

associated with DNA methylation in *Neurospora*, but it is unaffected in *dim-2* DNA methyltransferase mutants (67). Instead, genes encoding components of the RNAi machinery are required (68, 69). Our results would be consistent with these observations, if histone H3 K9 methylation were the direct consequence of RNAi, but DNA methylation is only an indirect consequence.

Thus, our results provide a possible link between RNAi and DNA methylation. dsRNA derived from repeated sequences might trigger RNAi, which would then initiate histone H3 lysine-9 methylation. Histone modification would then signal DNA methylation. This mechanism could guide eukaryotic DNA methyltransferases to specific regions of the genome, such as transposable elements, even though they have little sequence specificity in themselves (4). Such an arrangement could be reinforced by maintenance methyltransferase activity, as well as by the deacetylation of histones guided by methyl DNA binding complexes (70).

#### References and Notes

1. S. Henikoff, *Curr. Opin. Genet. Dev.* **2**, 907 (1992).
2. S. I. Grewal, S. C. Elgin, *Curr. Opin. Genet. Dev.* **12**, 178 (2002).
3. L. L. Wallrath, *Curr. Opin. Genet. Dev.* **8**, 147 (1998).
4. R. A. Martienssen, V. Colot, *Science* **293**, 1070 (2001).
5. M. K. Montgomery, S. Xu, A. Fire, *Proc. Natl. Acad. Sci. U.S.A.* **95**, 15502 (1998).
6. R. Jorgensen, *Trends Biotechnol.* **8**, 340 (1990).
7. E. Bernstein, A. A. Caudy, S. M. Hammond, G. J. Hannon, *Nature* **409**, 363 (2001).
8. A. J. Hamilton, D. C. Baulcombe, *Science* **286**, 950 (1999).
9. S. M. Hammond, E. Bernstein, D. Beach, G. J. Hannon, *Nature* **404**, 293 (2000).
10. T. Tuschl, P. D. Zamore, R. Lehmann, D. P. Bartel, P. A. Sharp, *Genes Dev.* **13**, 3191 (1999).
11. S. M. Hammond, S. Boettcher, A. A. Caudy, R. Kobayashi, G. J. Hannon, *Science* **293**, 1146 (2001).
12. M. Fagard, S. Boutet, J. B. Morel, C. Bellini, H. Vaucheret, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 11650 (2000).
13. T. Sijen *et al.*, *Cell* **107**, 465 (2001).
14. C. Lipardi, Q. Wei, B. M. Paterson, *Cell* **107**, 297 (2001).
15. T. Dalmay, A. Hamilton, S. Rudd, S. Angell, D. C. Baulcombe, *Cell* **101**, 543 (2000).
16. F. E. Vaistij, L. Jones, D. C. Baulcombe, *Plant Cell* **14**, 857 (2002).
17. C. Kidner, R. Martienssen, in preparation.
18. K. Lynn *et al.*, *Development* **126**, 469 (1999).
19. V. Wood *et al.*, *Nature* **415**, 871 (2002).
20. L. Aravind, H. Watanabe, D. J. Lipman, E. V. Koonin, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 11319 (2000).
21. Materials and methods are available as supporting online material on Science Online.
22. A. J. Klar, M. J. Bonaduce, *Cold Spring Harb. Symp. Quant. Biol.* **58**, 457 (1993).
23. K. Ekwall, G. Cranston, R. C. Allshire, *Genetics* **153**, 1153 (1999).
24. N. C. Steiner, K. M. Hahnenberger, L. Clarke, *Mol. Cell. Biol.* **13**, 4578 (1993).
25. K. Takahashi, S. Murakami, Y. Chikashige, O. Niwa, M. Yanagida, *J. Mol. Biol.* **218**, 13 (1991).
26. R. C. Allshire, E. R. Nimmo, K. Ekwall, J. P. Javerzat, G. Cranston, *Genes Dev.* **9**, 218 (1995).
27. K. Ekwall *et al.*, *J. Cell. Sci.* **109**, 2637 (1996).
28. R. C. Allshire, J. P. Javerzat, N. J. Redhead, G. Cranston, *Cell* **76**, 157 (1994).
29. A. A. Aravin *et al.*, *Curr. Biol.* **11**, 1017 (2001).
30. A. Schmidt *et al.*, *Genetics* **151**, 749 (1999).
31. B. Fishel, H. Amstutz, M. Baum, J. Carbon, L. Clarke, *Mol. Cell. Biol.* **8**, 754 (1988).
32. T. Volpe, G. Teng, R. Martienssen, unpublished data.
33. T. Dalmay, A. Hamilton, E. Mueller, D. C. Baulcombe, *Plant Cell* **12**, 369 (2000).
34. L. Jones *et al.*, *Plant Cell* **11**, 2291 (1999).
35. M. A. Matzke, A. J. Matzke, G. J. Pruss, V. B. Vance, *Curr. Opin. Genet. Dev.* **11**, 221 (2001).
36. J. Nakayama, J. C. Rice, B. D. Strahl, C. D. Allis, S. I. Grewal, *Science* **292**, 110 (2001).
37. K. Noma, C. D. Allis, S. I. Grewal, *Science* **293**, 1150 (2001).
38. S. Rea *et al.*, *Nature* **406**, 593 (2000).
39. S. I. Grewal, A. J. Klar, *Genetics* **146**, 1221 (1997).
40. I. M. Hall, S. Grewal, unpublished data.
41. A. H. Peters *et al.*, *Cell* **107**, 323 (2001).
42. K. Takahashi, E. S. Chen, M. Yanagida, *Science* **288**, 2215 (2000).
43. B. Reinhart, D. Bartel, *Science* **297**, 1831; published online 22 August 2002 (10.1126/science.1077183).
44. C. Catalanotto, G. Azzalin, G. Macino, C. Cogoni, *Genes Dev.* **16**, 790 (2002).
45. A. Akhtar, D. Zink, P. B. Becker, *Nature* **407**, 405 (2000).
46. M. Lachner, D. O'Carroll, S. Rea, K. Mechtler, T. Jenuwein, *Nature* **410**, 116 (2001).
47. A. J. Bannister *et al.*, *Nature* **410**, 120 (2001).
48. H. Nakagawa *et al.*, *Genes Dev.* **16**, 1766 (2002).
49. J. F. Partridge, B. Borgstrom, R. C. Allshire, *Genes Dev.* **14**, 783 (2000).
50. M. Baum, L. Clarke, *Mol. Cell. Biol.* **20**, 2852 (2000).
51. S. Jensen, M. P. Gassama, T. Heidmann, *Nature Genet.* **21**, 209 (1999).
52. R. F. Ketting, T. H. Haverkamp, H. G. van Luenen, R. H. Plasterk, *Cell* **99**, 133 (1999).
53. H. Tabara *et al.*, *Cell* **99**, 123 (1999).
54. N. Ayoub, I. Goldshmidt, R. Lyakhovetsky, A. Cohen, *Genetics* **156**, 983 (2000).
55. A.-V. Gendrel, Z. Lippman, C. Yordan, V. Colot, R. Martienssen, *Science* **297**, 1871 (2002).
56. C. Llave, K. D. Kasschau, M. A. Rector, J. C. Carrington, *Plant Cell* **14**, 1605 (2002).
57. B. J. Reinhart, E. G. Weinstein, M. W. Rhoades, B. Bartel, D. P. Bartel, *Genes Dev.* **16**, 1616 (2002).
58. K. Nabeshima *et al.*, *Mol. Biol. Cell.* **9**, 3211 (1998).
59. R. Allshire, personal communication.
60. M. Baum, V. K. Ngan, L. Clarke, *Mol. Biol. Cell.* **5**, 747 (1994).
61. P. Avner, E. Heard, *Nature Rev. Genet.* **2**, 59 (2001).
62. F. Sleutels, R. Zwart, D. P. Barlow, *Nature* **415**, 810 (2002).
63. J. T. Lee, N. Lu, *Cell* **99**, 47 (1999).
64. E. Heard *et al.*, *Cell* **107**, 727 (2001).
65. C. Maison *et al.*, *Nature Genet.* **30**, 329 (2002).
66. H. Tamara, E. U. Selker, *Nature* **414**, 277 (2001).
67. C. Cogoni *et al.*, *EMBO J.* **15**, 3153 (1996).
68. C. Cogoni, G. Macino, *Nature* **399**, 166 (1999).
69. P. K. Shiu, N. B. Raju, D. Zickler, R. L. Metzberg, *Cell* **107**, 905 (2001).
70. A. P. Bird, A. P. Wolffe, *Cell* **99**, 451 (1999).
71. We would like to thank T. Jenuwein, D. Bartel, and R. Allshire for sharing antibodies and unpublished data; J. Nakayama, M. Timmermans, V. Colot, and E. Heard for helpful discussions; and E. Head, B. Lehner, J. Hemish, and R. Krimper for technical help. I.M.H. is Arnold Beckman fellow of the Watson School of Biological Sciences; G.T. is a participant in the Watson School Undergraduate Research Program. This work was supported by an NCI postdoctoral training fellowship to T.V., grants from the Ellison Medical Foundation and NIH (GM59772) to S.G., and a grant from the National Science Foundation (NSF DBI-0077774) to R.M.

