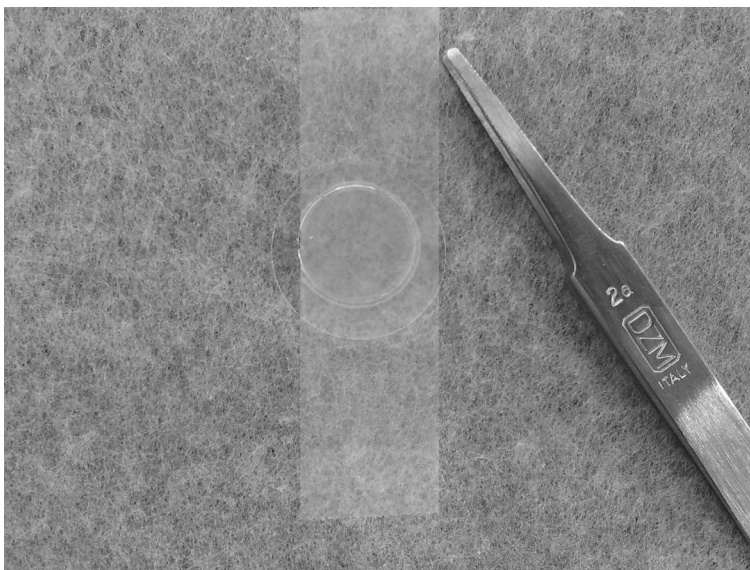


Fig 5. Illustration of preparing mica glued to glass cover slips. The mica is glued to a glass cover slip as described and then attached with tape to a lens cleaning tissue as shown. The tissue is used to pick up the fragile glass cover slip and the sample surfaces are also stored in this way.

9. Before using the mica-cover slip, the mica surface should be made as thin as possible, by cleaving as many layers as possible using the Scotch tape. The transparency of the mica now should be close to 100%, without cracks or other visible imperfections, otherwise the focusing step with the short

distance working objectives would be simply impossible (see Note 10).

10. Once the thickness of the mica-coverslip has been assessed (see Note 10), it can be treated by placing a 200 μ l drop of 0.25% w/v sodium tetrahydroborate on the mica for 20 minutes to reduce autofluorescence. Wash with water 4-5 times as shown in Figure 2. The surface is now ready for depositing the solution containing the DNA and protein complex to be analyzed.

Fluorescent DNA and proteins for combined SFM-Fluorescence: General considerations

General considerations for DNA and proteins: Fluorescence labels can be attached to either DNA and or protein. DNA is typically labeled by: (1) PCR using fluorescence nucleotides analogs for uniform labeling if the fluorophore does not interfere with protein interactions; (2) with PCR primers including a 5' fluorophore; (3) Conjugating biotin to DNA, also introduced by PCR with 5' biotin modified primers, bound to any of a variety of streptavidin coupled fluorophores.

Sample preparation for combined SFM-Fluorescence

1. Prepare a DNA, Protein, or DNA-protein complex binding reactions in appropriate functional conditions.
2. Dilute the binding reaction into 20 μ l of deposition buffer (10 mM HEPES, 10 mM $MgCl_2$, including 3 pM (Trans)FluoSpheres® (see note 11) markers) according to

guidelines given above in sections 3.4 to 3.8. Note, that the distribution of molecules should be 5 to 10 times sparser than in conventional SFM imaging in order to achieve a clear image of single fluorophores (Figure 6).

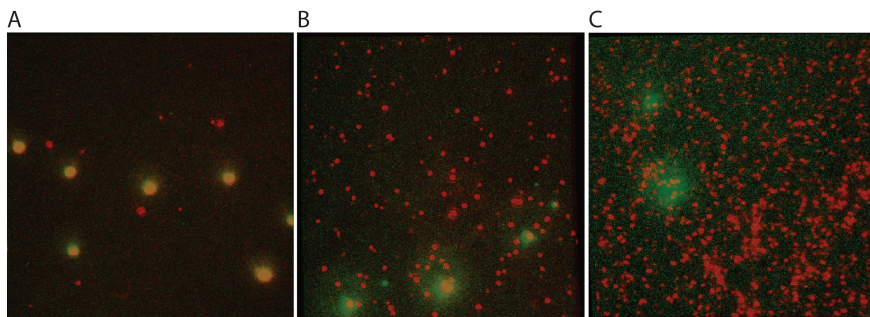


Fig 6. Examples of sparse, appropriate and excessive fluorescent signal. TransFluoroSpheres® (488/645) and single stranded DNA oligo coupled to Cy5 were deposited for combined SFM and TIRF imaging as described. The samples were imaged in tapping mode in air with a JPK NanoWizard scanner mounted on a Nikon TE2000 microscope. All three panels represent overlays of 35 x 35 µm optical images created by excitation at 488 and 633. (A) Example of deposition in which the Cy5 fluorescent signal is too sparse considering the size of area imaged. (B) An appropriate coverage of the area with Cy5 signal is plentiful, but clearly separated into individual spots. (C) Example of excessive coverage of the surface with signal coming from Cy5. Individual spots merge and cannot be distinguished in most cases. In grey scale image FluoSpheres® signal (green in color) appears as larger diffuse spots, Cy5 signal (red in color) appears as smaller defined spots.

3. Place a 20-40 µl drop of the diluted sample onto the freshly cleaved mica surface and let sit for 1 min.

4. Unbound material is washed off with water, about 1 pasture pipette full, as shown in Figure 2.

5. Blot off excess liquid by touching a tissue to the edge of the mica.
6. Dry the sample in a stream of filtered air.
7. If the use of deposition buffer needs to be omitted (for instance if dealing with cation-sensitive reactions), the sample can be diluted in reaction buffer and spermidine added just prior to deposition to a final concentration of 5 $\mu\text{g/ml}$. If any charged molecules need to be avoided in the reaction mixture, 3 pM fiducial markers should be deposited separately in total of 20 μl of deposition buffer, as described in steps 3-6. After washing and drying the mica, reaction mixture diluted in adequate buffer can be re-deposited in the same fashion.

General considerations for imaging with combined SFM-Fluorescence

Familiarity with the available fluorescence microscope set up is assumed. Specific operating instructions will vary depending on the set up and are in any case beyond the scope of this article. With respect to excitation light source we note that a mercury lamp is often sufficient for visualization of quantum dots, (Trans)FluoSpheres®, and to a limited extent single fluorophores like Alexa 633. Laser excitation specific to the dyes used is preferable for most single fluorophore visualization applications. In our set up the optical microscope is coupled to a NanoWizard®II scanner (JPK instruments). In this instrument correlation of fluorescence and topographic images is accomplished first by DirectOverlay™ software. Further

registration accuracy can be achieved using numerical computing software like MATLAB (16). Optical images are obtained at the highest magnification possible, usually as 60 mm fields. We typically use total internal reflection for optimal imaging of surface attached molecules. Topographic images are obtained as fields of 2-35 mm depending on the size of the objects deposited and the eventual analysis needed.

Using (Trans)FluoSpheres® to align topographic and optical images

1. Find an area of interest and optimize focus for optical signals. After mounting the sample in the holder on the microscope, find the fluorescent signal coming from the FluoSpheres® and biological sample. Focus on the fluorescent signal coming from the sample (smaller, less intense spots). Due to bleaching fluorescence from the sample, it is recommended to move into a new field to obtain an image that will be later used for overlay.

If using TransFluoSpheres® to bring the sample into focus, use an excitation wavelength of the fiducial marker that does not match the fluorophore coupled to the sample to optimize focus. Once that is done switch to the excitation wavelength that will reveal both of the sample and fiducial, adjust focus and proceed. When using TransFluoSpheres® obtain an image or image stack at both excitation wavelengths

2. Obtain an optical image with at least 3 (Trans)FluoSpheres® (see Note 12). For single molecule nanometer localization and

eventual quantification of fluorophore number based on intensity, obtain a stack of images (time lapse). The exposure time, intervals and frame number should be chosen such that complete bleaching of the sample is achieved (for example 300 ms exposure, 350 ms interval, 300 frames). Maximum intensity image created from the stack will be used for registration by overlaying with the SFM scan.

3. Monitor the progress of image acquisition; should the sample go out of focus or when noticeable drift occurs, the entire process needs to be repeated. In our experience drift occurs when the mica is not optimally thin and the objective exerts pressure on the cover slip to focus. Alternatively an unstable sample holder can result in drift.

4. Before obtaining the SFM image, minimize sources of vibration in the vicinity of the measurement. If using an immersion lens and a sample holder, make sure that both are not touching the sample during SFM imaging in order to minimize background vibrations. Retract the lens and if necessary detach the sample holder (depending on microscope set up this may or may not convey vibrations).

5. Obtain a topographic scan of the same area captured in the optical image at a resolution that will allow clear identification of the 20-40 nm diameter (Trans)FluoSpheres®. For example, optimal parameters for 35 x 35 μm field are 512 x 512 lines at scan rate 2-2.5 Hz.

6. Use the microscope software to overlay the two images so that the topographic image of the spheres (always smaller than the diffraction limited fluorescent spot) coincides with the center of the optical spots (Figure 7). This overlay now defines the register of the fluorescent and topographic images with accuracy in the 10's of nm range (depending on the microscope software routine applied) and can be used to identify fluorescent signals as belonging to specific topographic features (such as specific proteins or fluorescently labeled sites on DNA). If using TransFluoroSpheres®, use the (maximum intensity) image showing only the fiducial spheres to identify their position in the SFM scan.

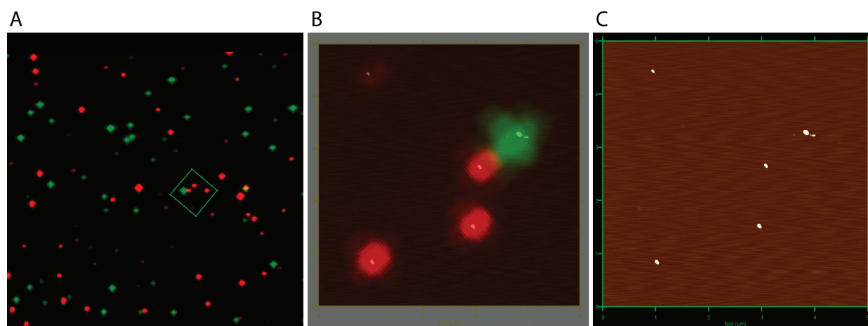


Fig 7. Aligning optical and topographic images using FluoSpheres®. A mixture of polystyrene spheres with three different colors: red, green and orange, were deposited for combined SFM and fluorescence imaging as described. The samples were imaged in tapping mode in air with a JPK NanoWizard scanner mounted on a Nikon TE2000 microscope. The density of FluoSpheres® shown is appropriate for both optical and topographic imaging. (A) Optical image, 60 μm x 60 μm , of a mixture of polystyrene spheres with three different colors: green and red channel were overlaid. Orange polystyrene beads are recognized by the colocalization of both signals. The indicated area (5 μm x 5 μm , green square) was chosen for scanning force microscopy. (B) Overlay of the fluorescence signal with the height image. (C) Topography image of the selected area, Z scale 0-30 nm.

The overlay in B shows that the optical and topographic images can be aligned by centering the height image of the FluoSpheres® in the optical signals from the same objects. Panels B and C present the area scanned rotated about 135 degrees clockwise relative its position in panel A.

7. Once the images are correctly registered in the microscope software, select a region of interest to obtain a SFM image of the desired resolution. Multiple small scan images, for example $2 \times 2 \mu\text{m}$ fields at 512×512 lines, can be collected. The best overlay/resolution combination is obtained by imaging large areas, we typically collect images as $35 \times 35 \mu\text{m}$ fields at 4096×4096 lines at a scan rate of 0.2 Hz (overnight). For nanometer localization this area should contain at least 3 fiducial spheres.
8. Overlay the (maximum intensity) optical image of the fluorescent signal corresponding to the biological sample with the high resolution SFM scan.

Notes

1. Here we describe the use SFM imaging of air-dried samples with the following standard settings: scanned in intermittent contact mode (air), using Silicon Tapping/Non-Contact Mode tips 125 mm in length with a spring constant of 25–75 N/m from Applied Nanostructures, drive frequency of the cantilevers on average 300 kHz, and a line rate of 2 Hz to acquire images.
2. Tetrahydridoborate is a flammable solid and should be handled according to all precautions indicated from the supplier. The solution should be made in a fume hood.
3. Because deposition is not always perfect but is easy to repeat it is a good idea to make several depositions of the same sample. When working with protein-DNA binding reactions depositions should be done at the same time to

keep the binding conditions, such as incubation time, constant. It is also a good idea to deposit different dilutions, differing by a factor of 2 to 5, at once to assure that one will be appropriate for analysis without requiring repetitive binding reactions and additional material.

4. After dilution, the sample should be transferred to mica as fast as possible to avoid changes in binding behavior due to changed salt conditions. Typically adding a small volume of binding reaction to a premeasured volume of deposition buffer for dilution and immediately pipetting onto mica takes less than 20 sec.
5. The optimal buffer conditions, stoichiometry of DNA and protein as well as expected number of complexes are best extracted from previous biochemical characterizations. As mentioned above these will indicate starting conditions as the amounts and concentration of DNA and proteins that are optimal for SFM imaging may be different. Based on the initial SFM results it may be necessary to vary the stoichiometry of DNA and protein to obtain sufficient protein-DNA complexes for analysis or to minimize background of unbound proteins. This is best done in small steps, such as 2 fold changes in concentration of proteins or DNA, in order to improve the image data that can be obtained.
6. If not enough material is deposited on mica small variations in the dilution step that may affect cation

concentrations by 2-fold or less can also make a big difference.

7. The sharpness of the tips determines the resolution and detail of the images obtained. Although specialized tips with end radius of curvature considerably less than the usual 10 nm (some as small as 2 nm) are available they are still rather expensive. In our experience for most applications the added resolution and detail are minimal and do not justify the extra cost. We use uncoated tips. Even though coatings that increase reflectivity of the back of the cantilevers should not influence the size of the end of the tips, in our experience coated tips produced poorer resolution.
8. Many commercial SFMs produce images with maximally 512 x 512 pixels, independent of the scan size. Thus larger scan sizes produce data with lower resolution. Higher resolution is achieved with smaller scan size (scan of 1 μm x 1 μm , 1 pixel \sim 2 nm vs. scan of 4 μm x 4 μm , 1 pixel \sim 8 nm). Due to the size of the scanning tips, we typically use standard silicon non-contact tips with a radius of curvature of about 10 nm, resolution does not improve much by decreasing scan size below 1 μm x 1 μm .
9. When using a UV light wear UV safety glasses and avoid skin exposure.
10. To check if the mica glued to glass thickness is appropriate, deposit the fluorescent test object like the

FluoSpheres® (3pM solution in deposition buffer), and check if it is possible to focus on the surface in the fluorescent microscope. If it is not possible, more mica layers need to be cleaved off.

11. We have also tested quantum dots as markers for aligning fluorescence and topographic images. However due to the relatively large percentage of dark quantum dots in the preparations we have used it is not easy to unambiguously align the patterns from the topographic and fluorescence image. Thus in our experience quantum dots are not a robust marker for alignment.
12. Dual color fluorescent imaging is most beneficial when dealing with samples described in this chapter. Therefore, we recommend using TransFluoSpheres®, which are characterized by broad excitation spectrum. This allows collecting images at a wavelength visualizing both fiducials and sample as well as a wavelength for fiducials only, minimizing sample bleaching. in combination with biological sample coupled with fluorophore which excitation spectrum partly overlays with that of the fiducial sphere (for example TransFluoSpheres® (488/675) and Cy5 (650))

Bibliography

1. Janicijevic, A., Ristic, D. and Wyman, C. (2003) The molecular machines of DNA repair: scanning force microscopy analysis of their architecture. *J Microsc.* **212**, 264-72.

2. Dame, R.T., Wyman, C. and Goosen, N. (2003) Insights into the regulation of transcription by scanning force microscopy. *J Microsc.* **212**, 244-53.
3. Erie, D.A., Yang, G., Schultz, H. C. and Bustamante, C. (1994) DNA bending by Cro protein in specific and nonspecific complexes: implications for protein site recognition and specificity. *Science* **266**, 1562-6.
4. Ristic, D., Modesti, M., van der Heijden, T., van Noort, J., Dekker, C., Kanaar, R. and Wyman, C. (2005) Human Rad51 filaments on double- and single-stranded DNA: correlating regular and irregular forms with recombination function. *Nucleic Acids Res.* **33**, 3292-302.
5. Sánchez, H., Wyman, C. (2015) SFMetrics: an analysis tool for scanning force microscopy images of biomolecules. *BMC Bioinformatics* **16**, 1-9.
6. Bustamante, C., Vesenka, J., Tang, C. L., Rees, W., Guthold, M. and Keller, R. (1992) Circular DNA molecules imaged in air by scanning force microscopy. *Biochemistry* **31**, 22-6.
7. Vesenka, J., Guthold, M., Tang, C. L., Keller, D., Delaine, E. and Bustamante, C. (1992) Substrate preparation for reliable imaging of DNA molecules with the scanning force microscope. *Ultramicroscopy* **42-44**, 1243-9.
8. Hansma, H.G. and Laney, D.E. (1996) DNA binding to mica correlates with cationic radius: assay by atomic force microscopy. *Biophys J.* **70**, 1933-9.
9. Han, W., Lindsay, S. M., Dlakic, M. and Harrington, R. E. (1997) Kinked DNA. *Nature* **386**, 563.
10. Rivetti, C., Guthold, M. and Bustamante, C. (1996) Scanning force microscopy of DNA deposited onto mica: equilibration versus kinetic trapping studied by statistical polymer chain analysis. *J Mol Biol.* **264**, 919-32.

11. Beerens, N., Hoeijmakers, J. H., Kanaar, R., Vermeulen, W. and Wyman, C. (2005) The CSB protein actively wraps DNA. *J Biol Chem.* **280**, 4722-9.
12. Ratcliff, G.C. and Erie, D.A. (2001) A novel single-molecule study to determine protein--protein association constants. *J Am Chem Soc.* **123**, 5632-5.
13. Wyman, C., Rombel, I., North, A. K., Bustamante, C. and Kustu, S. (1997) Unusual oligomerization required for activity of NtrC, a bacterial enhancer-binding protein. *Science* **275**, 1658-61.
14. van der Linden, E., Sanchez, H., Kinoshita, E., Kanaar, R. and Wyman, C. (2009) RAD50 and NBS1 form a stable complex functional in DNA binding and tethering. *Nucleic Acids Res.* **37**, 1580-8.
15. Janicijevic, A., Sugawara, K., Shimizu, Y., Hanaoka, F., Wijgers, N., Djurica, M., Hoeijmakers, J. H. and Wyman, C. (2003) DNA bending by the human damage recognition complex XPC-HR23B. *DNA Repair (Amst).* **2**, 325-36.
16. Dame, R.T., van Mameren, J., Luijsterburg, M. S., Mysiak, M. E., Janicijevic, A., Pazdzior, G., van der Vliet, P. C., Wyman, C. and Wuite, G. J. (2005) Analysis of scanning force microscopy images of protein-induced DNA bending using simulations. *Nucleic Acids Res.* **33**, e68.
17. Sánchez, H., Kertokallio, A., van Rossum-Fikkert, S., Kanaar, R., Wyman, C. Combined optical and topographic imaging reveals different arrangements of human RAD54 with presynaptic and postsynaptic RAD51-DNA filaments. *PNAS* **110**, 11385-90.

Chapter 5

Architectural plasticity of human BRCA2-RAD51 complexes in DNA break repair

H. Sánchez¹, M. W. Paul^{2,4}, M. Grosbart¹, S. E. van Rossum-Fikkert¹, J. H. G. Lebbink¹, R. Kanaar^{1,3}, A. B. Houtsmuller^{2,4}, C. Wyman^{1,3}

Department of ¹Molecular Genetics, ²Pathology and ³Radiation Oncology, Erasmus University Medical Center, 3000 CA Rotterdam, The Netherlands
⁴Erasmus Optical Imaging Centre, Erasmus University Medical Center, 3000 CA Rotterdam, The Netherlands

Published in Nucleic Acids Research 2017 45(8), 4507-4518

ABSTRACT

The tumor suppressor BRCA2 is a large multifunctional protein mutated in 50-60 % of familial breast cancers. BRCA2 interacts with many partners and includes multiple regions with potentially disordered structure. In homology directed DNA repair BRCA2 delivers RAD51 to DNA resulting in removal of RPA and assembly of a RAD51 nucleoprotein filament. Dynamic rearrangements of BRCA2 likely drive this molecular hand-off initiating DNA strand exchange. We show human BRCA2 forms oligomers which can have an extended shape. Scanning force microscopy and quantitative single molecule fluorescence define the variety of BRCA2 complexes, reveal dramatic rearrangements upon RAD51 binding and the loading of RAD51 patches on single strand DNA. At sites of repair in cell nuclei, super-resolution microscopy shows BRCA2 and RAD51 arranged in largely separate locations. We identified dynamic structural transitions in BRCA2 complexes suggested to facilitate loading of RAD51 onto RPA coated single strand DNA and subsequent release of BRCA2.