#### Supporting Online Material

www.sciencemag.org/cgi/content/full/1074973/DC1  
Materials and Methods

12 June 2002; accepted 9 August 2002

Published online 22 August 2002;

10.1126/science.1074973

Include this information when citing this paper.

## BRCA2 Function in DNA Binding and Recombination from a BRCA2-DSS1-ssDNA Structure

Haijuan Yang,<sup>1</sup> Philip D. Jeffrey,<sup>2</sup> Julie Miller,<sup>2,3</sup>  
Elsbeth Kinnucan,<sup>2</sup> Yutong Sun,<sup>1</sup> Nicolas H. Thomä,<sup>2</sup>  
Ning Zheng,<sup>2,3</sup> Phang-Lang Chen,<sup>4</sup> Wen-Hwa Lee,<sup>4</sup>  
Nikola P. Pavletich<sup>2,3\*</sup>

Mutations in the *BRCA2* (breast cancer susceptibility gene 2) tumor suppressor lead to chromosomal instability due to defects in the repair of double-strand DNA breaks (DSBs) by homologous recombination, but *BRCA2*'s role in this process has been unclear. Here, we present the 3.1 angstrom crystal structure of a ~90-kilodalton *BRCA2* domain bound to DSS1, which reveals three oligonucleotide-binding (OB) folds and a helix-turn-helix (HTH) motif. We also (i) demonstrate that this *BRCA2* domain binds single-stranded DNA, (ii) present its 3.5 angstrom structure bound to oligo(dT)<sub>9</sub>, (iii) provide data that implicate the HTH motif in dsDNA binding, and (iv) show that *BRCA2* stimulates RAD51-mediated recombination *in vitro*. These findings establish that *BRCA2* functions directly in homologous recombination and provide a structural and biochemical basis for understanding the loss of recombination-mediated DSB repair in *BRCA2*-associated cancers.

Germline mutations in *BRCA2* are responsible for a highly penetrant, autosomal dominant predisposition to breast cancer (1–3), and they also confer increased risk of early-onset ovarian, prostate, pancreatic, and male breast cancers (4). In addition, hypomorphic

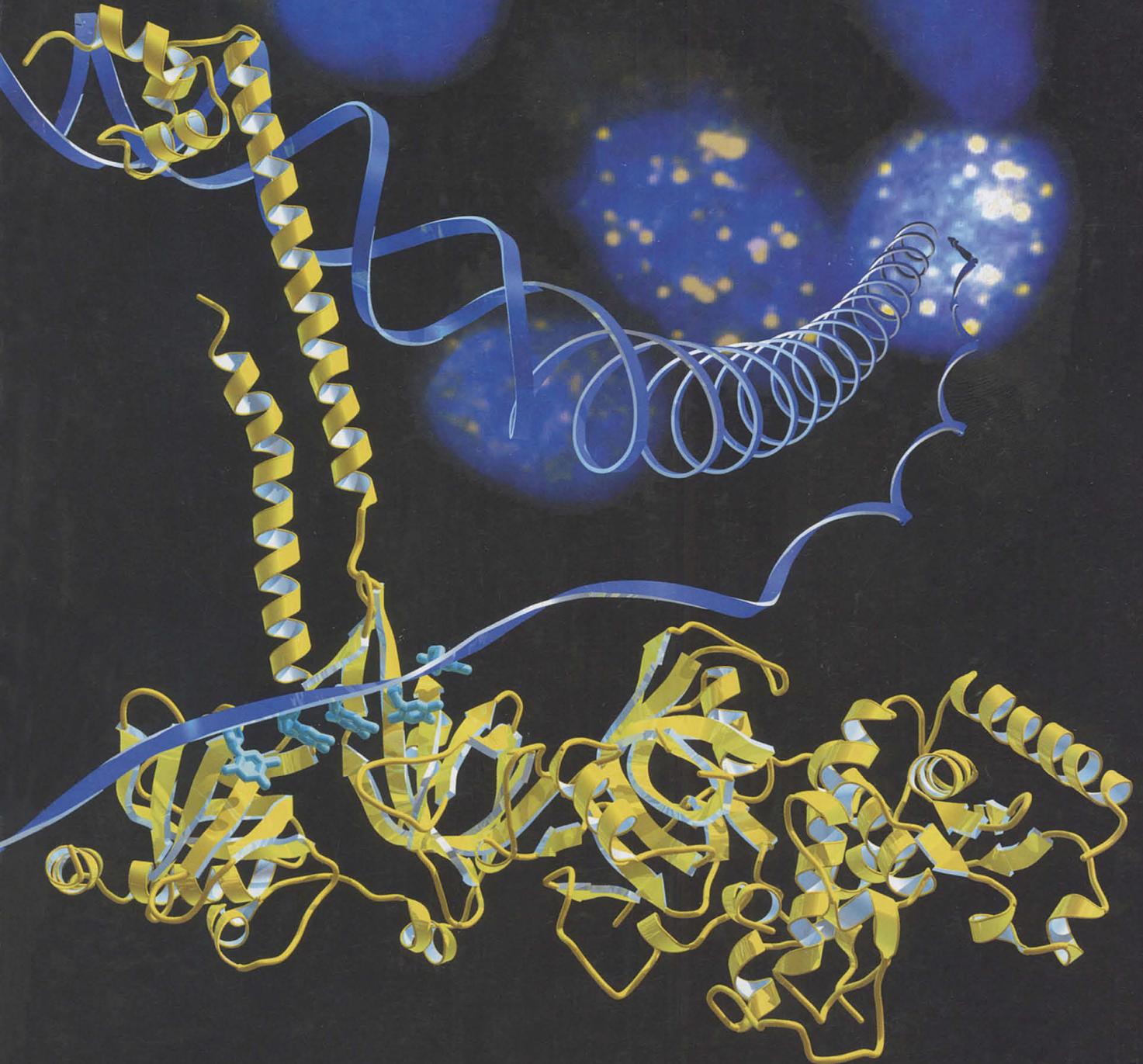
mutations in *BRCA2* have recently been found in cells from patients in the FANCB and FANCD1 subgroups of Fanconi's anemia, an autosomal recessive cancer susceptibility syndrome (5).

A role for *BRCA2* in the maintenance of

13 September 2002

# Science

Vol. 297 No. 5588  
Pages 1757-1944 \$9



AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE

genomic integrity is well established (6, 7). *BRCA2*<sup>-/-</sup> tumor cells and mouse embryo fibroblasts (MEFs) homozygous for hypomorphic *Brca2* alleles show increased sensitivity to ultraviolet light, ionizing radiation, methyl methanesulfonate, and other genotoxic agents (8–18). These cells accumulate chromosomal abnormalities including breaks, aberrant chromatid exchanges, translocations, micronuclei, chromosomal loss, and centrosome amplification (10, 12, 13). This genomic instability is thought to have a causative role in *BRCA2*-associated cancers (19). In support of its role in genomic integrity, deletion of *Brca2* in mice leads to the induction of the p53-p21<sup>Cip1</sup> DNA-damage checkpoint, proliferative defects, and embryonic lethality (8–10, 20, 21).

Recent cellular biological studies have shown that *BRCA2* is needed for the error-free repair of DNA double-strand breaks (DSBs) through homologous recombination, which is a process that maintains chromosomal integrity by repairing the break using the sister chromatid as template (22). Loss of *BRCA2* reduces the efficiency of homologous recombination-mediated DSB repair, resulting in the repair of the DSB lesions through error-prone mechanisms such as non-homologous end joining (NHEJ) or single-strand annealing (SSA) (15, 16, 23). In dividing cells, these repair mechanisms can lead to the chromosomal aberrations and loss of genetic information that are characteristic of *BRCA2*-null cells and other homologous recombination mutants.

A role for *BRCA2* in homologous recombination is supported by many lines of evidence. *BRCA2* physically interacts with *RAD51* (8, 24–26), which is a homolog of the *RecA* protein. *RecA* carries out homologous recombination in *Escherichia coli* and has a central role in DNA repair (27–29). Like *RecA*, *RAD51* can form a nucleoprotein filament with single-stranded DNA (ssDNA) and catalyze the pairing and subsequent exchange of ssDNA with homologous double-stranded DNA (dsDNA) (30–32). After DNA damage, *BRCA2* and *RAD51* colocalize to nuclear foci that are thought to be the sites of damage and repair (33, 34). *Rad51*-knockout mice have an embryonic lethal phenotype characterized by proliferation defects very similar to those of the *Brca2*-null mice (35, 36).

*BRCA2* binds *RAD51* through a partly degenerate 40-amino acid motif (BRC repeat), eight copies of which are interspersed throughout the middle ~1000 residues of *BRCA2* (25, 26, 37). Overexpression of a single BRC repeat in cells leads to decreased homologous recombination (38), radiation hypersensitivity, and loss of G<sub>2</sub>/M checkpoint control, presumably because of dominant interference with the function of endogenous wild-type alleles (26, 39).

The 3418-amino acid *BRCA2* shares no sequence homology with any other protein (1, 2). The BRC-repeat region is followed by an ~1000-residue COOH-terminal region that corresponds to the best-conserved portion of *BRCA2* across dog, mouse, rat, and chicken (40) orthologs [68% average identity for this region, compared with 42% for the entire protein (Fig. 1A)], as well as in putative orthologs found in *Arabidopsis* and rice (Fig. 1A). The COOH-terminal region also contains 27% of the tumor-derived missense mutations in the breast cancer information core (BIC) database (41); this indicates that it has an important role in the tumor-suppressor function of *BRCA2*. This region binds the 70-amino acid DSS1 (deleted in split-hand/split foot syndrome) protein (42), which was originally identified as one of three genes that map to a 1.5-Mb locus deleted in an inherited developmental malformation syndrome (43). DSS1 was independently isolated as a protein that binds *BRCA2* in a yeast two-hybrid screen and in HeLa cells, but the molecular function of DSS1 and the significance of its interaction with *BRCA2* have been unclear.

Here, we present the 3.1 Å crystal structure of a ~800-residue COOH-terminal domain of *BRCA2* bound to DSS1. The structure reveals that *BRCA2* contains multiple domains similar to ssDNA- and dsDNA-binding motifs. We extend these findings by showing that this *BRCA2* COOH-terminal domain, which we name *BRCA2*DBD (DNA/DSS1-binding domain), binds ssDNA in vitro and by presenting a 3.5 Å crystal structure of a *BRCA2*DBD-DSS1-oligo(dT)<sub>n</sub> complex. Our data indicate a direct role for *BRCA2* in homologous recombination, and we further develop this concept by showing that the *BRCA2*DBD stimulates the *RAD51*-mediated strand pairing and exchange reaction in vitro.

**Overall structure of the *BRCA2* COOH-terminal region.** *BRCA2* COOH-terminal polypeptides were coexpressed with DSS1 in Hi5 insect cells by using separate baculoviruses [Materials and Methods (44)]. Coexpression with DSS1 was necessary because the *BRCA2* that was overexpressed in insect cells in the absence of DSS1 was largely insoluble. Limited proteolysis of ~1000-residue human, mouse, and rat *BRCA2* COOH-terminal polypeptides bound to DSS1 was used to optimize the

domain boundaries for crystallization (Fig. 1, B and C). Crystals were obtained with a 736-residue mouse *BRCA2*DBD bound to DSS1 that diffracted to 3.1 Å, and with an 816-residue rat *BRCA2*DBD bound to DSS1 that diffracted to 3.4 Å and contained two complexes in the asymmetric unit (Table 1). Crystals, diffracting up to 3.5 Å resolution, were also obtained with a series of oligo(dT) ssDNA fragments bound to either the intact mouse *BRCA2*DBD-DSS1 complex, or to an engineered *BRCA2*DBD containing an internal deletion (*BRCA2*DBD<sup>ΔTower</sup>-DSS1) (Table 1).

The *BRCA2*DBD has five domains (Fig. 1D). Four of them are arranged in a linear fashion, packing successively and resulting in an extended, 115 Å-long structure. An additional domain protrudes from the core structure by 80 Å. The first domain is 190 amino acids and consists mostly of α helices (helical domain, Fig. 1, B and D). This is followed by three structurally homologous domains that contain the oligonucleotide/oligosaccharide-binding (OB) fold (OB1, OB2, and OB3, Fig. 1, B and D). The ~110-residue OB fold, which is present in most prokaryotic and eukaryotic ssDNA binding proteins (SSBs) (45–47), does not have a well-defined sequence motif associated with it, and its presence in *BRCA2* was not detected previously. OB2 contains a 130-amino acid insertion that adopts a tower-like structure (Tower domain) protruding away from the OB fold (Fig. 1, B and D). The Tower consists of a pair of long, antiparallel helices (the stem) that support a three-helix bundle (3HB) at their end. The 3HB contains the helix-turn-helix (HTH) motif and is similar to the DNA binding domains of the bacterial site-specific recombinases (48, 49) and of the eukaryotic Myb (50) and homeodomain transcription factors (51). The relative orientations of all the domains are very similar in the structures of the mouse and the rat orthologs and in the presence or absence of ssDNA (Fig. 1E).

The 70-residue DSS1 protein binds *BRCA2* in an extended conformation interacting with the helical domain, OB1 and OB2 (Fig. 1D). The binding is characterized by hydrophobic interactions and also by a large number of acidic DSS1 residues interacting with basic grooves on *BRCA2*. *BRCA2*-bound DSS1 does not have a hydrophobic core of its own.

The structure of the *BRCA2*DBD<sup>ΔTower</sup>-DSS1-oligo(dT)<sub>n</sub> complex shows clear electron density for five thymidine nucleotides interacting with OB2 and OB3 (Fig. 1E and fig. S1B). The *BRCA2*-ssDNA interactions are very similar to those seen in the structure of RPA, which is the most abundant eukaryotic SSB, bound to ssDNA (52, 53).

Tumor-derived missense mutations (41) are distributed throughout the five domains, with 25.0% of the *BRCA2*DBD mutations mapping to the helical domain, 25.4% to

<sup>1</sup>Department of Pharmacology, Sloan-Kettering Division, Joan and Sanford I. Weill Graduate School of Medical Sciences, Cornell University, New York, NY 10021, USA. <sup>2</sup>Cellular Biochemistry and Biophysics Program and <sup>3</sup>Howard Hughes Medical Institute, Memorial Sloan-Kettering Cancer Center, New York, NY 10021, USA. <sup>4</sup>Department of Molecular Medicine and Institute of Biotechnology, University of Texas Health Center, San Antonio, TX 78245, USA.