INTRODUCTION

Many proteins with regulatory roles, such as tumor suppressors like BRCA2, have regions of undefined structure sometimes described as intrinsically disordered. Not surprisingly due to its size and potentially disordered regions, the structure of the BRCA2 protein has been challenging researchers for more than two decades. Structural models obtained by X-ray crystallography have been determined for isolated parts of BRCA2. The DNA binding domain co-purified with the DSS1 (deleted in split-hand/split foot syndrome) protein and the BRC4 repeat fused with RAD51 show details of the interfaces between parts of BRCA2 and elements of the DNA recombination machinery (1,2). Full length BRCA2 interacts with 5-6 RAD51 molecules (3) and facilitates formation of RAD51 filaments on DNA to catalyze the strand exchange step of homologous recombination (HR) (3-5). The arrangement of BRCA2 dimers associated with RAD51 and single stranded DNA (ssDNA) has been determined from electron microscopy (EM) image reconstruction (6). BRCA2 associates with several additional partners including PALB2 that mediates interaction with BRCA1 (7,8). BRCA2 function will require dynamic rearrangement of these molecular complexes facilitating partner exchange throughout the process of HR.

Here we investigate the global arrangement of BRCA2, changes upon RAD51 binding and loading on DNA using single-molecule visualization techniques in vivo and in vitro. Scanning force microscopy (SFM) reveals three-dimensional structures of biomolecules in vitro in conditions compatible with biochemical

activity. Combining SFM with total internal reflection fluorescence microscopy (TIRF-SFM) (9) allows the identification and quantification of proteins in heterogeneous mixtures like the BRCA2-RAD51 complexes addressed here. In cells BRCA2 moves in the nucleus as oligomeric clusters together with RAD51 (10). Using single molecule localization by direct stochastic optical reconstruction microscopy (dSTORM) (11), we determined the arrangement of BRCA2, RAD51, and RPA in microscopically discernable structures (foci) formed in the cell nucleus after DNA damage induction. Although diffusing together, as reported previously (10), BRCA2 and RAD51 were largely separated at sites of DNA repair, in patterns that changed over time. These results reveal biological functions of the BRCA2 protein machinery manifest as the structural transitions and dynamic interactions observed by single molecule imaging of biochemical samples and super resolution in cells.

RESULTS

Human BRCA2 associates as oligomers

To study the architectural arrangements of BRCA2 protein, we purified full-length human BRCA2 protein fused with two maltose binding proteins (MBP) as described (3). Visualization by SFM reveals the ensemble of BRCA2 complexes and their structural variation at nm resolution after deposition in conditions compatible with biochemical activity (10 mM HEPES [pH 8.2], 112 mM NaCl) and in the absence of fixation agents. A variety of BRCA2 oligomers were evident in SFM images, as previously observed EM images, from which prominent dimers were

compiled for structure building (6). BRCA2 complexes appeared as a heterogeneous population of branched elongated structures (Fig. 1A). SFM imaging provides an accurate representation of protein association in solution and distribution of multimeric states. (29,30) The volume and the skeleton length of individual BRCA2 protein complexes were used to estimate the number of monomers by a semi-automatic analysis routine (20). The distribution of BRCA2 complexes (Fig. 1B), based on volume and skeleton length, had two main populations containing dimers, the most prominent, (1.6 ± 0.4 [SD] subunits) and likely tetramers or pentamers (4.6 ± 2.8 [SD]) with a skeleton length of 179 ± 114 (SD) and 403 ± 132 (SD) nm, respectively. Isolated monomers of BRCA2 were up to 100 nm in length (example boxed in Fig. 1A). Inspection of the oligomeric complexes revealed an intriguing, recurrent protruding V-shape domain. One arm of the V was significantly shorter than the other (35 ± 5 [SD] nm vs 44 ± 6 [SD] nm from 235 analyzed complexes, $p < 0.05$ two-sample t-test). This arrangement could result from folding one BRCA2 monomer near the middle of the molecule, or an association between two different monomers in parallel or antiparallel orientation with a small offset. Antibodies recognizing the MBP tags increased the volume in the center of the complex but not to the protruding ends of the elongated structures (Fig. 1C). This suggests that some of the C-terminal DNA binding domains could be protruding (see discussion). The association of BRCA2 into multimers and the arrangement of the elongated domains was independent of MBP as the shape or distribution of forms was unchanged after removal of the tags (Fig. 1D).

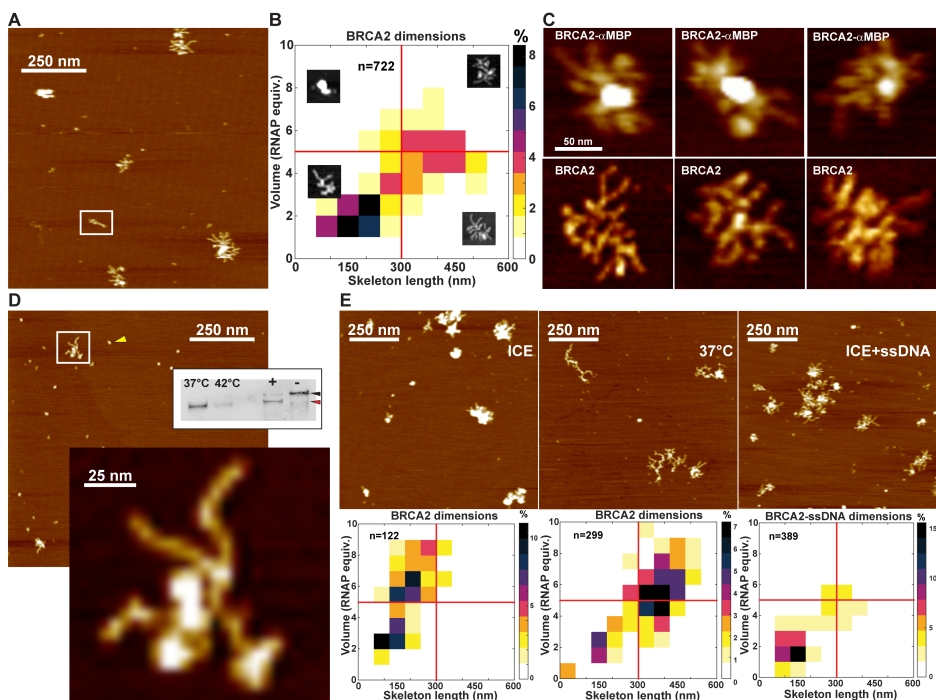


Fig 1. Human BRCA2 associates in high-order oligomers that are temperature sensitive. (A) Representative SFM image of 2xMBP-BRCA2. Several complexes from monomers (example boxed) to higher order oligomers showing a characteristic V-shape. Image size 1 x 1 mm. (B) Distribution of complexes based on volume and length. All individual complexes from images similar to those shown in panel A were measured and their distribution plotted in a two dimensional histogram with bin size 60 nm x 1 RNAP equiv. One RNAP (~450 kDa) equals approximately one 2xMBP-BRCA2 (470 kDa). Color indicates percentage of total complexes analyzed as shown in the scale bar. Red lines are guides to view bins. Inset images are zoomed examples (shown at different scales) representative of complexes in quadrants defined by the red lines. (C) SFM images of BRCA2 incubated with MBP antibody (upper row) compared to BRCA2 alone (bottom row). (D) The MBP tag does not influence BRCA2 oligomeric structure. SFM image of BRCA2 (white square and zoomed view below) after cleavage of the MBP tags by incubation with PreScission protease. The small round particles (yellow arrowhead) in the background are either protease or cleaved tag. Inset shows Western blot verification of tag cleavage. Samples are: lane "+" after and "-" before protease incubation. Western blot control and size comparison are: lane "37°C" extracts from human BRO-derived melanoma cells showing endogenous BRCA2 which is as expected degraded as shown in lane "42°C". Black arrowhead: MBP tagged BRCA2; red arrowhead: untagged BRCA2. Color intensity indicates height (from 0 to 1.5 nm, dark to light). (E) SFM images of BRCA2 and quantification of size distribution, in conditions as labeled; incubation on ice, *in situ* warming to 37°C, or incubation with 66 nt ssDNA.

The oligomeric and elongated appearance of BRCA2 suggested structural variability characteristic of proteins with intrinsically disordered regions. The structure of proteins with intrinsically disordered regions can vary dramatically due to relatively subtle change in conditions(31). We tested the effect of incubation at different temperatures and addition ssDNA on BRCA2 complex architecture. Notably, the protein complexes assumed a largely globular shape if the sample was deposited from incubation on ice (Fig. 1E). This transition was readily reversible as *in situ* warming to 37°C of the sample on mica, or incubation on ice with linear ssDNA oligonucleotides before deposition for imaging, exposed the branched structures. This suggests that self-association among BRCA2 domains, resulting in globular or extended structures, is readily reversible and dynamicconsistent with the behavior of intrinsically disordered regions.

Structural plasticity of Human BRCA2

Primary sequence analysis predicts the frequent occurrence of segments along the BRCA2 amino acid sequence that can be described as intrinsic disordered regions (IDRs) (32,33) (Fig. S1A-C). IDRs are often characteristic of flexible linkers involved in the assembly of macromolecular complexes (33,34). We compared the dimensions of the BRCA2 elongated structures with double-stranded DNA (dsDNA) and the coiled-coils of the RAD50 component of the human MRN complex (35) visualized by SFM. Full width at high maximum (FWHM) of extended BRCA2 was similar to dsDNA or RAD50 coiled coils indicating a true width of 2

or 3 nm after accounting for SFM tip distortions (27). The apparent persistence length (PL) of elongated BRCA2 (17 ± 1 [SD] nm) was less than dsDNA (56 ± 6 [SD] nm) or RAD50 coiled coils (30 nm (36)), showing the flexibility of BRCA2 as a polymer (Fig. S1D-E).

BRCA2-RAD51 complexes are regular elongated structures

BRCA2 functions in HR associated with RAD51. To describe the arrangement and variation in BRCA2-RAD51 complexes we used TIRF-SFM (9) and compared fluorescent RAD51, hereafter RAD51, (15,16,25) (Fig. S2A), BRCA2 and a mixture of the two, in the absence of DNA and nucleotide cofactor. BRCA2-RAD51 complexes were formed at a molar ratio of 1:24 and visualized by TIRF-SFM, after fixation with glutaraldehyde (Fig. 2A).

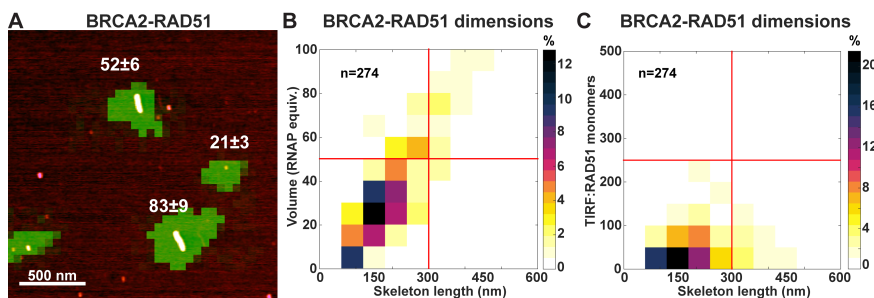


Fig 2. BRCA2-RAD51 complexes organize into filament-like structures (A) RAD51-BRCA2 complexes visualized by combined SFM and fluorescence microscopy (TIRF-SFM). SFM scan overlaid with fluorescence image (green) of BRCA2 incubated with fluorescent RAD51 and cross-linked with glutaraldehyde. Numbers indicate RAD51 monomers \pm SD for the complexes shown. (B) Volume and length of all complexes were measured and their distribution plotted in a two dimensional histogram with bin size 60 nm x 10 RNAP equiv. Color indicates percentage of total of complexes analyzed (n=274) as shown in the scale bar. (C) Distribution of RAD51-BRCA2 complexes based on number of fluorophores (RAD51) and length. All complexes were measured and their distribution plotted in a two dimensional histogram with bin size 60 nm x 50 RAD51 monomers. Color indicates percentage of total complexes analyzed (n=274) as shown in the scale bar.

In the absence of fixation limited BRCA2-RAD51 complexes were observed by SFM, indicating interactions that are dynamic or transient in these conditions.

The co-incubation resulted in protein complexes with an unanticipated filament-like appearance not apparent for either RAD51 or BRCA2 separately prepared with the same fixation protocol (Fig. S3). Analysis of TIRF-SFM images revealed filament like structures ranging from 30 nm to more than 500 nm long, and a main population with a volume equivalent to 20-30 RNAP equiv., including 50-100 RAD51 monomers based on fluorescence (Fig. 2B-C). The number of RAD51 monomers based on fluorescence intensity in BRCA2-RAD51 complexes was lower than in RAD51-DNA complexes with the same volume (see Materials and Methods and Fig. S2B for details of the quantification). The extra volume in the BRCA2-RAD51 complexes in the main population was equivalent to 12 to 26 BRCA2 monomers. Similar structures were observed using unlabeled RAD51 protein indicating that the regular elongated forms are not due to the fluorescent label (Fig. S4A). The elongated structures had an interrupted height profile repeating with 51 ± 67 (SD) nm periodicity along their contour, indicating the end-to-end association of smaller complexes. For comparison dimeric BRCA2-RAD51 complexes reconstructed from EM images are roughly oval at 26.5 nm X 16.5 nm (6). Additionally, BRCA2-RAD51 complexes were wider (FWHM= 17.8 ± 6 [SD] nm) and stiffer (PL= 698 ± 338 [SD] nm) than RAD51-dsDNA (FWHM= 8.75 ± 3.1 [SD] nm, PL= 91 ± 11 [SD] nm) filaments (Fig. S1D). Together the volume and fluorescence analysis of complexes observed by SFM indicated they consist of 4-5 RAD51 per BRCA2,

similar to the stoichiometry of on average 4.5 to maximal 6 RAD51 per BRCA2 determined by biochemical methods(3). The elongated structures indicated specific interactions, captured during fixation, resulting regular, ordered multimers with a unit length of about 50 nm.

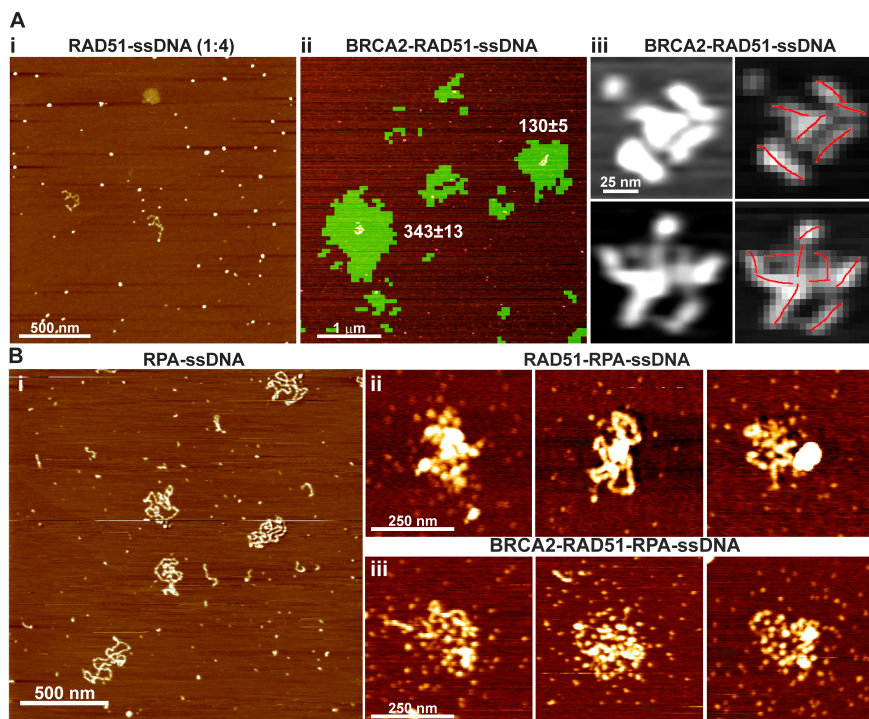


Fig 3. BRCA2 oligomers mediate RAD51 hand-off to DNA. (A) Loading of RAD51 on DNA (i) SFM image of Φ X174 Virion ssDNA incubated with RAD51 and cross-linked with glutaraldehyde. Color intensity indicates height (from 0 to 3 nm). (ii) TIRF-SFM image of Φ X174 Virion ssDNA incubated with BRCA2-RAD51 and cross-linked with glutaraldehyde. (iii) Examples of multiple BRCA2-RAD51 complexes bound to Φ X174 Virion ssDNA from images similar to panel B. The patch length (red line) was analyzed with 'SFMetrics' (B) Loading of RAD51 on RPA coated ssDNA (i) SFM image of Φ X174 Virion ssDNA incubated with RPA, visualized without fixation. (ii) RAD51-RPA-ssDNA complexes visualized by SFM without fixation. (iii) BRCA2-RAD51-RPA-ssDNA complexes visualized by SFM without fixation.

Multiple patches of BRCA2-RAD51 load on ssDNA

The organization of RAD51 by BRCA2, in the absence of DNA, into higher order structures larger than those previously described (6), suggested the possibility that RAD51 is loaded onto ssDNA in a manner distinct from the nucleation and growth mechanism commonly described (6,16,25,37,38). To observe the effect of BRCA2 on RAD51-DNA loading, we incubated ssDNA (Φ X174 Virion) with fluorescent RAD51 alone (Fig. 3A) or with BRCA2-RAD51 (Fig. 3B) in the presence of ATP and Mg^{2+} , necessary cofactors for dynamic filament assembly. We used sub-saturating concentrations of RAD51 (1 monomer per 4 nt) to highlight the possible mediator effect of BRCA2. DNA-protein complexes with fluorescent RAD51 were visualized by TIRF-SFM after fixation with glutaraldehyde (Fig. 3A) and analyzed for volume and fluorescence intensity (Fig. S4B). In these conditions, RAD51 alone did not stably bind to ssDNA and the DNA appeared mainly free of protein (31 out of 37 DNA molecules, Fig. 3Ai) with unbound RAD51 in the background. Incubation of BRCA2 with RAD51 resulted in multiple RAD51 patches on all ssDNA molecules (Fig. 3Aii, iii). The amount of fluorescent RAD51 in these DNA-protein complexes was determined by TIRF-SFM (Fig. S4B). Volume analysis indicated that BRCA2 was included, in addition to RAD51 detected by fluorescence, in the DNA bound structures. For example, in the population with a volume equivalent to 30-40 BRCA2 monomers. and 143 ± 101 ($n=17$) RAD51 monomers detected by fluorescence would include additional volume equivalent to 18-28 BRCA2 monomers. Both RAD51 and BRCA2 can bind ssDNA and either or both proteins could be bound in the complexes observed here.

This striking association of RAD51 with DNA only in the presence of BRCA2 reflects the importance of the BRCA2-RAD51 complex for initiating DNA strand exchange.

The DNA-protein complexes observed by SFM, including RAD51 and BRCA2, had patches with regular features (Fig. 3Aiii). The average length of the patches, delineated by a decrease in the cross-section height or a kink in the contour was 71 ± 24 (SD) nm (from 405 patches in 79 DNA-protein complexes). The dimensions of the DNA bound complexes (FWHM= 13.9 ± 6.7 [SD] nm, height= 3.8 ± 1.5 [SD] nm) were different from both the RAD51-dsDNA filaments (FWHM= 8.75 ± 3 [SD] nm, height= 2.5 ± 0.6 [SD] nm, n= 101) and the BRCA2-RAD51 complexes (FWHM= 17.8 ± 6 [SD] nm, height= 2.4 ± 0.9 [SD] nm, n=123). These dimensions, together with volume measurements and number of RAD51 estimated from fluorescence, indicate that BRCA2 not only facilitates RAD51 loading on ssDNA but that interaction with ssDNA also induced partial release of BRCA2.