\*To whom correspondence should be addressed. E-mail: nikola@xray2.mskcc.org

OB1, 8.6% to OB2, 8.6% to OB3, and 32.4% to the Tower (Fig. 1, B and F). They map to residues with structural roles, to residues with roles in binding DNA or DSS1, or to residues on the surface of the Tower, indicating that all five domains are important for the tumor suppressor function of BRCA2.

**The BRCA2 OB folds.** The three BRCA2 OB domains contain structures very similar to the canonical OB fold (45), which consists of a highly curved 5-stranded  $\beta$  sheet that closes on itself to form a  $\beta$  barrel ( $\beta 1$  to  $\beta 5$ , Fig. 2A). OB2 and OB3 each have the pronounced groove that is characteristic of the ssDNA binding sites of OB folds. The groove is formed by one face of the curved sheet and is demarcated by two loops, one between  $\beta 1$  and  $\beta 2$  (L12 loop) and another between  $\beta 4$  and  $\beta 5$  (L45 loop, Fig. 2A). The OB2 L12 loop contains the insertion that makes up the Tower domain. The Tower contributes an additional  $\beta$  strand (T $\beta 1$ ) to the groove (pairing with OB2 $\beta 1$ ), and increases the surface area of the OB2 groove. OB1 has a shallower groove, due in part to its L12 loop being disordered, and in part to its L45 loop folding away from the groove. The disordered OB1 L12 loop is  $\sim 20$  residues longer than its average length in OB folds.

The three OB folds of BRCA2 are most similar to the OB folds of RPA, as determined by the three-dimensional structural ho-

mology search program DALI (54), and by inspection of more than 10 protein structures that contain OB folds (Fig. 2A and legend). RPA, which is a heterotrimeric protein, has two OB folds that bind ssDNA with high affinity (DBD-A and -B) and two others (DBD-C and -D) that contribute to its overall affinity for ssDNA (52, 55–57). The closest structural homology is between OB2 of BRCA2 and the DBD-A of RPA. The two structures can be superimposed with a 1.4 Å root-mean-square deviation (rmsd) in the positions of 82 out of 109 C $\alpha$  atoms. The second closest structural homology is that between OB3 of BRCA2 and DBD-B of RPA, with an rmsd of 1.5 Å for 84 C $\alpha$  atoms. These rmsds are lower than those in between the four RPA OB folds. Although OB1 of BRCA2 is structurally more divergent, its closest structural homolog is also one of the RPA OB folds (DBD-D, 1.9 Å for 65 C $\alpha$ ).

The OB folds of the phage T4 gp32 protein (58) and the RPA DBD-C (56) also contain L12 loop insertions. Gp32 has a 45-residue zinc-stabilized helical domain that extends the surface area of its groove, and RPA's DBD-C has a  $\sim 30$ -residue zinc-stabilized ribbon domain that is positioned away from its groove. The precise functions of these insertions are not known.

**Arrangement of the OB folds.** OB2 and OB3 pack in tandem, with their individual

grooves aligned to create a continuous channel. This arrangement is similar in the RPA-ssDNA crystal structure, which consists of the DBD-A and DBD-B OB folds bound to an oligo(dC)<sub>8</sub> (RPA-ssDNA) complex (52), except for two differences. Relative to RPA, the BRCA2 OB3 is rotated by 65° along an axis running along the bottom of the channel, but this does not affect the continuity and overall dimensions of the OB2-OB3 channel (Fig. 2B). In addition, whereas in RPA the two OB folds do not interact in the absence of ssDNA (59), in BRCA2 OB2 and OB3 pack extensively in the absence of DNA, and their arrangement does not change upon ssDNA binding.

The OB2-OB3 packing involves both hydrophobic van der Waals contacts and hydrogen bond networks (1261 Å<sup>2</sup> total surface area buried, Fig. 2C). The residues that form the interface are highly conserved, accounting for 7 of the 26 OB2 and OB3 residues invariant across species (Fig. 1B and fig. S3). In addition to their conservation, the interface residues are also mutated in cancer, with 4% of tumor-derived BRCA2DBD missense mutations mapping to Val<sup>2829</sup>, Arg<sup>2893</sup>, Glu<sup>2921</sup>, and Asn<sup>3042</sup> (Fig. 2C). These observations indicate that the prearrangement of the two OB folds is important for BRCA2 function.

OB1 and OB2 also pack extensively (1271 Å<sup>2</sup> total surface area buried) through

**Table 1.** Statistics from the crystallographic analysis. Details of the crystallization and structure determination are provided in the supporting material (44).  $R_{sym} = \sum_h \sum_i |I_{hi} - I_h| / \sum_h \sum_i I_{hi}$  for the intensity ( $I$ ) of  $i$  observations of reflection  $h$ . Phasing power =  $\langle F_{\lambda i} \rangle / E$  where  $\langle F_{\lambda i} \rangle$  is the root-mean-square heavy atom structure factor and  $E$  is the residual lack of closure error.  $R_{cullis}$  is the mean residual lack of closure error divided by the dispersive or anomalous difference.  $R$ -factor =  $\sum |F_{obs} - F_{calc}| / \sum |F_{obs}|$ , where  $F_{obs}$  and  $F_{calc}$

are the observed and calculated structure factors, respectively.  $R$ -free =  $R$ -factor calculated by using a subset ( $\sim 5\%$ ) of reflection data chosen randomly and omitted throughout refinement. rmsd, root-mean-square deviations from ideal geometry and root-mean-square variation in the  $B$ -factor of covalently bonded atoms (shown only when individual  $B$ -factors were refined). The prefix m before BRCA2DBD-DSS1 indicates mouse, and r, rat ortholog.

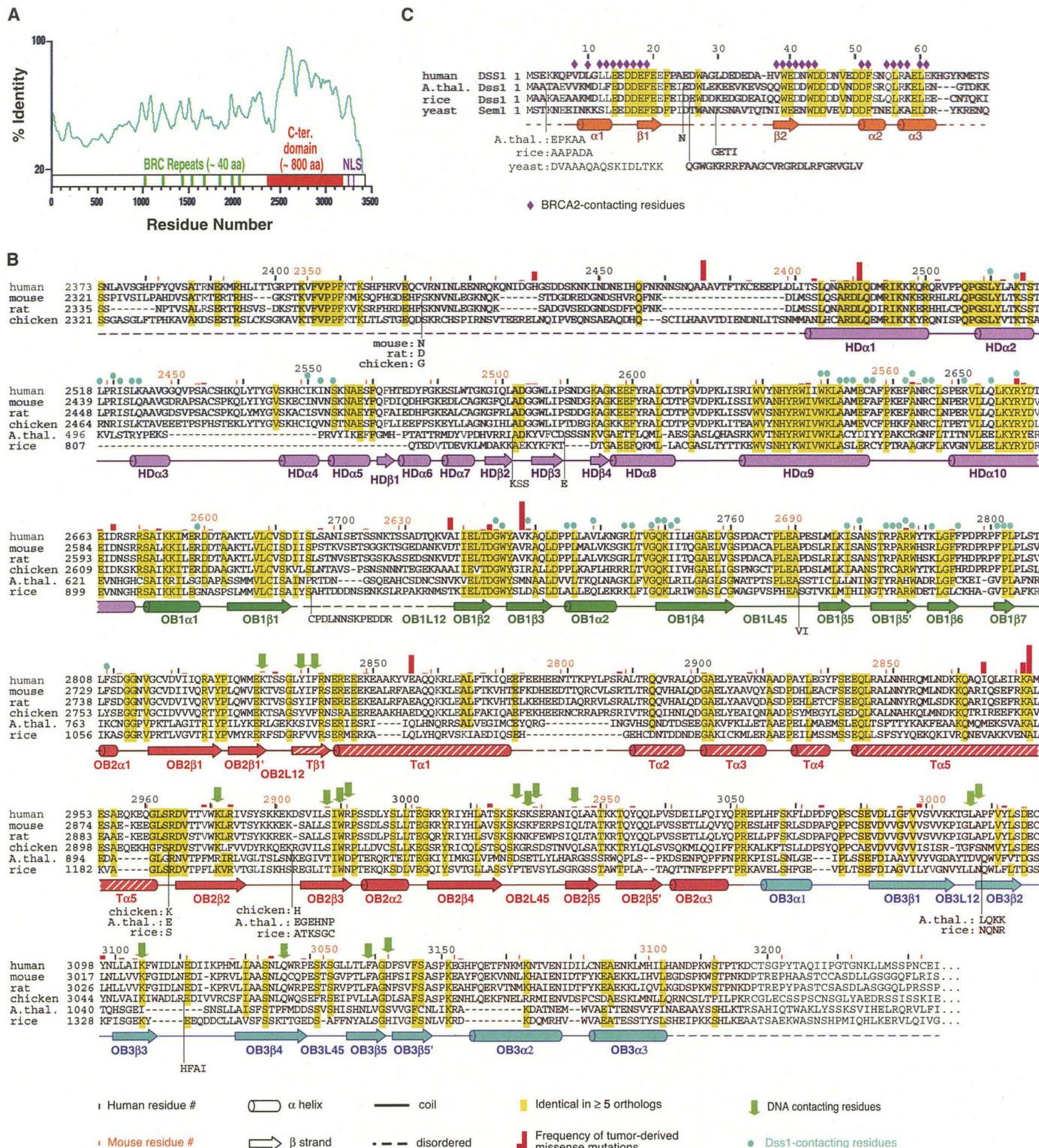
Complex Data set	mBRCA2DBD-DSS1					
	line Hg $\lambda$ 1	Hg $\lambda$ 2	Hg $\lambda$ 3	Native	rBRCA2DBD-DSS1 Native	mBRCA2DBD <sup>Tower</sup> -DSS1-(dT) <sub>9</sub> (5BrDU) <sub>4</sub>
Beam line	NLSX9A	NLSX9A	NLSX9A	CHESSA1	CHESSA1	CHESSA1
Wavelength (Å)	1.010	1.008	0.992	0.950	0.948	0.950
Resolution (Å)	25.0–3.8	25.0–3.8	25.0–3.8	25.0–3.1	25.0–3.4	25.0–3.5
Observations	137641	106198	106491	136363	107834	143258
Unique reflections	15429	14941	14914	27074	41234	24950
$R_{sym}$ (%)	7.0	7.6	6.8	4.6	8.2	6.4
Data completeness (%)	98.9	94.5	98.9	98.2	88.2	99.7
Redundancy	8.9	7.1	7.1	5.0	2.6	5.8
			<i>MAD analysis</i>			
Resolution (Å)	20–3.8	20–3.8	20–3.8			
Phasing power	–	0.96	0.90			
$R_{cullis}$ (dispersive)	–	0.84	0.85			
$R_{cullis}$ (anomalous)	0.91	0.87	0.88			
			<i>Refinement</i>			
Resolution (Å)				20.0–3.1	8.0–3.4	15–3.5
No. of reflections $ F  > 0 \sigma F$				26000	34394	23073
No. of atoms				5650	10096	5194
$R$ -factor (%)				25.6	24.4	24.8
$R$ -free (%)				31.0	29.5	28.0
rmsd bond lengths (Å)				0.015	0.008	0.014
rmsd bond angles (°)				1.80	1.55	1.98
rmsd $B$ -factors (Å <sup>2</sup> )				1.56	–	–

RESEARCH ARTICLES

van der Waals and hydrogen bond contacts. However, instead of a tandem arrangement, they pack head-to-head in a two-fold symmetric manner (Fig. 1D). The grooves of the two OB

forms are still continuous and collinear, as in the OB2-OB3 packing. A twofold symmetric packing, although with a different twofold axis, has also been seen with the *E. coli* SSB, which

forms a 2-2-2 symmetric homotetramer (60). The OB1-OB2 packing is further stabilized by a nine-residue segment of DSS1 that fits like a wedge between the two OBs, as discussed later.



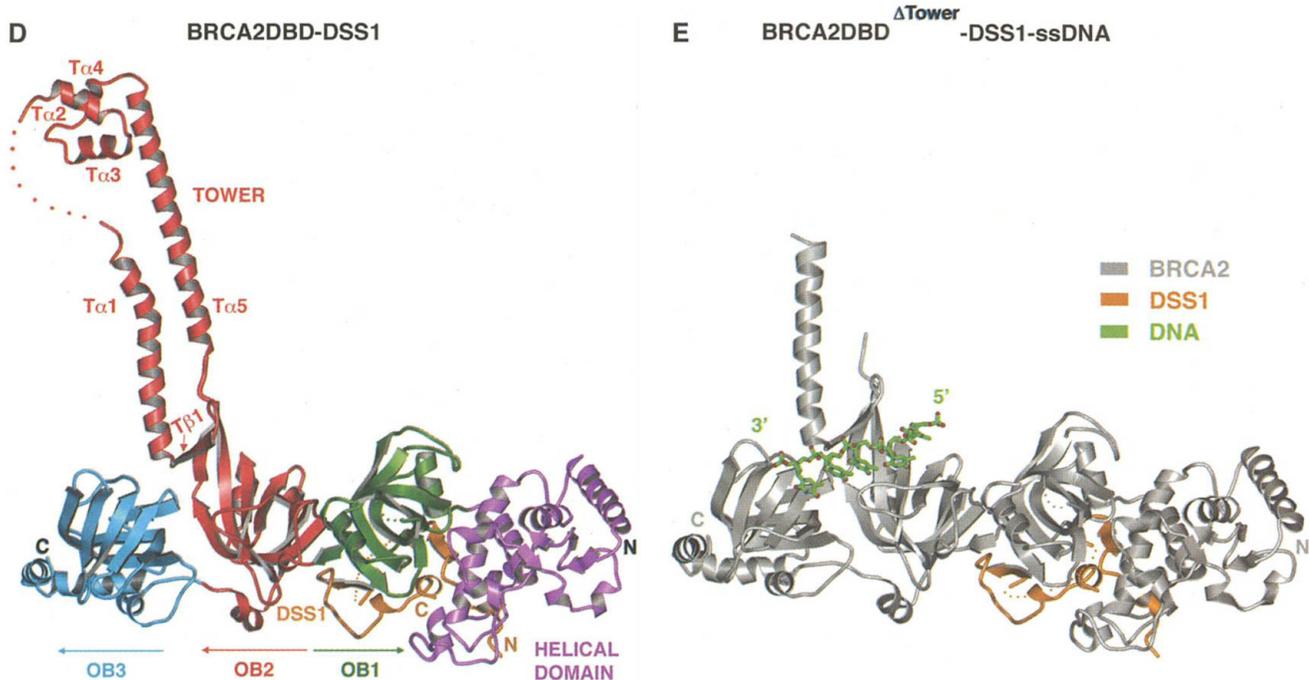
**Fig. 1.** Structures of the conserved BRCA2 COOH-terminal domain bound to DSS1 and ssDNA. (A) Graph of smoothed sequence identity across human, mouse, rat, chicken, *Arabidopsis* (accession number NM\_120241), and rice (accession number BAB64792) orthologs plotted as a function of human BRCA2 residues. The ~1000 NH<sub>2</sub>-terminal

residues are the least conserved and are absent in the reported sequences of the *Arabidopsis* and rice BRCA2-like genes. The eight BRC repeats are indicated in green, the crystallized COOH-terminal domain in red, and reported nuclear localization signals (NLS) in magenta. (B) Sequence conservation across rat (residues 2335 to 3151) and mouse