The apparent release of BRCA2 was more evident when the DNA protein complexes were visualized in the absence of fixation (Fig. 3B). Pre-incubation of the ssDNA with RPA (visualized in Fig. 3B without fixation) resulted in limited loading of RAD51 in the absence of BRCA2 (Fig. 3Bii, iii and Fig. S4E-F). Interestingly DNA electrophoretic mobility shift experiments suggest that BRCA2 could rearrange RPA on DNA (Fig. S4D, compare lanes 2 and 6). As expected, BRCA2 enhanced RAD51 loading onto RPA coated ssDNA (Fig. 3B and Fig. S4E-H). About 2 fold more RAD51 was loaded onto ssDNA preincubated with RPA in the presence of

BRCA2, 137 ± 42 (SD) vs 69 ± 42 (SD) RAD51 monomers per DNA molecule (Fig. S3F, H). The resulting protein DNA complexes also appeared very different from those formed on ssDNA alone. BRCA2 loaded RAD51 was more evenly distributed loaded onto RPA-ssDNA, not in defined patches (compare Fig. 3Aiii to Fig. 3Biii). In addition, based on fluorescence intensity estimates of RAD51 monomers and volume, BRCA2 was largely absent from these complexes (Fig. 3B). The architectural rearrangements and molecular hand-off of RAD51 from BRCA2 to DNA was apparently enhanced by interaction with RPA.

Distinct arrangements of BRCA2 and RAD51 at DNA damage sites in the cell nucleus

With conventional microscopy resolution RAD51 and BRCA2 are reported to co-localize at sites of DNA damage (10,19,39). However our in vitro results suggest that BRCA2 and RAD51 may separate once RAD51 is loaded onto DNA. To determine the arrangement of BRCA2, RAD51 and RPA in the nucleus of U2OS cells after DNA double-strand break induction, we used single molecule localization microscopy (dSTORM). This provides localization accuracy relevant to molecular dimensions. Typical accumulations of RAD51, often referred to as DNA repair foci, were clearly visible by fluorescence confocal microscopy (Fig. S5A and S6A). Localization of individual molecules by dSTORM revealed that RAD51 focal accumulations were frequently elongated some more than 400 nm long (Figure 4A-C). Subsequent two-color dSTORM imaging revealed that BRCA2 localizations were relatively diffuse with several clusters around the elongated

Figure 4

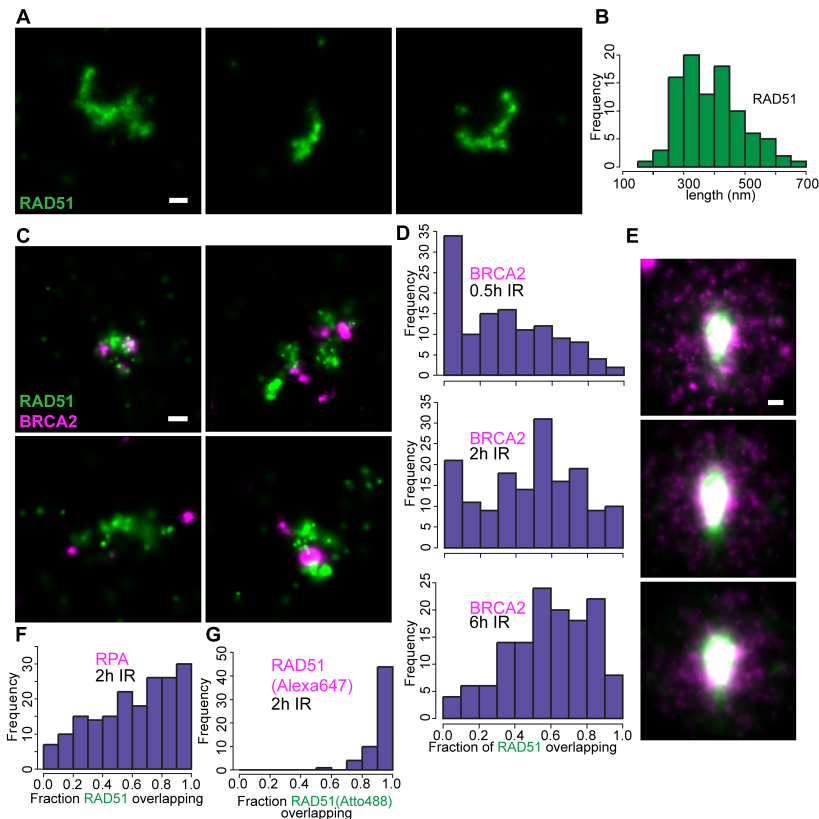


Fig 4. BRCA2 and RAD51 have differential localization patterns in ionizing radiation induced foci. (A) Super resolution (dSTORM) images of RAD51 foci detected by immunofluorescence in U2OS cells fixed 2 hours after treated with 6 Gy of ionizing radiation (IR). (B) Distribution of the RAD51 structures based on length of the major axis. (C) Examples of repair foci RAD51 (green, Atto488 conjugated secondary antibody) and BRCA2 (magenta, Alexa647 conjugated secondary antibody) acquired with dSTORM. (D) RAD51 localizations that overlap with BRCA2 were determined for every ROI (focus) and the distribution of foci based on fraction overlapping signals plotted in a histogram (Fig.S4D). (n= 15, 11, 11 cells respectively). (E) Summed projections of dSTORM images of repair foci aligned and rotated based on the RAD51 (green) signal, show the BRCA2 (magenta) signal spread around RAD51. Images represent the same time points as in D. (F) The distribution of foci based on fraction of overlapping signals for RAD51 and RPA. (G) As a control cells were stained for RAD51 and two different secondary antibodies showing, as expected, all foci with near complete overlap of signals in this analysis. Scale bars indicate 100 nm.

RAD51 accumulations (Fig. 4C). As expected from their relative nuclear concentration (10) and likely interaction stoichiometry (3) most BRCA2 clusters were smaller than RAD51 clusters, including fewer localizations.

The relative location of BRCA2 and RAD51 was analyzed quantitatively (Fig. 4C-E, S5). Each RAD51 focus was oriented with its long axis vertical with the most intense signal up and an overlay of all foci was compiled (Fig. 4E, see also Fig. S5E). BRCA2 localizations (magenta) were not concentrated in the center but clearly spread around the central RAD51 cluster. The spatial overlap between two different protein species in single molecule localization microscopy can be studied by quantifying the abundance of localizations of the one protein near localizations of the other (40). Here the fraction of RAD51 localizations that overlapped with the BRCA2 was quantified by counting the number of BRCA2 localizations within a radius of 50 nm from every RAD51 localization (Fig. S5C-D). From this a fraction of overlapping RAD51 localizations can be calculated for every DSB focus. For each time point after DNA damage induction by ionizing irradiation (IR) we classified foci based on the fraction of overlapping BRCA2 and RAD51 they included and plotted their distribution in histograms (Fig. 4D). As can be seen at all 3 time points after IR (0.5, 2 and 6 hours) most foci included at least some RAD51 signal that was not overlapping with the BRCA2 signal (cumulative total of all histogram bars except fraction overlapping = 1.0). At 0.5 hr after inducing breaks, with 6 Gy ionizing irradiation, the most frequent class of foci had no overlap in BRCA2 and RAD51 (30%). This pattern changed in time after

irradiation. The proportion of foci with most BRCA2 and RAD51 overlapping (histogram bin at 0.8 to 1.0 fraction overlapping) increased at 2 and 6 hours. In comparison similar two-color dSTORM imaging for RPA and RAD51 showed a different distribution within foci with varying overlap (Fig. 4F, S6C-D). As a control for our imaging, analysis methods and comparison of the histogram presentations, we stained RAD51 with a mixture of secondary antibody labeled with two different dyes which showed almost complete overlap in signal, as expected (Fig. 4G). We have reliably quantified the relative distribution of different proteins in DNA damage include foci, where BRCA2 and RPA were often separated from RAD51. Changes in these patterns can be correlated to dynamic interactions during the DNA repair process and used to test biological function.

DISCUSSION

We have determined the relative arrangement of essential DNA break repair proteins, BRCA2, RAD51 and RPA, in complexes on DNA from purified components and in the cell nucleus. These data identify steps in the molecular choreography of assembling functional repair machinery where it is needed. BRCA2 alone displays variable structures, depending on mild changes in conditions and to binding partners, features characteristic of proteins including intrinsically disordered regions. Binding to RAD51 induces dramatic reorganization into a regular complex capable of multimerization and active in loading RAD51 onto ssDNA. We observe at least partially release of RAD51 from BRCA2 to DNA in the presence of RPA. Super-resolution imaging in cell

nuclei also indicated that BRCA2 and RAD51 separate at the sites of DNA damage. BRCA2 however remained in the area of the DNA damage, as also observed in foci with standard microscopy, possibly indicating a role late in DNA break repair.

BRCA2 is responsible for delivering RAD51 to sites of DNA breaks in the nucleus. As such it has to scan nuclear volume for sites where its action is needed. At an approximate concentration of 10 nM, in the range estimated for BRCA2 and several other DNA repair proteins in the cell nucleus (10,41,42), each protein would have to search volumes equivalent to a sphere with radius of about 300 nm. Our direct SFM visualization shows that BRCA2 oligomers can adopt an extended configuration spanning up to several hundred nanometers. BRCA2 oligomers in this conformation effectively tether multiple C-terminal DNA binding domains. In addition association with multiple RAD51 monomers effectively provides additional DNA binding domains tethered in close proximity. These features would facilitate an effective search for DNA damage sites in the nucleus via frequent non-specific interactions (10). BRCA2 regions characterized as intrinsically disordered could allow fluctuating inter- and intra-molecular associations with low entropy cost promoting conformational flexibility and functional exchange of binding partners (43). Intrinsically disordered regions of BRCA2 may also facilitate its accumulation via interaction with poly(ADP-ribose) which has been suggested to seed the accumulation of intrinsically disordered proteins at the sites of DNA breaks (44).

In vitro single-molecule experiments have extensively studied the DNA binding behavior of RAD51 (16,25,38,45) where filament formation is described as a two-steps process involving nucleation and growth. Our previous live cell imaging data (10) indicated that most, if not all, nuclear RAD51 is traveling in complex with BRCA2 suggested that filament formation *in vivo* may be dominated by these complexes and not free RAD51. We show here that BRCA2 transforms RAD51 into a stiff multimeric DNA binding complex offering potential mechanistic advantages. The BRCA2-RAD51 complex is essentially a multimeric DNA binding complex that would facilitate loading of RAD51 onto ssDNA in competition with proteins like RPA (3-5). Although RAD51 *in vitro* can displace RPA on ssDNA, we note that *in vivo* the relative concentrations of these proteins likely make this ineffective. Modification of RAD51 by phosphorylation mitigates the need for mediators to replace RPA *in vitro* (46). However if post-translational modification not occur effectively after filament formation, additional help to replace RPA initiating filament formation would be needed. Even a single BRCA2 comprising three OB folds and multiple RAD51s forms a multimeric ssDNA binding complex that we suggest could effectively compete away ssDNA-bound RPA without the need of direct interaction between BRCA2 and RPA (47). The complete dependence on BRCA2 for loading RAD51 on ssDNA in the absence of RPA, observed here with limited RAD51 (Fig. 3), possibly suggests an additional role for this complex unfolding ssDNA secondary structures. In combination BRCA2-RAD51 complexes serve a series of functions: catching RPA-ssDNA, using the BRCA2 DNA-binding domains to promote replacement of RPA for RAD51, and delivering multiple RAD51

molecules for efficient filament formation (Fig. S4G). We anticipate that the intrinsically disordered regions of BRCA2 are essential for dynamic structural rearrangements that catalyze molecular hand-off events in the cell nucleus.

Although all detectable RAD51 diffuses in the cell nucleus in complex with BRCA2 (10) release to load RAD51 onto DNA at break sites is expected *in vivo*. The accumulation of DNA break repair proteins at sites of damage in the cell nucleus has been extensively studied by fluorescence microscopy. After DNA breaks are induced, proteins needed to perform DNA strand exchange in HR reorganize by rapidly associating in clusters, or repair foci, in a BRCA2 dependent manner (48). The limited resolution of standard light microscopy is unable to resolve the relative arrangement of RAD51, BRCA2, and other repair proteins in foci that all appear as (overlapping) spherical clusters. The increased resolution of dSTORM, acquired by separating the fluorescence signal of individual molecules, enables accurate localization of single molecules in densely labeled samples. We present here super resolution images of RAD51, BRCA2 and RPA detected by immunofluorescence in DNA damage induced foci. We observe RAD51 arranged in an elongated patterns, up to several 100 nm long. Though there is no independent evidence that these structures are RAD51-DNA filaments, this elongated appearance is expected for a protein that coats DNA, in a filament. The length of the elongated RAD51 structures is reasonable given that 450 nm (approximate median of the distribution Fig. 4B) is equivalent to a perfect RAD51-DNA filament including 900 nt of DNA and 300 RAD51 molecules (note number of localizations does not equal

number of molecules in our super-resolution methods employing indirect immunofluorescence staining and detection of fluorophore blinking).

Our immunofluorescence experiments show distinct patterns of BRCA2, RAD51 and RPA in DNA damage induced foci. The RAD51 primary antibody used is polyclonal and will recognize many epitopes or presentations of RAD51. The monoclonal BRCA2 antibody recognizes one of the RAD51 binding BRC repeats. It is possible this epitope is not as efficiently recognized if RAD51 is bound to some of the BRC repeats in BRCA2. Even if all BRCA2 is not detected all RAD51 should be detected whereby the distinct pattern of RAD51 and BRCA2 reveal different dynamic arrangements at DNA breaks sites. At the earliest times analyzed after damage (0.5 hr) BRCA2 is in small clusters separated from most of an elongated, much larger RAD51 cluster (Fig. 4D-E) consistent with a molecular hand-off upon loading RAD51 onto DNA. However, hand-off does not lead to BRCA2 absence from the foci, as BRCA2 clusters are observed also at later stages (2 and 6 hours after damage induction). This suggests a role for BRCA2 beyond delivering RAD51 and that BRCA2 and RAD51 may re-associate before leaving the site of DNA damage.

In our super resolution images the RPA signal is also largely separated from RAD51 as expected if RAD51 replaces RPA bound to DNA (Fig. 4F and S6C). Interestingly we observe both BRCA2 and RPA are present at the RAD51 damage induced foci at all three time points after damage induction. The retention of RPA at DSBs has been observed in yeast by CHIP, both at the DNA break site

and at repair donor locus (49). We observe RPA in dense clusters, with large number of localizations, in contrast to RAD51, which is present in an extended pattern. As is the case for the DNA repair foci observed with conventional microscopy, there is no proof that the foci presented here with dSTORM represent DNA bound proteins; however their patterns are consistent with the different roles of these proteins and the structures they are expected to form on DNA.

The relative patterns of BRCA2 and RAD51 change over time after DNA damage (50). Here we observe specifically that localization of BRCA2 and RAD51 within 50 nm of each other is more common at later times, 2 and 6 hr after damage induction. At the dose of DNA damage inflicted, some of the damage is irreparable and some cells will eventually die. Thus the later time points may represent repeated attempts to repair or stalled repair. The increased association of BRCA2 and RAD51 in time may represent additional BRCA2 functions, stabilizing RAD51 association with broken DNA or more intriguingly an active role in removal of RAD51 after repair. Distinguishing these possibilities awaits further analysis and the ability to follow a repair event over time. However the ability to define and quantify relative protein arrangements demonstrates the power of our analysis to determine function in cells during the DNA repair process. We propose that the malleable nature of BRCA2, including multiple intrinsically disordered domains, facilitate its functions to locate break sites in nuclear space, exchange RPA for RAD51 on ssDNA and deliver RAD51 as multimeric complexes for efficient assembly of strand exchange machinery.

MATERIALS AND METHODS

Proteins, DNAs, and bulk assays

BRCA2 tagged with two tandem repeats of the maltose binding protein (MBP) at the N-terminus was purified after transient transfection of human embryonic kidney 293T cells with phCMV1 plasmid containing the full length BRCA2 gene (kindly provided by S. Kowalczykowski) as described (3). Aliquots of BRCA2 with concentration range 50-100 ng/ml (as estimated using extinction coefficient at 280 nm of $365,160 \text{ M}^{-1} \text{ cm}^{-1}$) were stored at -80°C in elution buffer from the final purification column (50 mM HEPES [pH 8.2], 10% glycerol, 450 mM NaCl, 1 mM DTT, 0.5 mM EDTA). Protein identity was confirmed by mass spectrometry (ErasmusMC Proteomics) and Western blotting (see below). BRCA2 protein was purified by three different persons in the laboratory producing four distinct stocks used for the SFM characterization. Prediction of intrinsic disordered regions in primary sequence was done using two different prediction web tools: IUPred (<http://iupred.enzim.hu>) (12,13) and DisEMBLTM (<http://dis.embl.de>).

Human RAD51 and RPA were expressed in *Escherichia coli* and purified as described (14,15). RAD51 with a single cysteine exposed to the solvent was Alexa Fluor 488 (Invitrogen) labeled (degree of labeling 1.3) using maleimide chemistry and checked for biochemical activities (ATPase and D-loop formation) as described (15,16).

E. coli RNA polymerase (Boehringer Mannheim) was used as SFM size calibration standard.

Linear ssDNA (1000 bases) molecules were generated by PCR amplification of pBluescript DNA and degradation of one strand with lambda exonuclease I as described (9). Circular ssDNA (5386 bases) ΦX174 Virion DNA was purchased from New England Biolabs. Oligonucleotide sequence of the 66-mer was as described (17).

Electrophoretic mobility shift assays (EMSAs) were done by mixing Cy5-labeled 90 nt ssDNA (18) (90 nM in nt) with or without RPA (60 nM) in reaction buffer (50 mM Tris-HCl [pH 7.5], 1 mM DTT, 60 mM KCl, 2 mM CaCl₂, and 1 mM ATP) for 10 min at 37°C in a final volume of 10 µl. Then RAD51 (300 nM), BRCA2 (20 nM), a mixture of both, or protein storage buffer was added and incubation continued for additional 20 min. The reactions products were separated on a 5% non-denaturing polyacrylamide gel running in 0.5× TB buffer at 4°C. Gels were analyzed using a Typhoon Trio scanner exciting the dye-coupled DNA.

MBP cleavage, and immunodetection of BRCA2

Purified 2xMBP-BRCA2 aliquots (~1 mg) were incubated with 2 units of PreScission protease (GE Healthcare) or with protease storage buffer for 1 hr at 4°C. Reaction mixture was split for immunodetection and SFM imaging. Samples were run on a 3-8% NuPAGE Tris-Acetate gel (LifeTechnologies), which was blotted on PVDF membrane (Immobilon) for 2 hrs at 4°C at 300 mA. After 1 hr blocking with 5% milk, the membrane was incubated with the

primary antibody Anti-BRCA2 OP95, Calbiochem, 1:2000 dilution in 1% milk and 0.1% PBS buffer) overnight at 4°C. The secondary antibody (Anti-mouse HRP, Abcam, 1:2000 dilution in 1% milk and 0.1% PBS buffer) was incubated for 2 hrs at room temperature. The membrane was exposed with ECL (Pierce) and imaged with an Uvitec Alliance 2.7 instrument (Cambridge, UK). Cellular extracts from human BRO-derived melanoma cells treated for hyperthermia (a gift from H. Odijk) were used as molecular weight markers for endogenous BRCA2 identification (19). The protease reaction mixture was diluted ten times in BRCA2 storage buffer for SFM visualization in order to reduce background from protease molecules.

SFM microscopic analysis of protein, and protein-DNA complexes

Samples were imaged in air at room temperature by tapping mode SFM using a Nanoscope III or Nanoscope IV (Digital Instruments). Silicon Pointprobe tips were type NHC-W, resonance frequency 310-372 kHz, from Nanosensors supplied by Veeco Instruments, Europe. Images were collected at 2 µm x 2 µm, 512 X 512 pixels, and processed only by flattening to remove background slope.

BRCA2 protein was diluted four times from the frozen stock with HEPES buffer (final composition: 10-25 ng BRCA2, 22 mM HEPES [pH 8.2], 2.5% glycerol, 112 mM NaCl, 0.25 mM DTT, 0.12 mM EDTA). After incubation at 37°C or on ice, as described in the text, 50 mM spermidine was added as binding agent and the protein deposited on freshly cleaved mica for 1 min, washed 3 times with MilliQ water and dried with filtered air.

The volume, length, and contour profile (width and height) measurements were done using 'SFMetrics' software tool as described (20). Measurements are reported together with associated standard deviations. The volume of each complex was normalized to E. coli RNA polymerase (RNAP) with a known molecular weight of 450 kDa that appeared in SFM as regular round objects. In order to characterize the shape of the complex the minimum length occupied by each complex, defined as skeleton length, was measured. The contours of V-shape domains of BRCA2 oligomers were traced manually.