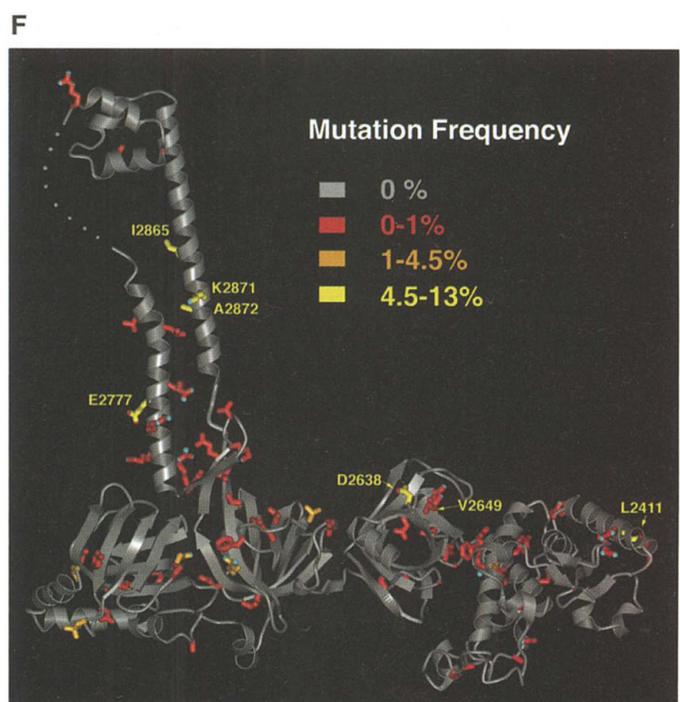
**Helical domain.** The helical domain has a core structure consisting of a four-helix cluster ( $\alpha 1$ ,  $\alpha 8$ ,  $\alpha 9$ ,  $\alpha 10$ ) and two successive  $\beta$ -hairpins ( $\beta 1$  to  $\beta 4$ ). An  $\sim 50$ -amino acid segment that contains four short helices ( $\alpha 2$  to  $\alpha 4$ ), meanders around the surface of the core structure, forming an additional layer of structural elements (Fig. 1D). The  $\alpha 9$  and  $\alpha 10$  helices pack with OB1 through van der Waals contacts involving hydro-

phobic and aromatic residues, and also through side-chain and backbone hydrogen bonds. The portions of the  $\alpha 9$  and  $\alpha 10$  helices that pack with OB1 correspond to some of the best-conserved segments of the entire BRCA2DBD (Fig. 1B and fig. S3). An unusual aspect of the helical domain is that its polypeptide chain is intertwined with DSS1, as discussed next.

**DSS1-binding sites on BRCA2.** DSS1 binds in two segments (residues 7 to 25 and 37 to 63), with an 11-residue middle portion being disordered in all the crystal forms. The BRCA2-interacting N and COOH-terminal segments are highly conserved, whereas the intervening sequence is not (Fig. 1C), and, conversely, most of the DSS1-interacting BRCA2 residues are highly conserved (Fig.



**Fig. 1 (continued).** (residues 2378 to 3114) BRCA2 fragments used in crystallization. Secondary-structure elements below the sequence are colored in magenta for the helical domain (HD in Fig. 1B), green for OB1, red for OB2, hatched-red for the Tower insertion in OB2, and blue for OB3. Colored dashed lines indicate regions that are disordered in the crystal structures of both orthologs. Black dashed lines indicate gaps in the alignment. Insertions in orthologs are dropped below the sequence. *Arabidopsis* (*A.thal.* in figure) and rice BRCA2 have no detectable sequence homology in the  $\text{NH}_2$ -terminal  $\sim 180$  residues and are omitted from this region of the alignment. Human and mouse residues are marked with a tick at every 10 amino acids and labeled with residue number at every 50 amino acids in black and orange, respectively. Residues identical in five or more orthologs are highlighted in yellow (except where the *Arabidopsis* and rice sequences do not align), DSS1-interacting residues are indicated by blue dots, and ssDNA contacting residues by green arrows. Histogram of 339 tumor-derived missense mutations (red bars) was obtained from the BIC database (41). Highest bar represents 45 mutations and lowest bars single mutation. (C) Sequence alignment of human, *Arabidopsis*, and rice DSS1 orthologs, and the yeast DSS1-like sequence Sem1. Human, mouse, and rat DSS1 orthologs are identical. The BRCA2-contacting residues of DSS1 are marked by magenta diamonds. (D) Overall view of the mouse BRCA2DBD-DSS1 structure. The helical domain is colored in magenta, OB1, OB2, and OB3 in green, red, and blue, respectively; the Tower domain, which is an insertion within OB2, is also shown in red, and DSS1 is in orange. The secondary structure elements of the Tower are labeled. Dotted lines represent disordered regions. A portion of the Tower (2796 to 2807) is disordered in the rat BRCA2DBD-DSS1 crystal form. Figures were prepared with the programs BOBSCRIPT (85), GL\_RENDER, and POVRAY. (E) Overall view of the mouse BRCA2DBD<sup>ΔTower</sup>-DSS1-oligo(dT)<sub>9</sub> ternary complex structure. The five nucleotides that are ordered are in green (also see fig. S1B), BRCA2 in gray, and DSS1 in orange. (F) The 339 tumor-derived missense mutations in the BIC database are mapped onto the mouse BRCA2DBD structure (a total of



75 residues in the BRCA2DBD are mutated). Mutated residues are color coded by their mutational frequency, which is indicated as a percentage of all BRCA2DBD missense mutations, and the seven most frequently mutated residues are labeled. DSS1 is omitted for clarity.

1B and fig. S3). The NH<sub>2</sub>-terminal DSS1 segment tunnels through the helical domain structure, traverses the helical domain-OB1 interface, and ends up packing with OB1 (Fig. 3A). The COOH-terminal segment first packs with the OB1-OB2 interface, then with OB1, and ends at the OB1-helical domain interface (Fig. 1D). Two DSS1  $\beta$  strands ( $\beta$ 1 and  $\beta$ 2) pair with OB1 ( $\beta$ 6 and  $\beta$ 5, Fig. 3A).

The unusual binding of the NH<sub>2</sub>-terminal DSS1 segment to the helical domain is mediated by both hydrophobic and charged residues (Fig. 3A). The DSS1 sequence Val<sup>8</sup>-Asp<sup>9</sup>-Leu<sup>10</sup>-Gly<sup>11</sup>-Leu<sup>12</sup>-Leu<sup>13</sup>, which forms a short  $3_{10}$  helix, is sandwiched between hydrophobic residues from the helical domain core structure ( $\alpha$ 9 and  $\alpha$ 10) on one side, and from the meandering helical domain segment on the other (Fig. 3, A and B). Immediately after the hydrophobic residues, five consecutive acidic DSS1 residues (14 to 18), all highly conserved, make salt bridges with the helical domain and OB1 residues, some of these interactions being entirely buried (Figs. 1C, and Fig. 3, A and B, and legend for interacting residues).