Purified 2xMBP-BRCA2 protein (~1 mg) was incubated with Anti-Maltose Binding Protein antibody (a-MBP ab9084, Abcam) (250 ng) for 50 min at 37°C. Unbound antibody was removed by size exclusion chromatography (Sephacrose CL-4B, Sigma). Fractions containing BRCA2-a-MBP complexes were visualized by SFM.

Apparent persistence length (P_l) was obtained assuming a worm-like chain model for semi-flexible polymers. In two dimensions (2D), the mean square of the end-to-end distance (R) can be written as a function of the contour length (C_l) and P_l :

$$\langle R^2 \rangle_{2D} = 4P_l C_l \left(1 - \frac{2P_l}{C_l} \left(1 - e^{-\frac{C_l}{2P_l}} \right) \right)$$

L measurements and P calculation were performed using 'Easyworm' software tool (21) from more than 100 complexes in each case.

Automatic peak-to-peak analysis was performed using 'peakfinder.m' MATLAB script (by Nathanael C. Yoder) and returned peaks at local maxima that were at least a half standard deviation of the sample above surrounding data and larger than the mean of the sample.

For in situ warming of BRCA2, mica containing the sample was placed on a Dry Block Thermostat UBD (Grant Instruments) at 37°C for 30 min.

Interactions between protein and DNA were performed in solution at the indicated concentrations before the sample was deposited on mica. When indicated sample was cross-linked with glutaraldehyde (0.12 %) for 5 min at 37°C, quenched with 50 mM Tris-HCl [pH 7.5], and deposited in the presence of 10 mM Mg^{2+} . BRCA2-RAD51 elongated complex were selected based on eccentricity above 0.7 (ratio of the distance between the foci of the ellipse modelling the protein complex and its major axis length).

Sample preparation for TIRF-SFM visualization

Nucleoprotein filaments were formed by incubating linear ssDNA (1000 bases, 1.5 μ M) molecules with RAD51 (1 μ M) in binding buffer (10 mM HEPES[pH 8.2], 60 mM KCl, 2 mM $CaCl_2$, 1 mM DTT, and 1 mM ATP) (9,15). BRCA2-RAD51 complexes were formed by incubating an excess of RAD51 (0.6 mM) with BRCA2 (0.025 mM) in buffer containing 12.5 mM HEPES [pH 8.2], 4 mM Tris-HCl [pH 7.5], 3.25% glycerol, 125 mM NaCl, 65 mM KCl, 0.4 mM DTT, 0.17 mM EDTA, taking into account contributions from

all protein storage buffers. This reaction buffer composition was kept constant in all subsequent incubations unless noted. After 30 min at 37°C, glutaraldehyde was added to a final concentration of 0.12%, followed by additional 5 min incubation, and quenched by adding Tris-HCl to a final concentration of 50 mM [pH 7.5]. The sample was diluted (from 1:20 to 1:200 in order to reduce fluorescent signal due to free RAD51) in deposition buffer (10 mM HEPES-KCl [pH 8.2], 10 mM Mg₂Cl) and deposited with 3 pM TransFluospheres on cleaved mica mounted on glass cover slips, as previously described (9,22,23). After one minute the mica surface was rinsed with MilliQ water and dried with filtered air. RAD51 loading onto DNA was also analyzed in ATP hydrolysis conditions by supplementing the filament formation reaction described above with 1 mM ATP and 2 mM Mg₂Cl (ATPase-Buffer). Circular ssDNA ΦX174 Virion DNA (2.4 mM in nt) was incubated with RPA (0.1 mM). When indicated RPA-DNA complexes were incubated with preassembled BRCA2 (0.025 mM)-RAD51 (0.6 mM) in ATPase-Buffer. Glutaraldehyde was added as indicated. Concentrations indicate final amount in the reactions. Samples were deposited for TIRF-SFM as described above and observed with a Nikon TE 2000U microscope as described (9). Scanning force microscopy was done with a NanoWizard®II scanner (JPK instruments) mounted on the fluorescence microscope (9,22,23). Correlation of fluorescence and topographic images was done with the JPK DirectOverlay™ software tool. Data analysis was performed with custom-made software written in MATLAB and available in <http://cluster15.erasmusmc.nl/TIRF-SFM-scripts>.

Quantification of fluorescent RAD51 monomers

The number of RAD51 monomers per complex was determined by quantification of fluorescence intensity by using the stepwise photobleaching of single fluorophores. In order to compensate for intensity variations due to differences in sample thickness, or constraints in the degrees of freedom of the fluorophore because of sample deposition on mica, a global estimation approach was used. Intensity of one fluorophore was defined per each field of view (FOV) after averaging the intensity steps in every region of interest (ROI) as follows; sequential frames (exposure time of 300 ms) from the same FOV were acquired until photobleaching of almost all fluorophores. This stack of images was used to make a maximum intensity projection (Figure S2Ai) that creates an output image where each pixel contains the maximum value over all images in the stack at the particular pixel location (Figure S2Aii). From each ROI the intensity trace over time was extracted from a sequence of 300 frames where only a few fluorophores remained blinking until complete bleaching. Intensity variations were estimated by a step fitting algorithm (9,24,25) (Figure S2Aiii, iv). Step sizes smaller than background and bigger than two times the median value were discarded from the analysis. The reported number of fluorophores is the maximum intensity in the ROI divided by the mean of average step sizes for each ROI (Figure S2Av, vi).

Stable nucleoprotein filaments with defined components were formed with RAD51 and 1000 nt ssDNA in the presence of ATP-Ca²⁺ (Fig. S2Bi). A perfectly formed RAD51 filament on this length DNA would consist of a maximum of 333 monomers (one RAD51 per 3 nucleotides) with a maximum length of 400 nm. Two-

dimensional histograms in Fig. S2Bii and iii show the varied population of filaments formed in these conditions (partially covered, fully covered, and association of more than one filament) as previously observed (9,15,26). For example, in the filaments grouped between 360 and 420 nm length, the main population (5 % of the total) had a volume of 30 to 40 RNAP equiv., representing 365 to 489 RAD51 monomers, and 200 to 300 fluorophores as estimated by optical microscopy. Taking into account that volume can be overestimated due to tip distortions (27), and fluorescence intensity underestimated because of surface-fluorophore interaction (9), our data is in reasonable agreement with the expected 333 RAD51 monomers per fully covered DNA molecule showing the consistency of our TIRF-SFM approach for protein quantification.

Cell culture (for super resolution microscopy)

U2OS cells were cultured in phenol-red free DMEM (Lonza) medium supplemented with 10% FCS, L-glutamine and Pen-Strep. For microscopy experiments cells were cultured on 24 mm round coverslips of 1.5 thickness ($170 \pm 5 \mu\text{m}$). Cells were irradiated the next day with 6 Gy of ionizing radiation using a ^{137}Cs source.

Immunostaining and sample preparation

After washing with PBS cells were fixed at the indicated time points with 2% PFA in PBS (Lonza) for 20 minutes, permeabilized with PBS with 0.1% Triton and blocked with PBS+ (0.5% BSA and 0.15% glycine). Primary antibodies were incubated at room temperature for two hours in PBS+. Primary antibodies

used for those experiments were: anti-RAD51 (rabbit polyclonal)(28), mouse monoclonal anti-BRCA2 (OP95, Calbiochem) and mouse monoclonal anti-RPA32 (ab2175, Abcam). For single color dSTORM experiments RAD51 antibody was directly conjugated with Alexa647 using APEX Antibody Labeling Kit (Molecular Probes, Thermo Scientific). For dual color dSTORM experiments secondary antibodies were used.

After multiple washing steps with PBS with 0.1% Triton coverslips were incubated for 1 hour with secondary antibodies recognizing either rabbit or mouse immunoglobulin conjugated with fluorescent dyes, Alexa647 (Thermo Scientific) or Atto488 (Rockland). To avoid background of free antibodies, the samples were postfixed for 20 min with 2% PFA in PBS. 100nm Tetraspeck beads (Life Technologies) were added to the fixed cells for channel alignment and drift correction. To perform dSTORM imaging coverslips were mounted in a coverslip holder and incubated in imaging buffer containing 25mM MEA (Sigma), 0.5mg/ml Glucose Oxidase (Sigma), 40 µg/ml Catalase (Sigma), and 10% w/v Glucose in Tris-HCl pH 8.0.

dSTORM imaging

Confocal and dSTORM data were acquired on a Zeiss Elyra PS1 system using a 100x 1.49NA α Plan Apochromat objective. Confocal images were acquired using an Argon laser and 633nm diode laser. For dSTORM 488 and 642 100mW diode lasers were used to excite the fluorophores together with, respectively, a BP 495-575 + LP 750 or LP 655 excitation filter. Additionally a 20mW 405 laser was used to increase blinking during dSTORM imaging.

Imaging was performed using HiLo illumination and the data was acquired using a 512 x 512 pixel Andor iXon DU 897 EMCCD camera. Per super resolution image 10 000-15 000 images were acquired with an acquisition time of 33 ms. Two color imaging was performed sequentially, first Alexa647 followed by Atto488.

dSTORM data analysis

A 2D Gaussian fitting algorithm (ZEN 2011, Carl Zeiss, Jena) was used to detect and localize the fluorescent events in the individual frames of the dSTORM movies. Overlapping fluorescent detections were discarded. Detections within a distance of 20 nm and 5 subsequent frames were assumed to originate from the same fluorophore and grouped together. Tetraspeck beads (100 nm) deposited on the cells before the experiment were used to correct for drift. To align the two color dSTORM images the positions of the same fiducials were used for an affine alignment. RAD51 dSTORM images were aligned, with the corresponding confocal image using a channel alignment algorithm in the ZEN2011 software. The confocal images were used to manually select the foci. Circular Regions of Interests with a 1 μ m diameter around the foci were selected using Fiji/ImageJ. The localizations of these ROIs were exported using an ImageJ macro as text files and imported into R using the RStudio for further analysis (R Development Core Team, 2008). Localizations were clustered using a DBSCAN algorithm (Michael Hahsler (2015). dbscan: Density Based Clustering of Applications with Noise (DBSCAN) and Related Algorithms. R package version 0.9-6. (<http://CRAN.R-project.org/package=dbscan>).

References

1. Pellegrini, L., Yu, D.S., Lo, T., Anand, S., Lee, M., Blundell, T.L. and Venkitaraman, A.R. (2002) Insights into DNA recombination from the structure of a RAD51-BRCA2 complex. *Nature*, **420**, 287-293.
2. Yang, H.J., Jeffrey, P.D., Miller, J., Kinnucan, E., Sun, Y.T., Thoma, N.H., Zheng, N., Chen, P.L., Lee, W.H. and Pavletich, N.P. (2002) BRCA2 function in DNA binding and recombination from a BRCA2-DSS1-ssDNA structure. *Science*, **297**, 1837-1848.
3. Jensen, R.B., Carreira, A. and Kowalczykowski, S.C. (2010) Purified human BRCA2 stimulates RAD51-mediated recombination. *Nature*, **467**, 678-U662.
4. Liu, J., Doty, T., Gibson, B. and Heyer, W.D. (2010) Human BRCA2 protein promotes RAD51 filament formation on RPA-covered single-stranded DNA. *Nature Structural & Molecular Biology*, **17**, 1260-1262.
5. Thorslund, T., McIlwraith, M.J., Compton, S.A., Lekomtsev, S., Petronczki, M., Griffith, J.D. and West, S.C. (2010) The breast cancer tumor suppressor BRCA2 promotes the specific targeting of RAD51 to single-stranded DNA. *Nature Structural & Molecular Biology*, **17**, 1263-1265.
6. Shahid, T., Soroka, J., Kong, E.H., Malivert, L., McIlwraith, M.J., Pape, T., West, S.C. and Zhang, X. (2014) Structure and mechanism of action of the BRCA2 breast cancer tumor suppressor. *Nat Struct Mol Biol*, **21**, 962-968.
7. Siaud, N., Barbera, M.A., Egashira, A., Lam, I., Christ, N., Schlacher, K., Xia, B. and Jasin, M. (2011) Plasticity of BRCA2

Function in Homologous Recombination: Genetic Interactions of the PALB2 and DNA Binding Domains. *Plos Genetics*, **7**.

8. Sy, S.M.H., Huen, M.S.Y. and Chen, J.J. (2009) PALB2 is an integral component of the BRCA complex required for homologous recombination repair. *P Natl Acad Sci USA*, **106**, 7155-7160.

9. Sanchez, H., Kertokallio, A., van Rossum-Fikkert, S., Kanaar, R. and Wyman, C. (2013) Combined optical and topographic imaging reveals different arrangements of human RAD54 with presynaptic and postsynaptic RAD51-DNA filaments. *P Natl Acad Sci USA*, **110**, 11385-11390.

10. 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. *et al.* (2014) BRCA2 diffuses as oligomeric clusters with RAD51 and changes mobility after DNA damage in live cells. *Journal of Cell Biology*, **207**, 599-613.

11. Heilemann, M., van de Linde, S., Schuttpelz, M., Kasper, R., Seefeldt, B., Mukherjee, A., Tinnefeld, P. and Sauer, M. (2008) Subdiffraction-resolution fluorescence imaging with conventional fluorescent probes. *Angew Chem Int Ed Engl*, **47**, 6172-6176.

12. Dosztanyi, Z., Csizmok, V., Tompa, P. and Simon, I. (2005) IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content. *Bioinformatics*, **21**, 3433-3434.

13. Fukuchi, S., Hosoda, K., Homma, K., Gojobori, T. and Nishikawa, K. (2011) Binary classification of protein molecules into intrinsically disordered and ordered segments. *Bmc Struct Biol*, **11**.

14. Henricksen, L.A., Umbricht, C.B. and Wold, M.S. (1994) Recombinant replication protein A: expression, complex formation, and functional characterization. *J Biol Chem*, **269**, 11121-11132.
15. Modesti, M., Ristic, D., van der Heijden, T., Dekker, C., van Mameren, J., Peterman, E.J.G., Wuite, G.J.L., Kanaar, R. and Wyman, C. (2007) Fluorescent human RAD51 reveals multiple nucleation sites and filament segments tightly associated along a single DNA molecule. *Structure*, **15**, 599-609.
16. Candelli, A., Holthausen, J.T., Depken, M., Brouwer, I., Franker, M.A., Marchetti, M., Heller, I., Bernard, S., Garcin, E.B., Modesti, M. *et al.* (2014) Visualization and quantification of nascent RAD51 filament formation at single-monomer resolution. *Proc Natl Acad Sci U S A*, **111**, 15090-15095.
17. van der Linden, E., Sanchez, H., Kinoshita, E., Kanaar, R. and Wyman, C. (2009) RAD50 and NBS1 form a stable complex functional in DNA binding and tethering. *Nucleic Acids Res*, **37**, 1580-1588.
18. Mazin, A.V., Zaitseva, E., Sung, P. and Kowalczykowski, S.C. (2000) Tailed duplex DNA is the preferred substrate for Rad51 protein-mediated homologous pairing. *Embo J*, **19**, 1148-1156.
19. Krawczyk, P.M., Eppink, B., Essers, J., Stap, J., Rodermond, H., Odijk, H., Zelensky, A., van Bree, C., Stalpers, L.J., Buist, M.R. *et al.* (2011) Mild hyperthermia inhibits homologous recombination, induces BRCA2 degradation, and sensitizes cancer cells to poly (ADP-ribose) polymerase-1 inhibition. *Proc Natl Acad Sci U S A*, **108**, 9851-9856.

20. Sanchez, H. and Wyman, C. (2015) SFMetrics: an analysis tool for scanning force microscopy images of biomolecules. *BMC Bioinformatics*, **16**, 27.
21. Lamour, G., Kirkegaard, J.B., Li, H., Knowles, T.P. and Gsponer, J. (2014) Easyworm: an open-source software tool to determine the mechanical properties of worm-like chains. *Source Code Biol Med*, **9**, 16.
22. Ristic, D., Sanchez, H. and Wyman, C. (2011) Sample preparation for SFM imaging of DNA, proteins, and DNA-protein complexes. *Methods in molecular biology*, **783**, 213-231.
23. Sanchez, H., Kanaar, R. and Wyman, C. (2010) Molecular recognition of DNA-protein complexes: A straightforward method combining scanning force and fluorescence microscopy. *Ultramicroscopy*, **110**, 844-851.
24. Kerssemakers, J.W., Munteanu, E.L., Laan, L., Noetzel, T.L., Janson, M.E. and Dogterom, M. (2006) Assembly dynamics of microtubules at molecular resolution. *Nature*, **442**, 709-712.
25. van Mameren, J., Modesti, M., Kanaar, R., Wyman, C., Peterman, E.J. and Wuite, G.J. (2009) Counting RAD51 proteins disassembling from nucleoprotein filaments under tension. *Nature*, **457**, 745-748.
26. Ristic, D., Modesti, M., van der Heijden, T., van Noort, J., Dekker, C., Kanaar, R. and Wyman, C. (2005) Human Rad51 filaments on double- and single-stranded DNA: correlating regular and irregular forms with recombination function. *Nucleic Acids Res*, **33**, 3292-3302.

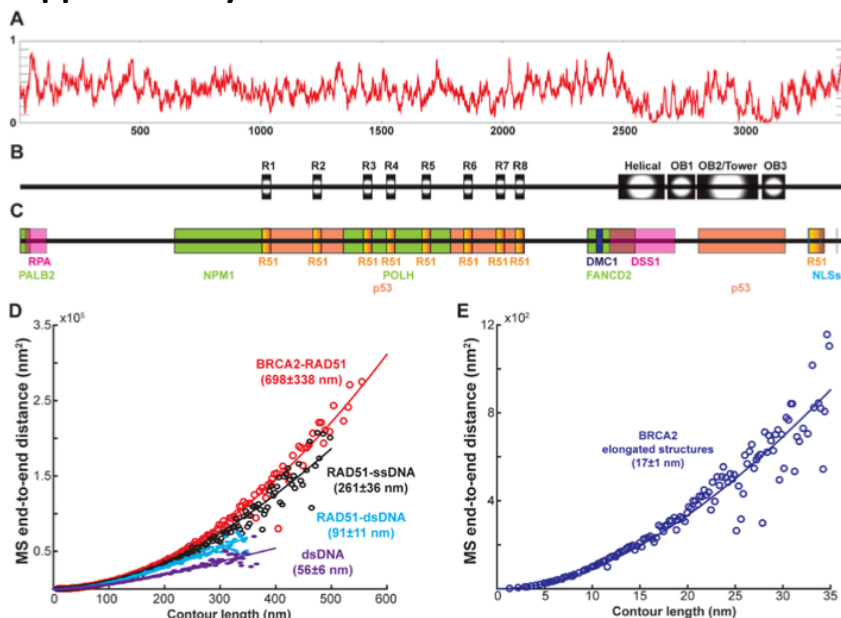
27. Bustamante, C., Keller, D. and Yang, G.L. (1993) Scanning Force Microscopy of Nucleic-Acids and Nucleoprotein Assemblies. *Curr Opin Struct Biol*, **3**, 363-372.
28. van Veelen, L.R., Cervelli, T., van de Rakt, M.W., Theil, A.F., Essers, J. and Kanaar, R. (2005) Analysis of ionizing radiation-induced foci of DNA damage repair proteins. *Mutat Res*, **574**, 22-33.
29. Janicijevic, A., Ristic, D. and Wyman, C. (2003) The molecular machines of DNA repair: scanning force microscopy analysis of their architecture. *J Microsc*, **212**, 264-272.
30. Ratcliff, G.C. and Erie, D.A. (2001) A novel single-molecule study to determine protein--protein association constants. *J Am Chem Soc*, **123**, 5632-5635.
31. Uversky, V.N. (2016) Dancing Protein Clouds: The Strange Biology and Chaotic Physics of Intrinsically Disordered Proteins. *J Biol Chem*, **291**, 6681-6688.
32. Cortese, M.S., Uversky, V.N. and Dunker, A.K. (2008) Intrinsic disorder in scaffold proteins: getting more from less. *Prog Biophys Mol Biol*, **98**, 85-106.
33. Oldfield, C.J. and Dunker, A.K. (2014) Intrinsically disordered proteins and intrinsically disordered protein regions. *Annu Rev Biochem*, **83**, 553-584.
34. Wright, P.E. and Dyson, H.J. (2015) Intrinsically disordered proteins in cellular signalling and regulation. *Nat Rev Mol Cell Bio*, **16**, 18-29.
35. de Jager, M., van Noort, J., van Gent, D.C., Dekker, C., Kanaar, R. and Wyman, C. (2001) Human Rad50/Mre11 is a flexible complex that can tether DNA ends. *Mol Cell*, **8**, 1129-1135.

36. van Noort, J., van Der Heijden, T., de Jager, M., Wyman, C., Kanaar, R. and Dekker, C. (2003) The coiled-coil of the human Rad50 DNA repair protein contains specific segments of increased flexibility. *Proc Natl Acad Sci U S A*, **100**, 7581-7586.
37. Holloman, W.K. (2011) Unraveling the mechanism of BRCA2 in homologous recombination. *Nature Structural & Molecular Biology*, **18**, 748-754.
38. van der Heijden, T., Seidel, R., Modesti, M., Kanaar, R., Wyman, C. and Dekker, C. (2007) Real-time assembly and disassembly of human RAD51 filaments on individual DNA molecules. *Nucleic Acids Research*, **35**, 5646-5657.
39. Bekker-Jensen, S., Lukas, C., Kitagawa, R., Melander, F., Kastan, M.B., Bartek, J. and Lukas, J. (2006) Spatial organization of the mammalian genome surveillance machinery in response to DNA strand breaks. *The Journal of cell biology*, **173**, 195-206.
40. Rossy, J., Cohen, E., Gaus, K. and Owen, D.M. (2014) Method for co-cluster analysis in multichannel single-molecule localisation data. *Histochem Cell Biol*, **141**, 605-612.
41. Essers, J., Hendriks, R.W., Wesoly, J., Beerens, C.E., Smit, B., Hoeijmakers, J.H., Wyman, C., Dronkert, M.L. and Kanaar, R. (2002) Analysis of mouse Rad54 expression and its implications for homologous recombination. *DNA Repair (Amst)*, **1**, 779-793.
42. Ghaemmaghami, S., Huh, W.K., Bower, K., Howson, R.W., Belle, A., Dephoure, N., O'Shea, E.K. and Weissman, J.S. (2003) Global analysis of protein expression in yeast. *Nature*, **425**, 737-741.

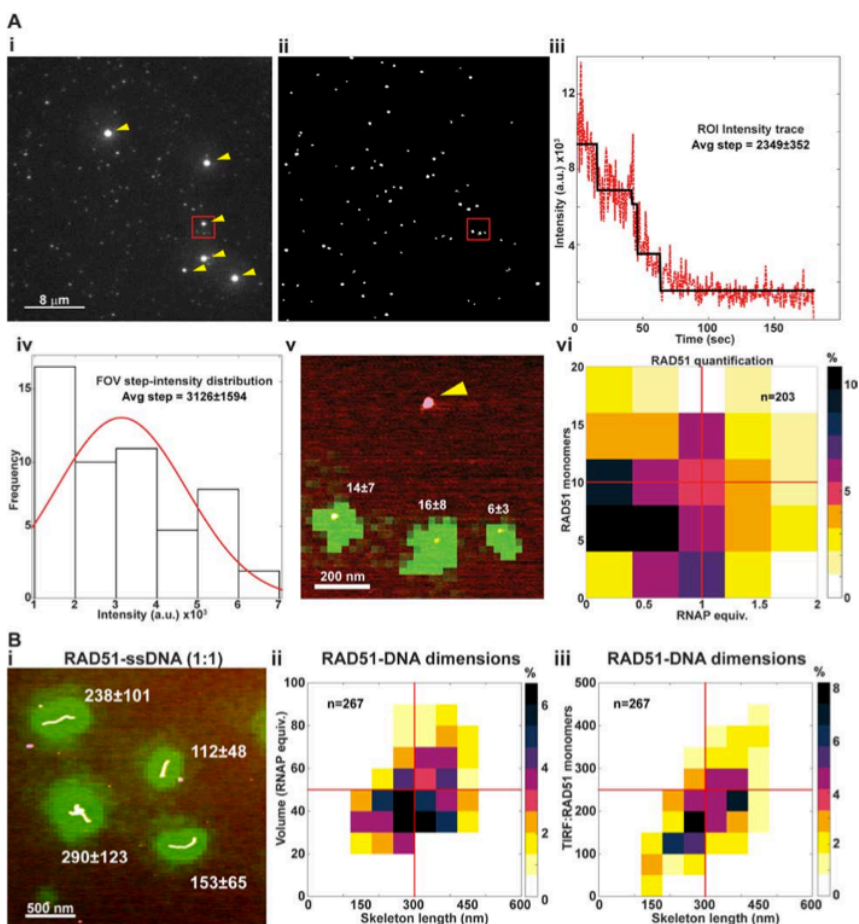
43. Shoemaker, B.A., Portman, J.J. and Wolynes, P.G. (2000) Speeding molecular recognition by using the folding funnel: the fly-casting mechanism. *Proc Natl Acad Sci U S A*, **97**, 8868-8873.
44. Altmeyer, M., Neelsen, K.J., Teloni, F., Pozdnyakova, I., Pellegrino, S., Grofte, M., Rask, M.B.D., Streicher, W., Jungmichel, S., Nielsen, M.L. *et al.* (2015) Liquid demixing of intrinsically disordered proteins is seeded by poly(ADP-ribose). *Nat Commun*, **6**.
45. Hilario, J., Amitani, I., Baskin, R.J. and Kowalczykowski, S.C. (2009) Direct imaging of human Rad51 nucleoprotein dynamics on individual DNA molecules. *Proc Natl Acad Sci U S A*, **106**, 361-368.
46. Subramanyam, S., Ismail, M., Bhattacharya, I. and Spies, M. (2016) Tyrosine phosphorylation stimulates activity of human RAD51 recombinase through altered nucleoprotein filament dynamics. *Proc Natl Acad Sci U S A*, **113**, E6045-E6054.
47. Sing, C.E., Olvera de la Cruz, M. and Marko, J.F. (2014) Multiple-binding-site mechanism explains concentration-dependent unbinding rates of DNA-binding proteins. *Nucleic Acids Res*, **42**, 3783-3791.
48. Tarsounas, M., Davies, D. and West, S.C. (2003) BRCA2-dependent and independent formation of RAD51 nuclear foci. *Oncogene*, **22**, 1115-1123.
49. Wang, X. and Haber, J.E. (2004) Role of *Saccharomyces* single-stranded DNA-binding protein RPA in the strand invasion step of double-strand break repair. *PLoS Biol*, **2**, E21.
50. Essers, J., Houtsmuller, A.B., van Veelen, L., Paulusma, C., Nigg, A.L., Pastink, A., Vermeulen, W., Hoeijmakers, J.H.J. and Kanaar, R. (2002) Nuclear dynamics of RAD52 group homologous

recombination proteins in response to DNA damage. *Embo J*, **21**, 2030-2037.

Supplementary Data

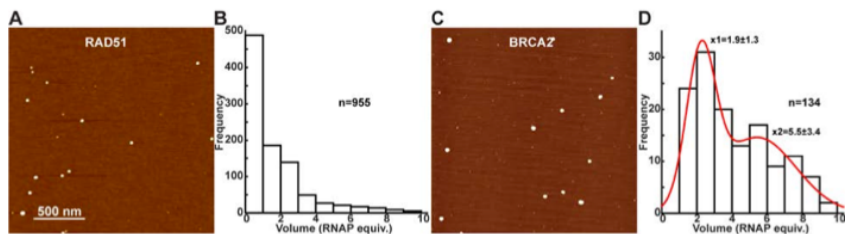


Supplementary Figure 1. Intrinsic disordered regions span the entire BRCA2 sequence. (A) Amino acid sequence of human BRCA2 analysed by IUPred web tool (1). Computed disorder probability (red line) is plotted against residue number. (B) Folded domains described in BRCA2. R1-R8: BRC repeats; Helical: Helical domain, OB: Oligosaccharide binding domain; Tower: Tower domain. (C) Interaction regions identified between BRCA2 and PALB2 (2), RPA (3), nucleophosmin (NPM1) (4), RAD51 (R51) (5,6), polymerase eta (7), p53 (8), DMC1 (9), FANCD2 (10), DSS1 (11), and the nuclear localization signals (NLSs) (12). (D) BRCA2 is a flexible polymer that produces rigid rods upon RAD51 interaction. Mean square (MS) end-to-end distances plotted as a function of the contour length (circles) of RAD51-BRCA2 complexes ($n=123$), dsDNA ($n=107$), RAD51-dsDNA filaments ($n=101$), and RAD51-ssDNA ($n=105$) filaments in nanometers (nm). Lines are fits (with coefficient of determination, indicating goodness of the fit, >0.9) of the worm-like chain model to the data. Apparent persistence length for each complex is indicated between brackets \pm SD. (E) Similar analysis for the BRCA2 V-shape domains ($n=111$)

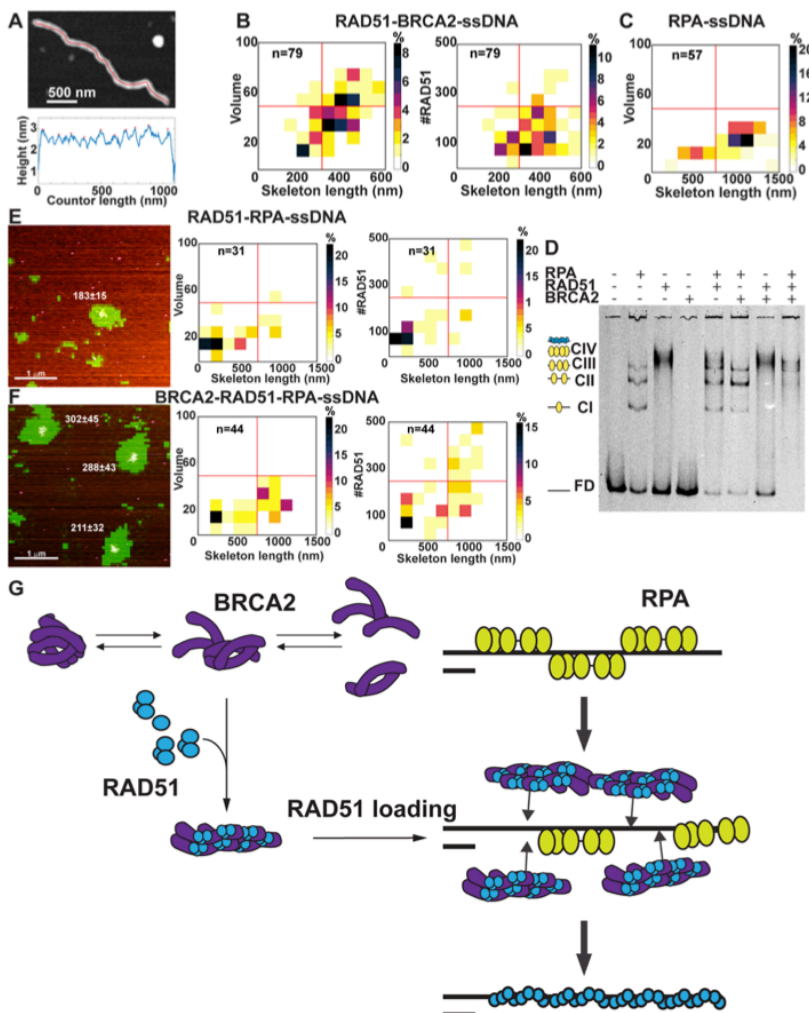


Supplementary Figure 2. Quantification of RAD51 monomers by fluorescence. (A) Analysis of RAD51 protein. (i) TIRF image (maximum intensity projection from a stack of 300 frames) of fluorescent RAD51. Yellow arrowheads indicate fluorescent fiducials. Red box shows the region enlarged in panel v. (ii) Binary mask obtained from image in panel A showing selected regions of interest (ROIs) for quantitative analysis. (iii) Example intensity trace from one ROI in panel ii. Number of fluorophores were determined based on intensity steps from bleaching and blinking of single AF488 dyes. (iv) Distribution of step intensities from all ROIs in panel ii presented as histogram with bin size of 1000 intensity arbitrary units (a.u.). FOV: Field of view. (v) RAD51 proteins visualized by TIRF-SFM. SFM scan overlaid with S2fluorescence image (green) from boxed area in panels i and ii. Numbers indicate RAD51 monomers \pm SD per ROI. Yellow arrowhead indicates a fluorescent fiducial. (vi) Distribution of RAD51 protein with respect to volume and number of fluorophores. Individual complexes from images similar to panel (v) were measured and their distribution plotted in a two dimensional histogram with bin size 0.4 RNAP equiv. \times 4 fluorophores. Color indicates percentage of total ROIs analyzed ($n=203$) as shown in the color bar. Red lines are guides to view bins. (B) Analysis of RAD51-ssDNA

complexes. (i) Calibration of TIRF-SFM set-up for quantification of fluorescent RAD51. SFM scan overlaid with fluorescence image (green) of ssDNA incubated with RAD51 and cross-linked with glutaraldehyde. Numbers indicate the amount of RAD51 monomers \pm SD per protein complex. (ii) Volume and length of individual complexes from images similar to panel (i) were measured and their distribution plotted in a two dimensional histogram with bin size 60 nm x 10 RNAP equiv. Color indicates percentage of total complexes analyzed (n=267) as shown in the color bar. (iii) Distribution of RAD51-ssDNA complexes based on number of fluorophores (RAD51) and length. Bin size 60 nm x 50 RAD51 monomers.

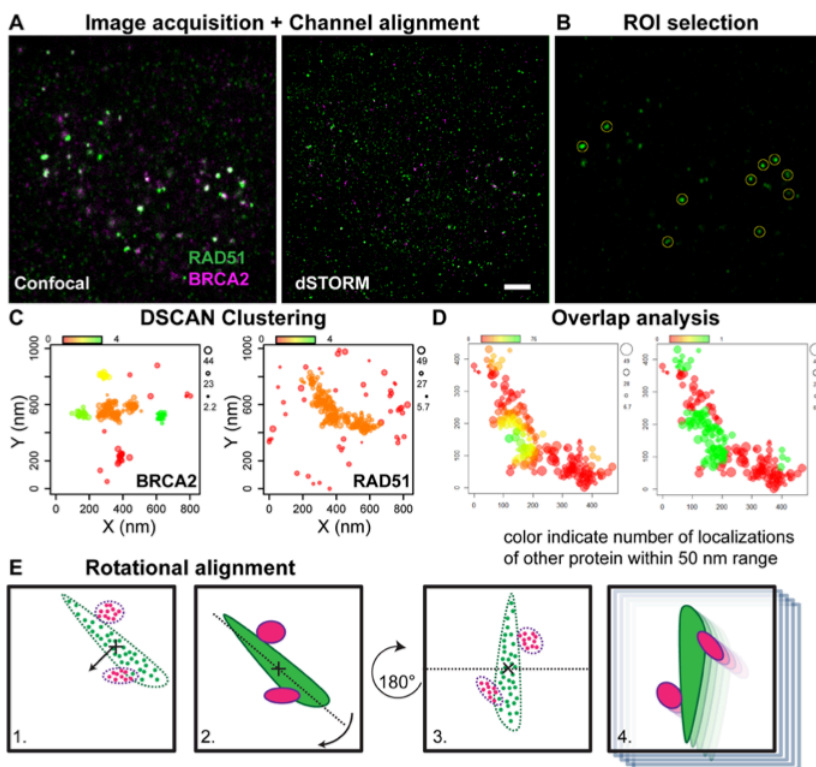


Supplementary Figure 3. RAD51 and BRCA2 alone do not make elongated structures. (A) Representative SFM image of RAD51 incubated with glutaraldehyde. Image size 2 x 2 μ m. (B) Measurements (n=955) of RAD51 volume from images such as (A) presented as distribution in a histogram with bin size of 1 RNAP equiv. (C) Representative SFM image of BRCA2 incubated with glutaraldehyde. Image size 2 x 2 μ m. (D) Measurements (n=134) of BRCA2 volume from images such as (C) presented as distribution in a histogram with bin size of 1 RNAP equiv. Red line is the Gaussian fit with two populations centered around 1.9 and 5.5 RNAP equiv., respectively.



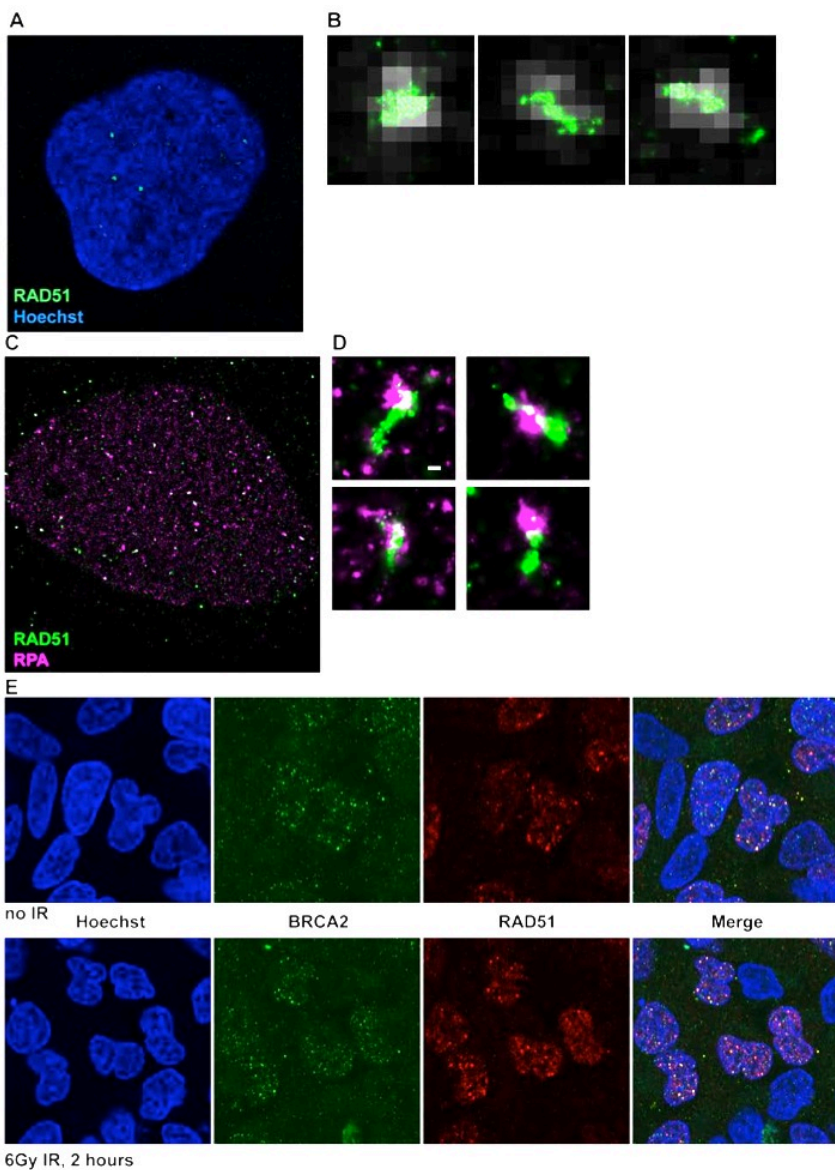
Supplementary Figure 4. (A) Example of elongated BRCA2-RAD51 (unlabeled) complex with the central contour line indicated in red. Height profile along the red contour line is shown in the graph below with automated peak detections indicated by red circles. (B) Distribution of RAD51-BRCA2-ssDNA complexes based on volume, length, and number of fluorophores. All complexes from images similar to panel (ii) and (iii) in Figure 3A of the main text were measured and their distribution plotted in a two dimensional histogram with bin size 60 nm x 10 RNAP equiv. (left) and 60 nm x 50 RAD51 monomers (right). (C) Distribution of RPA-ssDNA of complexes based on volume and length. RPA-ssDNA complexes had an average volume of 21 ± 14 (SD) RNAP equiv. (D) BRCA2 remodels RPA and helps RAD51 loading on ssDNA. DNA mobility shift assays after incubation with (+) or without (-) proteins ; RPA, RAD51 and BRCA2. RPA binds to the fluorescence oligonucleotide with one (CI), two (CII), three (CIII), or four (CIV) monomers

(lane 2). In the presence of BRCA2 the amount of CI and free DNA (FD) was reduced with a concomitant increase in CII, indicating a reorganization of the proteins bound to the DNA (lane 6). RAD51 in molar excess shifted more than 50% of the DNA (lane 3), however DNA pre-incubation with RPA inhibited drastically the binding of RAD51 (lane 5). Addition of pre-incubated BRCA2-RAD51 to the reaction recovered the loading of RAD51 on RPA bound DNA (lane 8). Yellow ellipses represent RPA monomers and blue ellipses RAD51 monomers. (E) Effect of BRCA2 on RAD51 loading on RPA covered ssDNA. RAD51-RPA-ssDNA complexes visualized by TIRF-SFM without fixation. and distribution of RAD51-RPA-ssDNA complexes based on volume and length, or number of fluorophores (RAD51) and length. (F) BRCA2-RAD51-RPA-ssDNA complexes visualized by TIRF-SFM without fixation and . distribution of BRCA2-RAD51-RPA-ssDNA as above. Numbers in white indicate RAD51 monomers \pm SD for the DNA-protein complexes shown. (G) Cartoon illustrating ideas for the dynamic structural rearrangement of oligomeric BRCA2 leading to RAD51 loading on DNA covered by RPA.