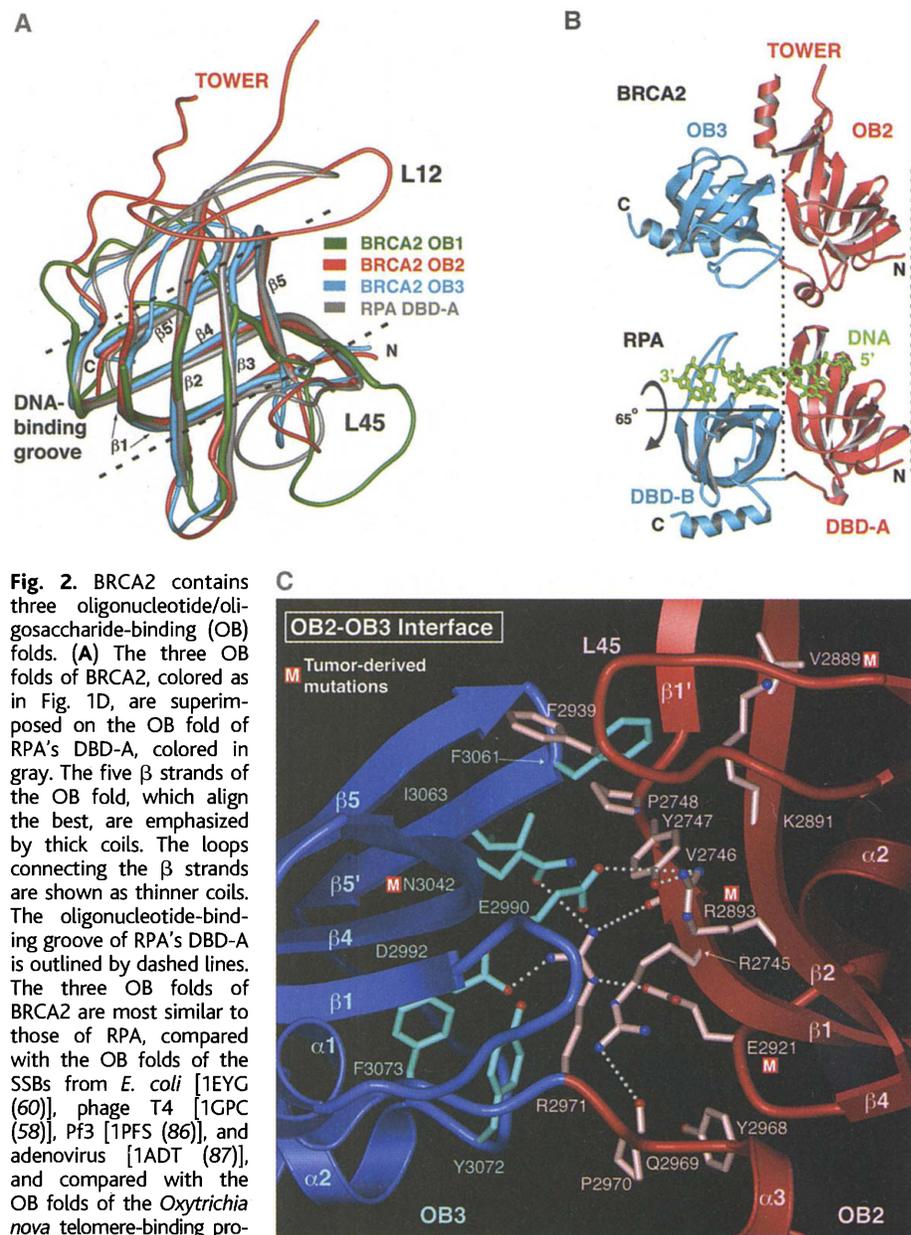
The rest of DSS1 recapitulates the theme of clusters of acidic residues with interspersed hydrophobic and aromatic residues interacting with BRCA2 (Fig. 3, B and C; see legend for list of interactions). Overall, the two DSS1 segments appear to follow the path of a long, winding groove on the helical domain and OB1 surface that is rich in basic, aromatic and hydrophobic residues (Fig. 3, B and C). This DSS1-binding groove and the canonical OB-fold groove of OB1 are roughly on opposite sides of OB1 (Fig. 1D).

**Tower structure and similarity to DNA binding domains.** The stem of the Tower consists of two long helices ( $T\alpha$ 1 and  $T\alpha$ 5) that pack antiparallel without any coiling. They are anchored on OB2 through a Tower  $\beta$  strand ( $T\beta$ 1) forming a three-stranded  $\beta$  sheet with OB2 $\beta$ 1 and OB2 $\beta$ 2, as well as through van der Waals contacts and hydrogen bond networks involving four invariant residues [fig. S2A and (44)].

$T\alpha$ 1 and  $T\alpha$ 5 pack through interdigitated hydrophobic side chains, many of which are partially exposed to the solvent. The solvent-exposed portions of these hydrophobic residues segregate to one side of the helix-helix interface, where there are several additional exposed hydrophobic and aromatic residues [see (44) for list]. In the crystals of the mouse BRCA2, this patch of exposed hydrophobic residues mediates the formation of a crystallographic dimer, where two pairs of helices pack at a  $\sim 145^\circ$  angle (fig. S2B). We have not been able to detect a stable BRCA2DBD dimer in solution, but we cannot rule out the possibility that a dimer may form on the DNA or in the context of full-length BRCA2. It is also possible that the patch of exposed hydrophobic residues is an interaction site for other parts of BRCA2 or other macromolecules.

The 3HB domain (helices  $T\alpha$ 2,  $T\alpha$ 3, and  $T\alpha$ 4) is supported by  $T\alpha$ 5, the NH<sub>2</sub>-terminal portion of which is incorporated into the hydrophobic core of the 3HB. A 12-amino acid segment between  $T\alpha$ 1 and the start of 3HB is disordered (Fig. 1D). The structure of the 3HB is most similar to the 3HB DNA binding domains of the site-specific recombinase family exemplified by the *Salmonella* Hin recombinase (48) and the *E. coli*  $\gamma\delta$  resolvase (49) in the SCOPE database (61). The entire 35-residue

BRCA2 3HB can be superimposed on 35 of the 36 residues of the Hin recombinase 3HB with a C $\alpha$  rmsd of 2.1 Å (Fig. 4). The two 3HBs have very similar hydrophobic core packings and, in fact, their structural alignment produces 18% sequence identity, most of which maps to structural residues. The differences between the two structures are primarily due to the  $T\alpha$ 3 helix, which is positioned further away from the 3HB hydrophobic core to make space for its interactions with  $T\alpha$ 5 (Fig. 4). Alignments with the



**Fig. 2.** BRCA2 contains three oligonucleotide/oligosaccharide-binding (OB) folds. (A) The three OB folds of BRCA2, colored as in Fig. 1D, are superimposed on the OB fold of RPA's DBD-A, colored in gray. The five  $\beta$  strands of the OB fold, which align the best, are emphasized by thick coils. The loops connecting the  $\beta$  strands are shown as thinner coils. The oligonucleotide-binding groove of RPA's DBD-A is outlined by dashed lines. The three OB folds of BRCA2 are most similar to those of RPA, compared with the OB folds of the SSBs from *E. coli* [1EYG (60)], phage T4 [1GPC (58)], Pf3 [1PFS (86)], and adenovirus [1ADT (87)], and compared with the OB folds of the *Oxytrichia nova* telomere-binding protein [1OTC (88)], of Asp tRNA-synthase [1ASY (89)], of RecG [1GM5 (90)], and of staphylococcal nuclease [1SNC (91)]. (B) The structure of the BRCA2 OB2-OB3 region was aligned with the structure of the RPA fragment consisting of DBD-A (red) and DBD-B (blue) bound to DNA (green) (52), by superimposing OB2 on DBD-A (shown by dashed lines). Although not used in the superposition, DBD-B of RPA ends up in a position very similar to that of OB3 of BRCA2, except for an  $\sim 65^\circ$  rotation about an axis roughly collinear with the bound DNA. The chain direction is indicated by the labels N and C. (C) OB2 and OB3 pack across an extensive interface, involving hydrophobic residues and hydrogen bond networks (white dotted lines). Many of the interface residues are highly conserved, and several have been found mutated in cancer (indicated by the letter M).

3HB domains of  $\gamma\delta$  resolvase (49), Myb (50), and engrailed homeodomain (51) give C $\alpha$  rmsds of 2.1 Å for 31 residues, 1.6 Å for 27 residues, and 2.1 Å for 29 residues, respectively.