Supplementary Figure 5. Analysis of dSTORM super resolution data. (A) Two color confocal images and dSTORM images were acquired using the same microscope and objective lens. dSTORM images were acquired as describe in the Material and Methods section. (B) Using a channel alignment algorithm within the Zeiss ZEN software, one of both channels of the confocal images was aligned with the dSTORM image. Subsequently ROIs were manually selected from confocal images of the RAD51 staining. The single

molecule localizations of these ROIs were analyzed in R-Studio. (C) The localization data from the ROIs was clustered using a Density Based Clustering Algorithm with Noise (DBSCAN) algorithm. In this algorithm localizations are clustered together based on the local density of localizations with a defined radius, epsilon, and minimum number of points within that radius. The clustering was used to quantify the size and shape of the clusters. Furthermore localizations outside of any cluster were filtered out for further analysis. The plots show the dSTORM localizations of a BRCA2/RAD51 focus. The colors indicate the identified cluster by DBSCAN. Localizations with a cluster number 0 (red) are considered background. (D) To assess the overlap between RAD51 and respectively BRCA2 or RPA; from every RAD51 localization the number of localizations of the other species was measured within a 50 nm radius, this can be visualized by a gradient scatter plot, showing regions with much overlap between both proteins and regions without overlap. A threshold was set on the number of localizations within 50 nm of every RAD51 localization at 1% of all localizations of either BRCA2 or RAD51. This results in a scatter plot with either overlapping (green) or not overlapping (red) localizations. For every ROI the fraction of localizations of RAD51 that overlapped was quantified. (E) Individual images of DSB foci were aligned by rotation and visualized as one image. For every individual image the following algorithm was performed: 1) The center of mass of the RAD51 signal was determined and the localizations translated accordingly to have the center of mass in the middle of the image. 2) The major axis of the RAD51 signal was determined and the angle of this major axis with respect with the vertical axis was used to rotate the localization with the center of mass as rotation point. 3) For RAD51 the number of localizations above and below the horizontal axis was measured. If the amount of localizations above the horizontal axis was smaller than below the axis the image was rotated for another 180 degrees. 4) Finally the coordinates of the rotated dSTORM images were merged and visualized as a single super resolution image.



Supplementary Figure 6. (A) Confocal image of a U2OS cell nucleus treated with 6Gy IR , fixed after 2 hours and stained with RAD51 antibody. The DNA is visualized with Hoechst.

(B) Examples of repair foci imaged both with confocal (grey) and dSTORM (green). Showing the large increase in resolution of dSTORM compared to confocal. (C) A two-color dSTORM image of RAD51 with RPA treated with 6Gy IR and fixed after two hours. (D) Examples of DSB foci showing the localization of RAD51 and RPA together. Scale bar is 100 nm. (E) Confocal images of cells immunostained for BRCA2 and RAD51 of cells treated without and with ionizing radiation (IR). Showing co-localization of BRCA2 and RAD51 in DSB foci. The DNA is visualized by Hoechst.

Chapter 6

Nanoscale mapping: uncovering the structure of BRCA2 protein by imaging individual domains.

M. Grosbart¹, B. Verhagen¹, H. Sánchez¹, S.E. van Rossum¹, R. Kanaar^{1,2}, C. Wyman^{1,2}

Department of ¹Molecular Genetics and ²Radiation Oncology Erasmus University Medical Center, P. O. Box 2040, 3000CA Rotterdam, The Netherlands

Abstract

BRCA2 is a crucial component of dsDNA break response. Loss of the protein is lethal in cells and impaired function in mammals results in genomic instability and increased risk of developing cancer. Parts of BRCA2's role in Homologous Recombination have been uncovered and linked to structural features. However, structural studies are hindered by the structural nature of the protein and difficulties in obtaining preparations of desired concentrations and purity. In our previous work, we demonstrated that BRCA2 functions as dynamic, elongated oligomers. To investigate which parts of BRCA2 are responsible for flexibility and/or involved in oligomerization, we generated truncated versions of the protein and examined them using Scanning Force Microscopy (SFM). Our findings point to multiple interaction interfaces on various parts of the protein that all contribute to formation of irregular, elongated BRCA2 oligomers. We conclude that the N-terminal portion of the protein is usually positioned on the exterior of the complexes and heavily contributes to formation of elongated protrusions observed in most BRCA2 oligomers. The middle region, referred to as R because it encompasses the RAD51 binding BRC repeats, substantially contributes to formation of oligomers, apparently involving multiple transient interactions. Changes in the R domain domain in response to RAD51 allowed us to speculate on BRCA2/RAD51 complex formation.

Introduction

BRCA2 is a tumor suppressor protein with an established role in development of hereditary cancer (Nathanson et al., 2001). Mechanistically it is responsible for accumulation of RAD51 at the damage site (Davies et al., 2001; Thorslund et al., 2010), RAD51 filament formation (Liu et al., 2010), and rearrangement of RPA on processed, ssDNA overhangs (Jensen et al., 2010; Sanchez et al., 2017), but many of its functions and properties are yet not clearly understood.

BRCA2 is a large protein (384 kDa) predicted to contain stretches of intrinsic disorder throughout its whole length of 3418 amino acids (Sanchez et al., 2017). BRCA2 can be roughly divided into three parts including different known interaction domains: The C-terminal third includes ssDNA and DSS1-binding motifs (Yang et al., 2002), the middle third includes 8 BRC repeats responsible for RAD51 binding (Pellegrini et al., 2002; Sharan et al., 1997; Wong et al., 1997) and the N-terminal third includes the PALB2 binding domain (Oliver et al., 2009)(Fig 1). Crystallographic reconstructions have been obtained only for parts of BRCA2, namely the ssDNA binding domain co-purified with DSS1 (Yang et al., 2002) and one BRC repeat fused with RAD51 (Pellegrini et al., 2002). These fragmentary data, provided little insight into understanding the overall architecture of BRCA2, and the N-terminal part of the protein, thought to be mostly intrinsically disordered, remains structurally uncharted. Recently, the arrangement of BRCA2 dimers associated with RAD51 and single stranded DNA (ssDNA) has been determined from electron

microscopy (EM) image reconstruction (Shahid et al., 2014). In our SFM studies, though dimers were prominent, BRCA2 was observed as an array of oligomeric forms (Sanchez et al., 2017).

Because SFM images of BRCA2 show all forms present and their distribution, this gives a different view of the proteins structure compared to bulk analysis of class averaging methods described above. SFM images reveal BRCA2 as always oligomeric, with complexes comprising 2 up to 6 monomers. The architecture of BRCA2 oligomers can be described as branched and elongated, and the heterogeneity of conformation states suggests high structural dynamics and/or plasticity. Most notably, said architecture is dramatically altered by relatively mild changes in conditions. BRCA2 molecules collapsed and lost their elongated form at 0°C, and formed a multiprotein, regular rod-like structures when bound to RAD51 (Sanchez et al., 2017). Thus BRCA2 appeared as an array of oligomeric, elongated forms, prone to dramatic conformation changes upon ligand binding and environmental changes. These features are reminiscent of IDP's which have a propensity for rapid structural changes upon relatively mild altering in environmental conditions that can include temperature, pH, ionic strength or red-ox status of the solution (Nott et al., 2015; Smith and Jelokhani-Niaraki, 2012).

In order to further understand and characterize the architectural arrangement of BRCA2, we started by dividing the protein into 3 parts and investigated these independently via SFM.

We divided BRCA2 roughly into three parts: the C-terminal domain (C), the middle including the BRC repeats (R), and the N-

terminal domain (N). We designed pHCMV1 plasmids encoding these as MBP-tagged BRCA2 versions: with either the N-or C-terminal deleted (BRCA2-NR and BRCA2-RC) and the domains separately (BRCA2-N, BRCA2-R and BRCA2-C). (Fig 1.) Our SFM analysis indicated that the R domain must contribute to oligomerization. BRCA2-N was the most elongated and BRCA2-NR had the most pronounced elongated, branched characteristics, which were retained even in low temperature, suggesting the N terminal domain can be extended and elongated. BRCA2-R, while being di- and tetrameric on its own, became prevalently mono- and dimeric after incubation with RAD51. This was in contrast to the full length BRCA2, where addition of RAD51 induced formation of oligomeric rod-like arrangement. This suggests that require N and C termini contribute to oligomerization in the RAD51 bound state.

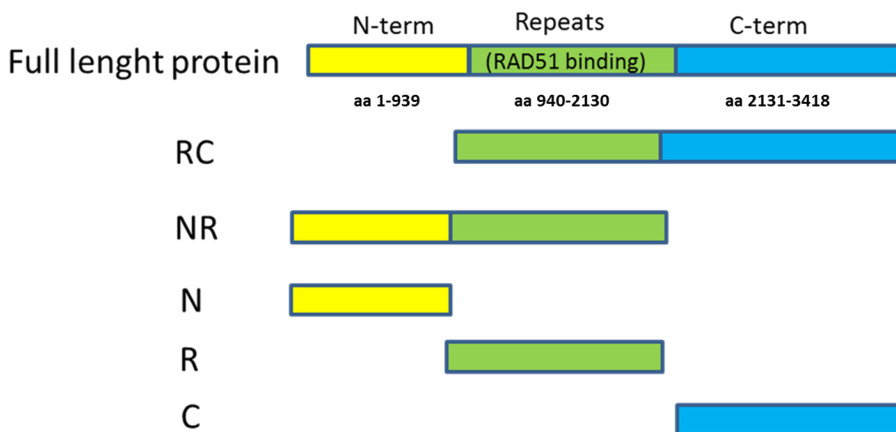


Fig.1. A schematic diagram of BRCA2, highlighting locations of functionally crucial domains. Below, schematic representations of truncation variants created for SFM studies.

The work presented here increases understanding of the structure of BRCA2, but leads to new questions and shows an urgent need for further analysis of the inter- and intra-molecular interactions involved in the structural variation observed.

Results

Both N and C terminal domains contribute to oligomer formation

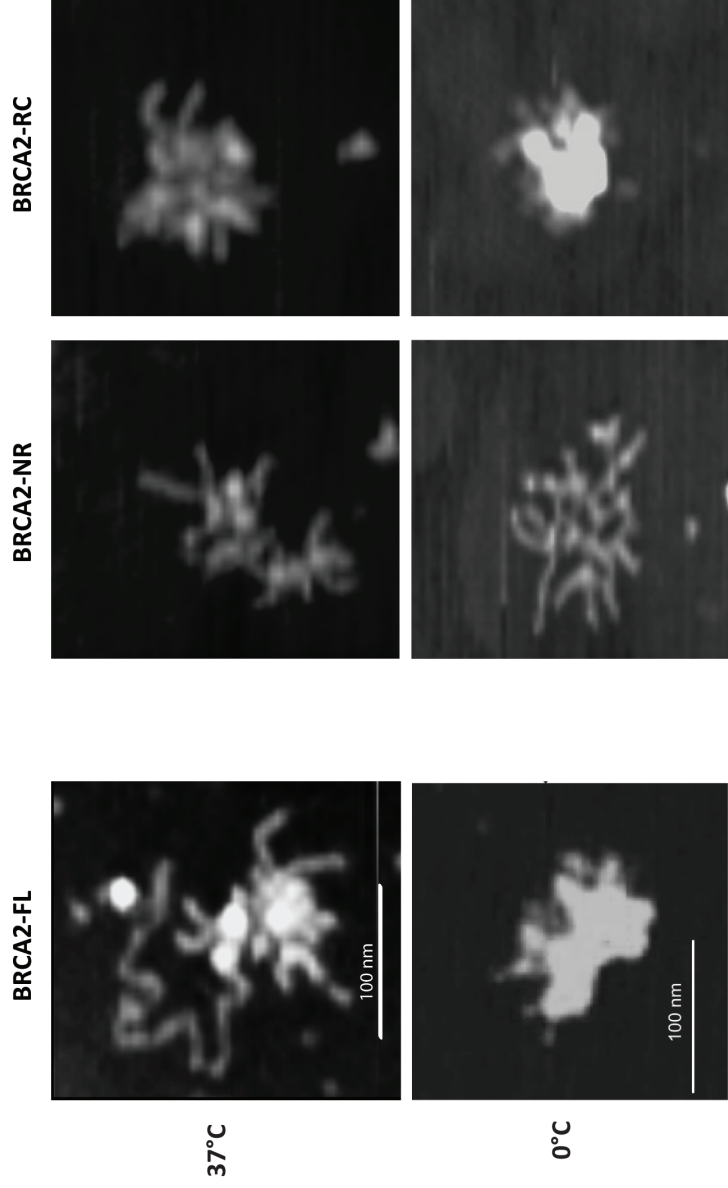
SFM imaging in near physiological conditions revealed that BRCA2 appears as a heterogeneous mixture of oligomers, comprising between two and six BRCA2 monomers. The shape of these molecules is irregular and often elongated, with visible protrusions extending for up to 50 nm (Sanchez et al., 2017) (Fig 2A). In order to investigate which domains are responsible for oligomerization and the peculiar shape of BRCA2 we generated MBP-tagged truncation variants of the protein: BRCA2-NR with C-terminal deletion (aa 1-2130 ,missing aa 2131-3418) and BRCA2-RC with N-terminal deletion (aa 2131-3418, missing aa 1-939) (Fig 1). All truncation variants and full-length (FL) BRCA2 were successfully purified from HEK 293T cells. The proteins were deposited for imaging in identical conditions and similar concentration, following the protocols we have used in previous studies of BRCA2 architecture (Sanchez et al., 2017). Quantitative analysis included measuring the volume of individual molecules (determining oligomeric state) and their skeleton length (indicative of elongation) (Fig 2B). SFM imaging of both BRCA2-NR and BRCA2-RC revealed familiar, branched structures, overall very

similar to those formed by full length BRCA2 (Fig 2A, upper panel). Although multimeric forms were observed as shown in figure 2A, from the volumetric analysis it was clear that the predominant fraction of molecules was monomeric (Fig 2B). Previous studies noted a lack of monomers in the FL BRCA2 preparations (Sanchez et al., 2017; Shahid et al., 2014). Molecules with volume corresponding to one BRCA2-NR/RC were irregular, appeared collapsed and did not display specific architectural arrangement. After closer scrutiny, it became apparent that the dimensions of BRCA2-NR vary considerably from BRCA2-RC. Based on volume and skeleton length, the distribution of the oligomeric forms of the BRCA-NR and BRCA-RC is similar to each other (and to FL BRCA2 (Sanchez et al., 2017)), ranging from dimers to pentamers (Fig 2B). There is a discrepancy, however, in the respective skeleton length, where BRCA2-NR tended to be more elongated than BRCA2-RC for comparable volume (Fig 2B, center and right). The same is true for the monomeric fractions of both proteins (Fig 2B). The N-terminal domain is largely responsible for the elongated protrusions in BRCA2 oligomers. Interestingly, the abundance of monomeric forms in both truncation variants preparations suggests that interaction between the N- or the C-terminus promote oligomerization.

Temperature-dependent conformation change of BRCA2 is abolished in the absence of the C-terminus

Rapid structural alterations upon environmental changes, like changes in osmotic strength, temperature or pH of the solution, are characteristic for IDP's (Nott et al., 2015; Smith and Jelokhani-Niaraki, 2012; Uversky, 2016). BRCA2's structure differs

A)



B)

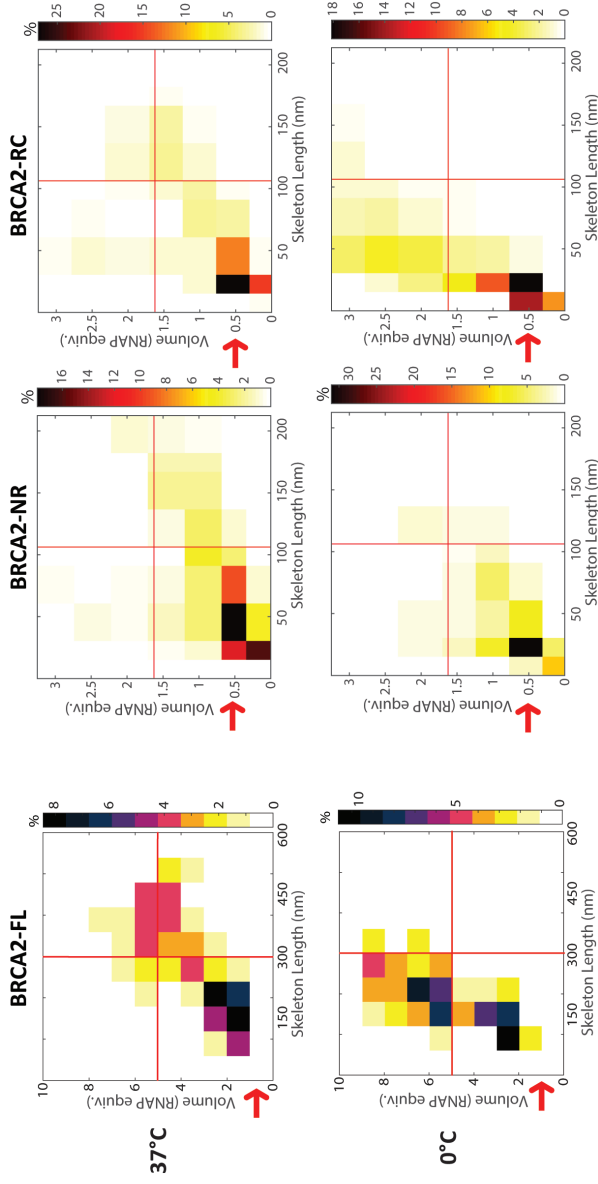


Fig 2.A) Topography of BRCA2 FL (left), BRCA2-NR (center) and BRCA2-RC (right) at 37°C (upper panel) and 0°C (lower panel) as seen in SFM. B) Distribution of complexes based on volume (expressed as RNA Polymerase equivalents, 1~500 kDa) and skeleton length at 37°C (upper panel) and 0°C (lower panel) for BRCA2 FL, BRCA-NR and BRCA2 RC. Red arrows indicate the expected position of monomers. Color represents frequency as indicated in the scale to the right. All distributions represent $n \geq 500$.

dramatically in 37°C comparing to 0°C, a trait uncommon for structurally stable proteins, but easily explainable by high content of intrinsically disordered regions. At 0°C, most of BRCA2's elongated features are not visible and the molecules assume a globular shape, as if protrusions would be collapsed (Fig 2A, lower panel)(Sanchez et al., 2017). When incubated in 0°C, BRCA2-RC also appeared collapsed, globular oligomers (Fig 2B). Quantification confirmed these observations, showing that while the distribution of oligomeric states of BRCA2-RC was the same as incubated in 37°C, but the distribution of skeletal lengths was greatly simplified and shorter (Fig 2B, right). BRCA2-NR did not undergo structural alterations in 0°C. The structural transition of BRCA2-NR is less pronounced, with a sustained strong presence of elongated structures (Fig 2B, center). Molecules with obvious elongations remained at 0°C (Fig 2A). This suggests that the C-terminal domain includes an easily modulated (here by temperature) interaction/oligomerization site.

Individual domains of BRCA2 have distinct spatial characteristics and oligomerization states

To determine which domain/s include oligomerization interface/s, we created BRCA2 truncation variants comprising the C (aa 2131-3418), N (aa 1-939) or the R (940-2130) regions (Fig 1) and successfully purified them from HEK 293T cells. SFM imaging of BRCA2-C revealed it to be monomeric and spherical (Fig 3A, B). BRCA2-N was a mixture of monomers/dimers. The molecules were largely elongated and heterogeneity in their arrangement suggested a degree of flexibility (Fig 3A, B). Strikingly, the molecules were never branched, but rather appeared as linear

flexible structures in different states of elongation. The skeleton length of the predominant monomeric/dimeric population increased with the volume suggesting end- to-end association rather than side-to-side, parallel dimerization. This result, together with dimensional characteristics of BRCA2-NR in 0°C, is consistent with the idea that the N terminal region of BRCA2 is largely responsible for elongation and protruding elements observed. The RAD51-binding region, BRCA2-R, formed irregular spheres. Volumetric analysis identified majority of the structures as dimeric or tetrameric (Fig 3 A, B). The R region is clearly prone to oligomerization

Reorganization of BRCA2-R upon RAD51 binding

RAD51 binding is the biologically crucial characteristics of BRCA2 (Pellegrini et al., 2002; Sharan et al., 1997; Wong et al., 1997). Eight RAD51-binding interfaces (BRC repeats) are located in the central part of BRCA2, we designated as R for repeats. Full length BRCA2 binds 5 to 6 RAD51 molecules *in vitro* (Jensen et al., 2010; Sanchez et al., 2017). The BRCA2-R domain we produced has M_w of 231 kDa, while RAD51 is significantly smaller, 37 kDa. If the expected 6 RAD51 monomers would bind per BRCA2-R, the volumetric change would be easily detectable by SFM. We performed a BRCA2-R/RAD51 binding reaction with saturating amounts of RAD51 to ensure maximal binding and imaged the resulting complexes. The structures were not distinguishable from BRCA2-R by simple inspection, about the same size and overall globular structure (Fig 3A)., and the best candidate for oligomerization promotor in BRCA2.

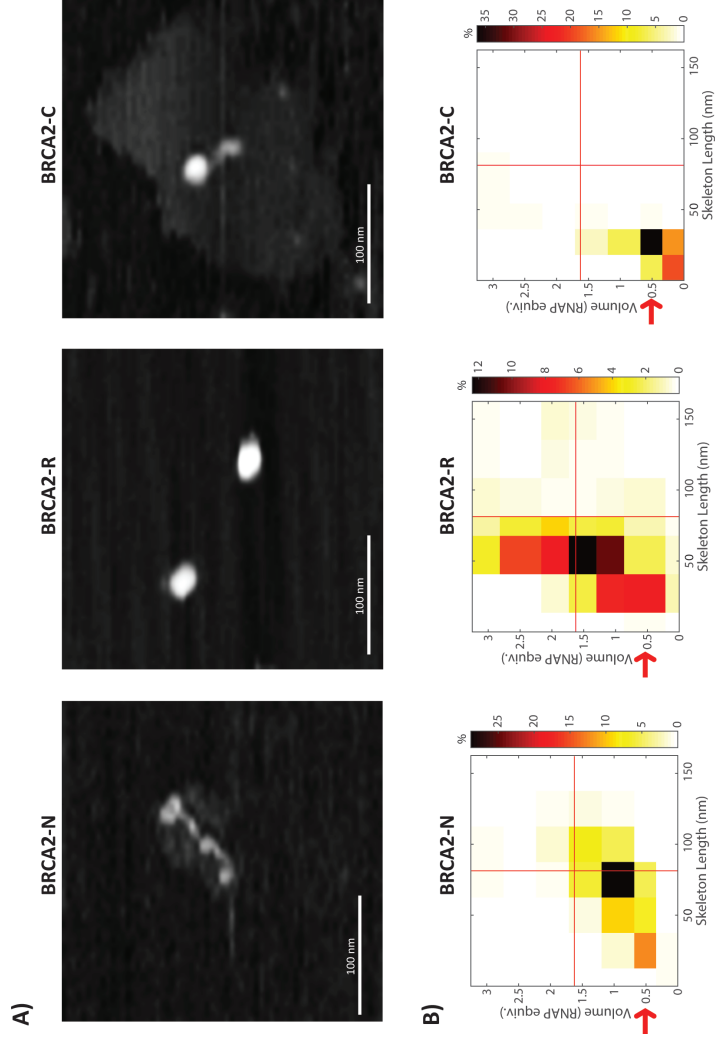


Fig 3. A) Topography of BRCA2 N (left), BRCA2-R (center) and BRCA2-C (right) at 37°C as seen in SFM. **B)** Distribution of complexes based on volume (expressed as RNA Polymerase equivalents, 1~500 kDa) and skeleton length at 37°C for BRCA2 N, BRCA2-R and BRCA2 C. Red arrows indicate the expected position of monomers. All data was collected with $n \geq 500$

Volume analysis showed a clear difference between BRCA2-R and BRCA2-R/RAD51 complexes. In the presence of RAD51 the volume distribution changed, the complexes were smaller and the tetrameric population was greatly diminished. The most prevalent complexes had a volume equivalent to BRCA2-R dimers, or BRCA2-R monomers in complex with 5-6 RAD51 molecules. Using SFM alone we are unable to discriminate between the two possibilities; however, by cross-referencing the result with a volumetric profile of BRCA2-R and RAD51 protein imaged in the same conditions, we can easily appreciate the distinct differences (Fig 4).

Notably, the skeleton length of the BRCA2/RAD51 complexes was significantly increased, suggesting elongation of the complexes in comparison to BRCA2-R alone (Fig 4). This result suggests RAD51 binding induces rearrangements in the R region. Complex formation with RAD51 causes dramatic structural rearrangements in FL BRCA2 (Sanchez et al., 2017). BRC repeats are the primary RAD51 binding interface and the R region, where they are situated, most probably drives oligomerization of BRCA2, as proposed in the previous chapter. FL BRCA2 bound to RAD51 is apparently oligomeric *in vivo* (Reuter et al., 2015) and *in vitro* (Sanchez et al., 2017). It is possible that the oligomerization interfaces within the R region are only partly disrupted by RAD51 binding, or that oligomerization of BRCA2-RAD51 complexes involves different interactions than oligomerization of BRCA2 alone.

The work presented here is a part of an ongoing project and by no means finished. More BRCA2 variants have to be

generated, imaged and analyzed in order to fully comprehend the structural and architectural properties of BRCA2.

Discussion

BRCA2 is a tumor suppressor, implicated in a variety of familial and non-familial cancers (Nathanson et al., 2001). The important role of this protein in coordinating homologous recombination by delivering and loading RAD51 recombinase onto the damaged DNA strand is undebatable. However, question remain about the physical mechanism behind this function. Specifically the importance of the structural plasticity of BRCA2 can now be addressed. Until recently, not much was known about the structure of FL BRCA2. Two parts of BRCA2 have been crystalized. One is the conserved, C-terminal portion of the protein (Yang et al., 2002). The other is one of the BRC RAD51-binding interfaces, located in the middle (R) section of the protein (Pellegrini et al., 2002; Rajendra and Venkitaraman, 2010). These crystalized and well-structured parts, however, constitute less than 1/3 of this

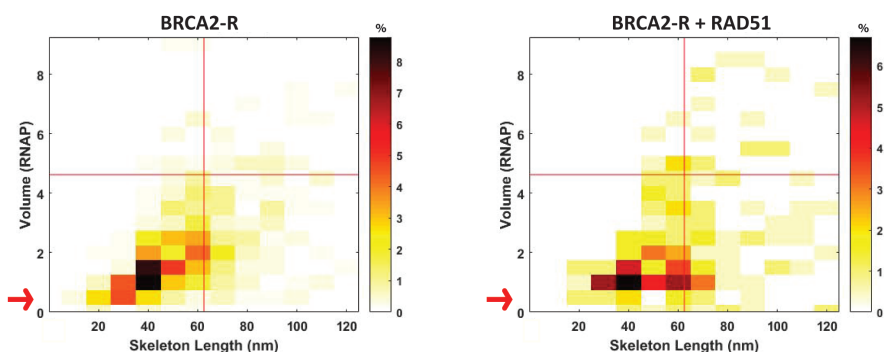


Fig 4. Distribution of of BRCA2-R and BRCA2-R/RAD51 complexes complexes based on volume (expressed as RNA Polymerase equivalents, 1= \sim 500 kDa) and skeleton length at 37°C. . Red arrows indicate the expected position of monomers. All data was collected with $n \geq 500$

450 kDa protein. The structure of the entire N-terminal portion and complete R have not been determined by high resolution methods, and even their precise properties are not known. The size of the protein and the high content of IDPR's has hindered answering these questions for a long time.

Today, structural reconstructions (Shahid et al., 2014) and SFM images of FL BRCA2 molecules (Sanchez et al., 2017) are available. One of the most striking characteristics of BRCA2 shown in these studies (supported by *in vivo* data (Reuter et al., 2015)) is that it commonly exists in oligomeric form. The protein is predicted, based on the amino acid sequence, to include many regions of disorder. IDP's are believed to be in structural flux, and can become ordered in complex with a cofactor, usually a partner protein (Uversky, 2016). That would easily explain difficulties in structural characterization using conventional methods. It might also be, that the oligomeric state BRCA2 assumes in close to physiological conditions represents self-stabilization (Sanchez et al., 2017). In any case, it is apparent that internal oligomerization interfaces exist within BRCA2. Where exactly, and whether they are located in a specific region of the protein or rather scattered along the molecule, is a fundamental question that we have begun to address.

For our initial dissection, BRCA2 was roughly divided into three parts: C-terminal, containing ssDNA- and DSS1-binding interfaces, the "middle", R region with 8 BRC repeats responsible for RAD51 binding; and the largely unstructured, RPA- and PALB1-binding N terminus (Oliver et al., 2009; Pellegrini et al., 2002; Yang

et al., 2002). If the oligomerization interfaces were to be situated in one of these regions, their successive removal should result in breakage of oligomers and subsequent appearance of monomers. With that in mind we created BRCA2-NR, BRCA2-RC, BRCA2-N, BRCA2-R and BRCA2-C truncation variants (Fig 1) and utilized SFM imaging to study their architectural arrangement.

For both BRCA-NR and BRCA2-RC oligomer formation was disturbed but not prevented. A substantial portion of these proteins, appeared as oligomers, with volume ranging from di- to pentamers—stoichiometrically identical to FL protein. These oligomers assumed a similar, heterogeneous appearance as the previously studied FL protein, with characteristic elongated protrusions in a branched arrangement (Sanchez et al., 2017) (Fig 2A). Careful quantification revealed subtle differences between these two truncation variants: BRCA2-NR seemed more elongated and flatter whereas BRCA2-RC was more compact and bluky (Fig 2B). Conversely, the greater skeleton length of BRCA2-NR indicates that the N-terminal region is the elongated portion of the protein. Because oligomers of BRCA2 form whenever the R region is present, the R region is involved in this self association.

Given the heterogeneity of the BRCA2 oligomers observed in SFM, as well as the known properties of IDP's, one can postulate that the structural dynamics could facilitate frequent exchange of monomers in solution (Sanchez et al., 2017; Uversky, 2016). It is tempting to think that the removal of one or the other terminus as in BRCA2-NR and -RC could impair this *process* of structural

dynamics, rather than changing the simple presence or absence of an oligomerization interface.

We have reported that the structural arrangement of BRCA2 changes with temperature, on ice compared to 37°C the protrusions collapse and oligomers assume a globular shape (Sanchez et al., 2017). Only BRCA2-RC mimics the FL protein in this behavior, while low temperature has no BRCA2-NR (Fig 2A, B). The collapse of elongated protrusions is dependent on the presence of the C domain. Since BRCA2-NR does not exhibit any change in quantifiable dimensions at 0°C, it is tempting to argue that interactions between the N and C domain (and, perhaps, the R and C domain), are stabilized at lower temperatures.

The three BRCA2 regions separately provide further insight into their architectural contributions. The BRCA2-N region was distinctly elongated (Fig 3A), suggesting that it is responsible for the protrusions observed in the FL protein. BRCA2-N dimers were also prevalent (Fig 3B) with dimensions consistent with monomers arranged end-to-end, rather than side-to-side. The BRCA2-R fragment alone was oligomeric in nature (Fig 3B) and as noted above confirms the R region is responsible for BRCA2 self association, at least in the absence of RAD51. The precise character and location of the oligomerization interface is, at this point, unknown. BRCA2-C appeared as small, globular, and predominantly monomeric (Fig 3A, B). This is consistent with this region including known structured domains and implies its role in multimerization is through interaction with the N and or R regions.

The R region includes the BRC repeats responsible for RAD51 binding (Pellegrini et al., 2002; Sharan et al., 1997; Wong et al., 1997). We have reported previously that RAD51 binding causes rearrangement of BRCA2, resulting in formation of filamentous complexes (Sanchez et al., 2017). Addition of RAD51 to the BRCA2-R changed the distribution of oligomers, simplifying the distribution and favoring dimers (Fig 4). We were unable, based on volume alone, to unambiguously determine the composition of these complexes. Their measured volume is consistent with BRCA2 dimers, or BRCA2 monomers in complex with eight RAD51's. The former interpretation, however, seems more plausible, as it is not likely to have all of the BRC repeats occupied by RAD51 at a given moment, with the estimated stoichiometry of four to six RAD51 molecules per BRCA2 monomer (Jensen et al., 2010; Sanchez et al., 2017). After interaction with RAD51, BRCA2-R appeared as small, slightly elongated, globular complexes. It is tempting to speculate that larger complexes of RAD51-FL BRCA2 involve interaction of RAD51 binding sites outside of the BRC repeats of one BRCA2 with RAD51 bound to different to another BRCA2 molecule. Further analysis to determine structural characteristics of BRCA2-NR/RAD51 and BRCA2-RC/RAD51 complexes are much needed to resolve the multiple BRCA2-BRCA2 and BRCA2-RAD51 interactions possible in dynamic complex architecture and multimerization.

Materials and Methods

Protein Cloning and Purification

A pHCMV-2 plasmid containing BRCA2 tagged with two tandem repeats of the maltose binding protein (MBP), a generous gift from the Kowalczykowski lab, was used as a template to create the BRCA2-NR, BRCA2-RC, BRCA2-N and BRCA2-C constructs. In turn, the BRCA2-RC was used as a template for the BRCA2-R construct. All construct were made with NEB Q5 site-directed mutagenesis kit (E0552S) and transformed into home-made DH5 α competent cells. Colonies were picked, screened and the DNA fully sequenced. Primers used to make the constructs were:

BRCA “NR”

Forward 5'-TAACTCGAGATGGCCGCGGACTC-3'

Reverse 5'-AAATTCTTTACTGCAGGTTTTTTCC-3'

BRCA “RC”

Forward 5'-AAATTATCAAATAACTTAAATGTTGAAGG-3'

Reverse 5'-GGTTGCTTGTTTATCACCTGTGTC-3'

BRCA “N”

Forward 5'-TAACTCGAGATGGCCGCGGACTC-3'

Reverse 5'-GGTTGCTTGTTTATCACCTGTGTC-3'

BRCA “R”

Forward 5'-TAACTCGAGATGGCCGCGGACTC-3'

Reverse 5'-AAATTCTTTACTGCAGGTTTTTTCC-3'

BRCA2 “C”:

Forward 5'-AAATTATCAAATAACTTAAATGTTGAAGG-3'

Reverse 5'-GCGGCCGCCGGGCCCCCTGGAAC-3'

The constructs were then transiently transfected into human embryonic kidney 293T cells and the proteins were purified as described (Jensen et al., 2010). In short, 30 15 cm plates containing HEK 293T cells were brought up to 70% confluence in DMEM + 10% FBS (Invitrogen). The cells were transfected with 10 µg of the plasmid/ 10×10^6 cells using calcium phosphate co-precipitation method (Sambrook and Russell, 2006) and harvested 31 hours post-transfection. Cells were harvested in 300 ml Lysis Buffer (50 mM HEPES (pH 7.5), 250 mM NaCl, 5 mM EDTA, and 1 mM DTT, 1% Igepal CA-630, 3 mM $MgCl_2$, 1 mM ATP, 1 mM PMSF and Protease Inhibitor Cocktail (Roche)). After rotating the cell lysate for 15 minutes the mixture was spun down at 10,000 g for 15 minutes in Sorval centrifuge, the supernatant collected and incubated with 6 ml amylose resin (New England Biolabs) overnight. The resin was then centrifuged at 2,000 g in a swinging bucket rotor (JS 5.3 Beckman), washed one time with Wash Buffer (50 mM HEPES pH 7.5, 250 mM NaCl, 0.5 mM EDTA and 1 mM DDT), packed into disposable plastic columns (Pierce) and washed extensively with Wash Buffer. The protein was eluted with 10 mM maltose in Elution Buffer (50 mM HEPES (pH 7.5), 250 mM NaCl, 0.5 mM EDTA, 10% glycerol, and 1 mM DTT) and loaded onto HiTrap Q column (GE Life Sciences) and eluted with Storage Buffer (50 mM HEPES pH 8.2, 0.5 mM EDTA, 10% Glycerol, 450 mM NaCl and 1 mM DTT). The identity of the protein was confirmed by

Western Blot: for FL BRCA2, BRCA2-NR, BRCA2-RC and BRCA2-R using an anti-BRC repeats antibody (Calbiochem), for BRCA2-N anti-N-terminal antibody (Bio-vision) and an anti-C-terminal antibody (Calbiochem) for BRCA2-C. 20 μ l aliquots were stored in -80°C.

Human RAD51 was expressed and purified as described (Henricksen et al., 1994).

BRCA2-RAD51 complex formation

The BRCA2-R/RAD51 complex formation was performed in the conditions used for BRCA2 FL/RAD51 complex formation (Sanchez et al., 2017). 50nM BRCA2-R was mixed with 20 μ M RAD51 in total volume of 20 μ l of Binding Buffer (12.5 mM HEPES, pH 8.2,

4 mM Tris-HCl, pH 7.5, 3.25% glycerol, 125 mM NaCl, 65 mM KCl, 0.4 mM DTT, 0.17 mM EDTA) and incubated for 30 minutes at 37°C.

SFM analysis

Protein preparations were diluted four times with HEPES pH 8.2 and incubated for 30 minutes at 37°C or 0°C. 50 μ M spermidine was added as binding agent and the protein deposited on freshly cleaved mica for 1 min, washed three times with MilliQ water and dried with filtered air. Samples were imaged in air at room temperature by tapping mode SFM using a Nanoscope III or Nanoscope IV (Digital Instruments). Silicon Pointprobe tips were type NHC-W, resonance frequency 310–372 kHz (Nanosensors) Images were collected at 2 μ m \times 2 μ m, 512 \times 512 pixels, and

processed only by flattening to remove background slope. Images were analyzed using SFMetrics software tool (Humberto).

References

Davies, A.A., Masson, J.Y., McIlwraith, M.J., Stasiak, A.Z., Stasiak, A., Venkitaraman, A.R., and West, S.C. (2001). Role of BRCA2 in control of the RAD51 recombination and DNA repair protein. *Mol Cell* 7, 273-282.

Henricksen, L.A., Umbricht, C.B., and Wold, M.S. (1994). Recombinant replication protein A: expression, complex formation, and functional characterization. *J Biol Chem* 269, 11121-11132.

Jensen, R.B., Carreira, A., and Kowalczykowski, S.C. (2010). Purified human BRCA2 stimulates RAD51-mediated recombination. *Nature* 467, 678-683.

Liu, J., Doty, T., Gibson, B., and Heyer, W.D. (2010). Human BRCA2 protein promotes RAD51 filament formation on RPA-covered single-stranded DNA. *Nat Struct Mol Biol* 17, 1260-1262.

Nathanson, K.L., Wooster, R., and Weber, B.L. (2001). Breast cancer genetics: what we know and what we need. *Nat Med* 7, 552-556.

Nott, T.J., Petsalaki, E., Farber, P., Jervis, D., Fussner, E., Plochowietz, A., Craggs, T.D., Bazett-Jones, D.P., Pawson, T., Forman-Kay, J.D., et al. (2015). Phase transition of a disordered

nuage protein generates environmentally responsive membraneless organelles. *Mol Cell* 57, 936-947.

Oliver, A.W., Swift, S., Lord, C.J., Ashworth, A., and Pearl, L.H. (2009). Structural basis for recruitment of BRCA2 by PALB2. *EMBO Rep* 10, 990-996.

Pellegrini, L., Yu, D.S., Lo, T., Anand, S., Lee, M., Blundell, T.L., and Venkitaraman, A.R. (2002). Insights into DNA recombination from the structure of a RAD51-BRCA2 complex. *Nature* 420, 287-293.

Rajendra, E., and Venkitaraman, A.R. (2010). Two modules in the BRC repeats of BRCA2 mediate structural and functional interactions with the RAD51 recombinase. *Nucleic Acids Res* 38, 82-96.

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., et al. (2015). BRCA2 diffuses as oligomeric clusters with RAD51 and changes mobility after DNA damage in live cells. *J Cell Biol* 208, 857.

Sambrook, J., and Russell, D.W. (2006). Calcium-phosphate-mediated Transfection of Eukaryotic Cells with Plasmid DNAs. *CSH Protoc* 2006.

Sanchez, H., Paul, M.W., Grosbart, M., van Rossum-Fikkert, S.E., Lebbink, J.H.G., Kanaar, R., Houtsmuller, A.B., and Wyman, C. (2017). Architectural plasticity of human BRCA2-RAD51 complexes in DNA break repair. *Nucleic Acids Res* 45, 4507-4518.

Shahid, T., Soroka, J., Kong, E., Malivert, L., McIlwraith, M.J., Pape, T., West, S.C., and Zhang, X. (2014). Structure and mechanism of action of the BRCA2 breast cancer tumor suppressor. *Nat Struct Mol Biol* 21, 962-968.

Sharan, S.K., Morimatsu, M., Albrecht, U., Lim, D.S., Regel, E., Dinh, C., Sands, A., Eichele, G., Hasty, P., and Bradley, A. (1997). Embryonic lethality and radiation hypersensitivity mediated by Rad51 in mice lacking Brca2. *Nature* 386, 804-810.

Smith, M.D., and Jelokhani-Niaraki, M. (2012). pH-induced changes in intrinsically disordered proteins. *Methods Mol Biol* 896, 223-231.

Thorslund, T., McIlwraith, M.J., Compton, S.A., Lekomtsev, S., Petronczki, M., Griffith, J.D., and West, S.C. (2010). The breast cancer tumor suppressor BRCA2 promotes the specific targeting of RAD51 to single-stranded DNA. *Nat Struct Mol Biol* 17, 1263-1265.

Uversky, V.N. (2016). Dancing Protein Clouds: The Strange Biology and Chaotic Physics of Intrinsically Disordered Proteins. *J Biol Chem* 291, 6681-6688.

Wong, A.K., Pero, R., Ormonde, P.A., Tavtigian, S.V., and Bartel, P.L. (1997). RAD51 interacts with the evolutionarily conserved BRC motifs in the human breast cancer susceptibility gene brca2. *J Biol Chem* 272, 31941-31944.

Yang, H., Jeffrey, P.D., Miller, J., Kinnucan, E., Sun, Y., Thoma, N.H., Zheng, N., Chen, P.L., Lee, W.H., and Pavletich, N.P. (2002). BRCA2

function in DNA binding and recombination from a BRCA2-DSS1-ssDNA structure. *Science* 297, 1837-1848.

Appendix

Summary

Proteins, the working tools of each and every cell, are able to carry out and catalyse all processes essential for cellular metabolism and survival. Every protein is equipped with a set of specific properties, e.g. catalytic or structural features, that enable them to fulfil specific tasks necessary for the proper functioning of the cell. Moreover, proteins and peptides can assemble in intricate, multicomponent arrays, that further increases the efficiency and specificity of performed tasks, thus functioning as biological nanomachines. As in every machine, or tool, structure and structural dynamics are inseparable with protein' function. With the recent developments in techniques utilized to study protein structure, advances in understanding protein function had been remarkable, however, big, flexible proteins prove difficult to study and very often remain outside the scope of conventional biochemical investigations. The aim of this thesis was to employ Scanning Force Microscopy (SFM) to investigate structure-to-function relations of proteins which size and structural dynamics are limiting factors in most biochemical approaches.

Chapter 1 is an introductory outline, where the relationship between DNA and proteins, and the concept of a nanomachine is presented. DNA damage repair as a process of central interest in this thesis is briefly explained. The concept of proteins as tools is then discussed and the importance of structure-to-function relations explained in detail. Finally, various methods of studying protein architecture and structural dynamics are presented and contrasted with SFM.

Chapter 2 investigates the structure of bacterial condensin MukB, member of the Structural Maintenance of Chromosomes (SMC) family. While the overall structure of this protein was known from Electron Microscopy imaging, information on structural dynamics and, in fact, mechanism of DNA condensation, was scarce. Using SFM imaging we observed relatively big dynamics within the protein' globular domains and structural constriction within the hinge domain. By imaging DNA-MukB complexes in different stages of protein saturation we were able to determine four distinctive DNA-binding modes. Moreover, we identified a new DNA-binding interface located within the hinge region of MukB. Overall, we confirmed that DNA condensation is a step-wise process and define this process topologically.

Chapter 3 focuses on another member of the SMC family, the MRE11/RAD50 (MR) complex. Intensively researched for more than three decades, the role of this protein in Homologous Recombination (HR) DNA repair pathway remains elusive. A versatile enzyme, with confirmed functions in early break detection and checkpoint activation, DNA resection and end tethering, it is a large protein complex with elongated, flexible coiled coil protrusions. We introduced point mutation into the MR' ATPase domain, hampering either ATP binding or hydrolysis, and imaged the complex with SFM in the presence and absence of the nucleotide. Our results show that both ATP binding and hydrolysis is crucial for MR oligomerization. Also, by observing the changes in MR' structural dynamics when unable to process ATP, we speculate that ATP-driven conformation changes crucial for proper functioning of the complex comprise of a series of dynamic steps,

rather than one rapid event. Further imaging of MR variants in the presence and absence of ATP and various DNA substrates further proved, that ATP hydrolysis, not only binding, is required for DNA tethering. Additionally, we revealed a new conformation state that corresponds to ssDNA processing and possibly signaling functions.

Chapter 4 is an updated, expanded, step-by step manual on state-of-the art imaging technique developed by our group, TIRF-SFM. In brief, this method is a combination of optical single-molecule Total Internal Reflection Fluorescence(TIRF) microscopy and Scanning Force Microscopy (SFM). By particular sample preparation, carefully placing and tuning the two imaging systems together, and eliminating interfering vibrations, it is possible to overlay the optical and topographical data, thus obtaining detailed information on both the sample's spatial characteristics and composition.

Chapter 5 tackles one of the most mysterious DNA repair proteins, BRCA2. Large, flexible and structurally disordered, BRCA2 has baffled structural biologists for decades. By employing SFM, we show that BRCA2 is a dynamic entity, which can exist in many conformational states. We investigate rapid structural rearrangements upon environmental changes and interaction with substrates—DNA and RAD51 recombinase. Combined TIRF-SFM microscopy allowed us to examine BRCA2-DNA/BRCA2-RAD51 complexes in great detail and reconstruct the mechanism of BRCA2-mediated RAD51 loading on ssDNA. Finally, super-resolution light microscopy showed distinct, separate locations of BRCA2 and RAD51 within a cell nucleus at different repair stages.

Chapter 6 is a continuation of the single-molecule imaging based research into BRCA2 architecture. In the previous chapter we show the prevalent oligomeric state of BRCA2 and drastic reorganizations of oligomers upon environmental changes and ligand binding. The focus of this work is to investigate which regions of BRCA2 are responsible for its seemingly vital propensities: flexibility and structural dynamics. We roughly divided BRCA2 protein into three essential parts: the N-terminal (N), the BRC-repeats (R) and the C-terminal (C) domains and created truncation variants missing one or two of these domains. SFM imaging of these proteins points to multiple interaction interfaces on various parts of the protein that all contribute to formation of irregular, elongated BRCA2 oligomers. We concluded that while all of the three domains seem to contribute to oligomer formation, the R domain is the most crucial one in that respect. The N domain is most likely responsible for the extended parts of the BRCA2 complexes. The dynamic changes upon interaction of the R fragment with RAD51 pointed to direct involvement of the N and C domains into the BRCA2-RAD51 filament formation.

Samenvatting

.Eiwitten, het functionele gereedschap van elke cel, zijn in staat alle processen die nodig zijn voor celmetabolisme en –overleving uit te voeren. Ieder eiwit heeft een aantal specifieke eigenschappen, bijvoorbeeld katalytische of structurele kenmerken, die het in staat stelt een specialistische taak te vervullen, wat nodig is voor het goed functioneren van de cel. Verder kunnen eiwitten en peptiden zich samenvoegen tot complexe, multi-componente structuren, die deze taak nóg efficiënter en preciezer kunnen uitvoeren, in feite werkend als biologische nanomachines. Zoals bij iedere machine of stuk gereedschap, zijn de structuur en structurele dynamiek van een eiwit onlosmakelijk verbonden met diens functie. Met behulp van de meest recente technische ontwikkelingen, gebruikt voor het bestuderen van eiwit -structuren, is opmerkelijke voortgang geboekt bij het begrijpen van eiwit-functies. Echter, grote, flexibele eiwitten blijken moeilijk te bestuderen en vallen vaak buiten het blikveld van conventioneel biochemisch onderzoek. Het doel van dit proefschrift is het gebruik van Scanning Force Microscopy (SFM) ter bestudering van structuur-naar-functie relaties van eiwitten waarvan de grootte en structurele dynamiek een obstakel vormen voor het gebruik van de meeste biochemische methodes.

Hoofdstuk 1 is een inleidend overzicht, waarin zowel de relatie tussen DNA en eiwitten, als het concept van de nanomachine wordt uiteengezet. Vervolgens wordt het proces van het herstellen van DNA-schade, van het hoogste belang voor dit proefschrift, kort en bondig besproken. Ook wordt het concept van eiwitten als gereedschap beschreven en het belang van

structuur-naar-functie relaties in detail uitgelegd. Tenslotte worden verschillende methodes ter bestudering van de architectuur en structurele dynamiek van eiwitten besproken en vergeleken met SFM.

Hoofdstuk 2 onderzoekt de structuur van het bacteriële condensin MukB, dat deel uitmaakt van de 'Structural Maintenance of Chromosomes' (SMC) familie. Hoewel de algemene structuur van dit eiwit reeds bekend is door middel van elektronenmicroscopie, is informatie over de structurele dynamiek en het mechanisme van DNA-condensatie, ervan zeer beperkt. Met het gebruik van SFM hebben wij een relatief grote dynamiek binnen het 'globular domain' van het eiwit MukB, alsook structurele constrictie binnen het 'hinge domain' waargenomen. Het in beeld brengen van DNA-MukB complexen in vier verschillende stadia van proteïne-verzading stelde ons in staat vier verschillende vormen van DNA-binding vast te stellen. Bovendien hebben wij een nieuwe DNA-bindende interface geïdentificeerd in het 'hinge domain' van MukB. Samenvattend, hebben wij kunnen bevestigen dat DNA-condensatie een stapsgewijs proces is, en hebben dit proces topologisch kunnen beschrijven.

Hoofdstuk 3 focust op een ander lid van de SMC familie: het MRE11/RAD50 (MR) complex. Ondanks meer dan drie decennia intensief onderzoek, blijft de rol van dit eiwitcomplex in het 'Homologous Recombination' (HR) DNA-reparatie proces enigszins ongrijpbaar. Dit veelzijdige enzym, met bekende functies in 'early break detection', 'checkpoint activation', DNA-resectie en tethering, is een groot eiwitcomplex met uitgestrekte, flexibele 'coiled coil' ledematen. Wij introduceerden een puntmutatie in MR's ATPase domein, om ófwel ATP-binding ófwel hydrolyse te

verhinderen, en scanden het complex met SFM. Zowel in de aanwezigheid als de afwezigheid van het nucleotide. Onze resultaten laten zien dat zowel ATP-binding als hydrolyse cruciaal zijn voor MR-oligomerisatie. Voorts – door het observeren van de veranderingen in MR's structurele dynamica wanneer het niet in staat is ATP te verwerken – kunnen wij vaststellen dat ATP-gerichte conformatieveranderingen cruciaal zijn voor het goede verloop van een complex samenspel van stappen, in plaats van een enkele plotselinge gebeurtenis. Verdere scans van MR-varianten, in de aan- en afwezigheid van ATP alsook verschillende DNA-substraten, heeft verder aangetoond dat behalve ATP-binding, ook –hydrolyse nodig is voor DNA tethering. Voorts hebben wij een nieuwe conformatie-staat kunnen onthullen die overeenkomt met ssDNA processing en mogelijke signalering-functies.

Hoofdstuk 4 is een up-to-date, uitgebreide, stap-voor-stap handleiding voor een innovatieve imaging-techniek ontwikkeld door ons team: TIRF-SFM. Kortweg is deze methode een combinatie van optische, op één cel gerichte, Total Internal Reflection Fluorescence (TIRF) microscopie en Scanning Force Microscopie (SFM). Door het speciaal prepareren van het sample, het voorzichtig samen plaatsen en afstemmen van de twee imaging-systemen en het elimineren van interfererende trillingen, is het mogelijk de optische en topografische data over elkaar te leggen, waarmee gedetailleerde informatie over zowel de ruimtelijke karakteristieken als de samenstelling van de sample wordt verkregen.

Hoofdstuk 5 tackelt één van de meest enigmatische DNA hersteleiwitten: BRCA2. De omvang, flexibiliteit en structurele

vormloosheid van BRCA2 stelt structurele biologen al decennia lang voor raadsels. Met het gebruik van SFM hebben wij kunnen aantonen dat BRCA2 een dynamische entiteit is, die in een veel verschillende conformaties kan voorkomen. Wij onderzoeken snelle structurele herschikkingen naar aanleiding van veranderingen in het milieu en de reactie met substraten (DNA en RAD51 enzym). Gecombineerde TIRF-SFM microscopie stelt ons in staat BRCA2-DNA/BRCA2-RAD51 complexen in detail te onderzoeken, alsook het mechanisme van het laden van RAD51 op ssDNA door middel van BRCA2 te reconstrueren. Tenslotte heeft super-resolutie optische microscopie laten zien dat BRCA2 en RAD51 zich op bepaalde, verschillende locaties binnen de celkern bevinden gedurende de diverse stadia van het herstelproces.

Hoofdstuk 6 is een voortzetting van het onderzoek door middel van het in beeld brengen van een enkel molecuul in het onderwerp van de architectuur van BRCA2. In het voorgaande hoofdstuk lieten wij de meest voorkomende oligomerische staat van BRCA2 en de drastische reorganisatie van oligomeren, na veranderingen in het milieu en ligand-verbinding, zien. Het doel van dit werk was erachter komen welke delen van BRCA2 verantwoordelijk zijn voor haar schijnbaar vitale kenmerken: flexibiliteit en structurele dynamica. Wij verdeelden het BRCA2-eiwit in ruwweg drie essentiële delen: het N-terminale (N), de BRC-repeats (R), en het C-terminale (C) domein, en creëerden ingekorte varianten waarin één of twee van deze domeinen ontbreken. SFM-scans van deze eiwitten wijzen op meerdere interactie-interfaces op verschillende delen van het eiwit die allemaal aan de vorming van onregelmatige, langwerpige BRCA2 oligomeren. Wij hebben geconcludeerd dat hoewel alle drie de

domeinen aan het creëren van oligomeren lijken bij te dragen, het R-domein in dit opzicht het meest cruciaal is. Het N-domein is waarschijnlijk verantwoordelijk voor de uitstekende delen van BRCA2-complexen. De dynamische veranderingen bij interactie van het R-fragment met RAD51 wijst erop dat het N- en C-domein direct betrokken zijn bij de BRCA2-RAD51 filament-formatie.

CURRICULUM VITAE

Personal details

Name: Małgorzata Grosbart

Date of birth: November 10th, 1985

Place of birth: Warsaw, Poland

Education

2011-2013 MSc, Erasmus University Medical Center, Molecular Medicine research Master. Thesis title: "Effects of ATP on functional architecture of the DNA repair complex MRE11/RAD50".

2004-2009 MSc Biophysics, University of Łódź, Faculty of Biology and Environmental Protection. Thesis title: "The antioxidant effects of Tamoxifen and Raloxifene in erythrocyte membranes against physiological oxidants hypochlorite, peroxynitrite and peroxy radical".

Professional training

2012-2016 Erasmus University Medical Center, PhD position in the department of Molecular Genetics

2011-2012 The Hague University of Applied Sciences, laboratory and teaching assistant

2009-2011 University hospital USK-WAM, Łódź, laboratory technician

PORTFOLIO

Name: Małgorzata Grosbart

PhD period: June 2012 - December 2016

Department: Molecular Genetics

Promoters: Prof. dr. Claire Wyman

Prof. dr. Roland Kanaar

In-depth courses

In Vivo Cellular Imaging, 2012

Workshop on Photoshop and Illustrator CS5, 2013

PhD-course Genetics, 2013

PhD-course Cell and Developmental Biology, 2014

PhD-course Biophysics and Biochemistry, 2014

Research Management for PhD students and Postdocs, 2015

Seminars and Workshops

Department weekly work discussion, 2012-2016

Department journal club, 2012-2016

Molecular Genetics cluster weekly work discussion, 2012-2016

MGC PhD Workshop, 2012 poster presentation, 2013-2016 talks

Dutch SPM day, 2013

NWO-nano day, 2013, 2014

International Conferences

Dutch Biophysics—The Annual Dutch Meeting on Molecular and Cellular Biophysics, Veldhoven, The Netherlands, 2012, 2013, 2015, 2016 poster presentations, 2016 talk

CHAINS—Chemistry as Innovating Science, Veldhoven, The Netherlands, 2014 poster presentation

25th MGC Symposium, Leiden, The Netherlands, 2015 talk

"DNA transactions" 14th meeting on the Site-Specific Recombination, Transposition and DNA Dynamics, Orleon, France, 2014 poster presentation

Responses to DNA damage: from molecule to disease, Egmond aan Zee, The Netherlands, 2016 poster presentation

ABCAM conference on Mechanisms of recombination, Alicante, Spain, 2016 poster presentation

List of publications

Grosbart, M., Sánchez, H., Ristic, D., Wyman, C. (2018) Imaging of DNA and protein by SFM and combined SFM-TIRF microscopy, In: Single Molecule Analysis: Methods and Protocols, Second Edition, ed. Peterman E.J.G., Springer Science + Business Media, New York, NY USA.

Sánchez, H., Paul, M.W., Grosbart, M., van Rossum-Fikkert, S.E., Lebbink, J.H., Kanaar, R., Houtsmuller, A.B. and Wyman, C. (2107) Architectural plasticity of human BRCA2-RAD51 complexes in DNA break repair, *Nucleic Acids Res*, 545(8):4507-4518

Kumar M., Grosbart, M., Nurse, P., Bhang, S., Wyman, C.L., and Mariani, K.J. (2017) The bacterial condensin MukB compacts DNA by sequestering supercoils and stabilizing topologically isolated loops, *J Biol Chem*, 292(41):16904-16920

Kim, J.H., Grosbart, M., Anand, R., Wyman, C., Cejka, P. and Petrini, J.H. (2017) The Mre11-Nbs1 Interface Is Essential for Viability and Tumor Suppression, *Cell Rep*, 18(2):496-507

Dinant, C., Ampatzidis-Michailidis, G., Lans, H., Tresini, M., Lagarou, A., Grosbart, M., Theil, A.F., van Cappellen, W.A., Kimura, H., Bartek, J., Foustieri, M., Houtsmuller, A.B., Vermeulen, W. and Marteijn, J.A. (2013) Enhanced chromatin dynamics by FACT promotes transcriptional restart after UV- induced DNA damage, *Mol Cell* 51(4):469-79

Acknowledgments

So many people took part in, contributed to, and helped in the work that this book is a result and culmination of, that I can not possibly aspire to include them all within these pages. Knowing that the list of thanks will be, by necessity, incomplete, I may as well keep it short.

First and foremost, to my promoter, colleague and mentor Claire Wyman, for allowing and helping me to complete this body of work in her group, and for the fact that her only fault was that she didn't treat me as harsh as my laziness and disorganization would call for. I'm pretty sure I'll never have a boss as supportive and inspiring again.

The rest of senior staff in the Biochemistry/Biophysics team: Roland Kanaar, Joyce Lebbink, Dik van Gent, as well as colleagues that accompanied me and my work during these few years, I am forever grateful for all the help and insightful conversations regarding my research projects in particular, science in general and other subjects altogether. Special mention and gratitude is due to Humberto Sánchez who has taught me more about science, lab work and nanobiological techniques than ten years of studies ever could.

To Wim Vermeulen and his lab, especially Jurgen Martijn, big thanks for pulling me into the world of microscopy and introducing me into the realities of serious research. The year I've spent under your supervision was priceless in many aspects.

To my grandmother, Stanisława Grosbart, for being literally the only person to believe in me back when it all begun, fifteen years ago.

Things wouldn't have gone half as smooth if it wasn't for my wonderful parents-in-law, Oda and Willie. Thank you for everything, material and immaterial.

Beata Cymerman, my best friend—thank you for being with me, staying by my side, regardless of distance or events that could have so easily came between us. It is truly something, the bond we share.

Finally and, perhaps, most importantly, to my partner and friend, Niels. To have your loving support made me proud and joyous during the good times, and sustained me when life was at its most difficult. *jak kwiatem, każdym słowem twym się pieczę*