The 3HBs of the Hin recombinase,  $\gamma\delta$  resolvase, Myb, and homeodomains bind the major groove of dsDNA. Although they all bind to DNA with sequence specificity, the Myb motif is closely related to the SANT motif, which has been proposed to be a non-specific dsDNA binding motif on the basis of its occurrence in the SWI-SNF family of chromatin remodeling proteins, in N-CoR and in TFIIB (62).

The Tower contains four of the seven most frequently mutated BRCA2DBD residues. Ala<sup>2872</sup> and Ile<sup>2865</sup> mediate T $\alpha$ 1-T $\alpha$ 5

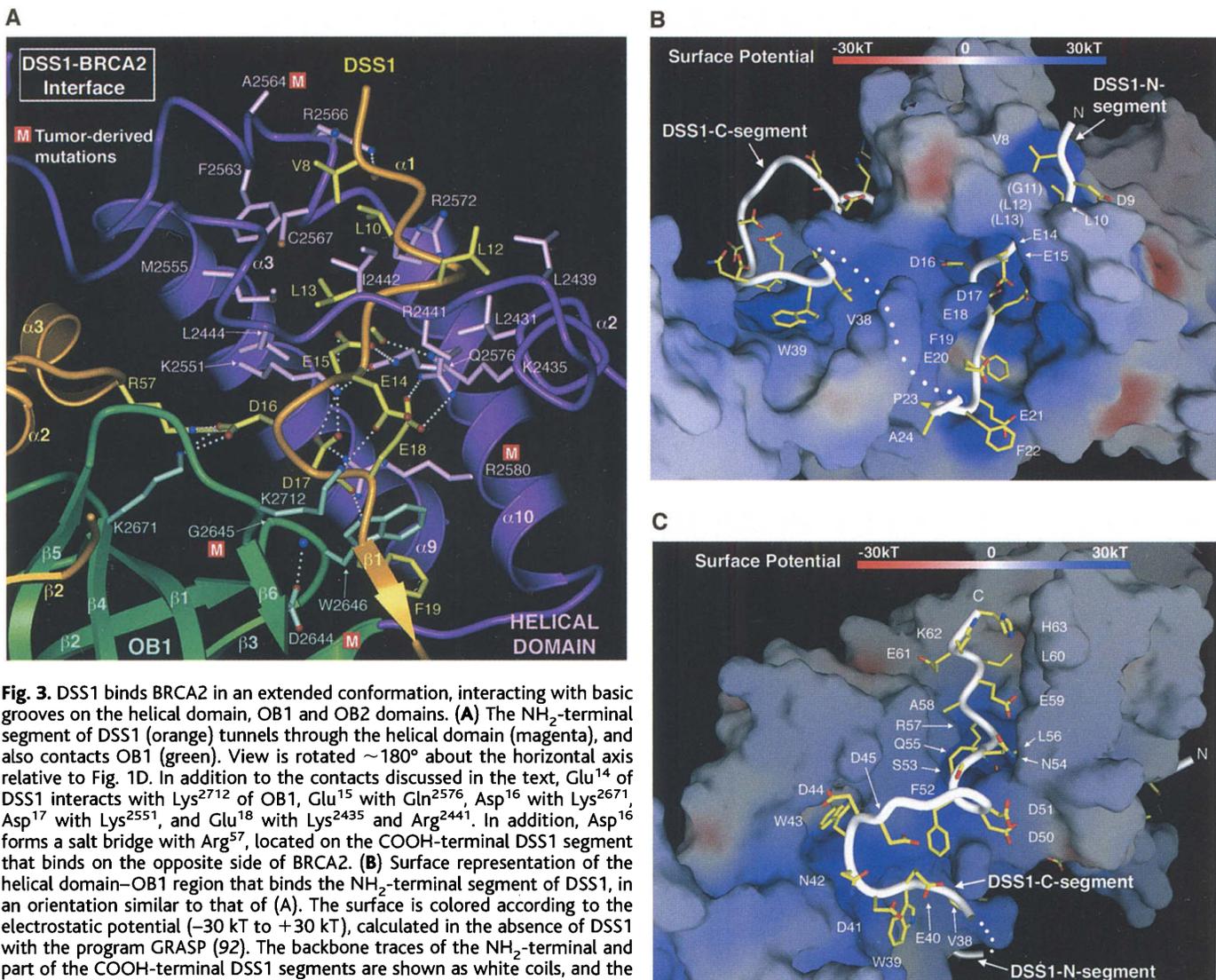
packing, and Glu<sup>2777</sup> and Lys<sup>2871</sup> are solvent exposed on the stem of the Tower (Fig. 1, B and F). This, coupled to the conservation of residues at the base of the Tower and the 3HB (Fig. 1B and fig. S3), indicates that the Tower has an important role in the tumor suppressor function of BRCA2.

#### DNA binding activity of the BRCA2DBD.

Using the native gel electrophoretic mobility-shift assay (EMSA) with end-labeled DNA probes, we found that the BRCA2DBD-DSS1 complex has affinity for oligo(dT), oligo(dC), and mixed sequence ssDNA but not for dsDNA [Fig. 5A and (63)]. As has been reported for other SSB proteins, the BRCA2DBD had negligible affinity for oligo(dA) and oligo(dG), presumably because

of their high propensity to form secondary structures (53, 64).

To investigate the optimal size of ssDNA for binding, we used oligo(dT) ranging in size from 8 to 36 nt in 4-nt increments, and a 60-nt oligo(dT). As shown in Fig. 5B, we observed three shifted bands as a function of oligonucleotide size and BRCA2DBD-DSS1 concentration (0.1, 0.3, 0.6, 1.0, 3.0, 6.0, and 10.0  $\mu$ M). We refer to the band with the slowest mobility as SC (slow-complex), the band with the fastest mobility as FC1 (fast complex 1), and the band with an intermediate mobility as FC2 (fast complex 2). SC could be observed with all sizes of oligo(dT), but its stability increased significantly with increasing size. It was barely observable with



**Fig. 3.** DSS1 binds BRCA2 in an extended conformation, interacting with basic grooves on the helical domain, OB1 and OB2 domains. **(A)** The NH<sub>2</sub>-terminal segment of DSS1 (orange) tunnels through the helical domain (magenta), and also contacts OB1 (green). View is rotated  $\sim 180^\circ$  about the horizontal axis relative to Fig. 1D. In addition to the contacts discussed in the text, Glu<sup>14</sup> of DSS1 interacts with Lys<sup>2712</sup> of OB1, Glu<sup>15</sup> with Gln<sup>2576</sup>, Asp<sup>16</sup> with Lys<sup>2671</sup>, Asp<sup>17</sup> with Lys<sup>2551</sup>, and Glu<sup>18</sup> with Lys<sup>2435</sup> and Arg<sup>2441</sup>. In addition, Asp<sup>16</sup> forms a salt bridge with Arg<sup>57</sup>, located on the COOH-terminal DSS1 segment that binds on the opposite side of BRCA2. **(B)** Surface representation of the helical domain-OB1 region that binds the NH<sub>2</sub>-terminal segment of DSS1, in an orientation similar to that of (A). The surface is colored according to the electrostatic potential ( $-30$  kT to  $+30$  kT), calculated in the absence of DSS1 with the program GRASP (92). The backbone traces of the NH<sub>2</sub>-terminal and part of the COOH-terminal DSS1 segments are shown as white coils, and the side chains, in yellow. The white-dotted line indicates the 11-residue disordered segment between the NH<sub>2</sub>- and COOH-terminal segments of DSS1. DSS1 residues that are entirely buried below the surface of BRCA2 are indicated in parenthesis. The side chains for residues 46 to 49 of the COOH-terminal segment have poor electron density and are omitted. **(C)** Surface representation of the helical domain-OB2 region that binds the COOH-terminal segment of DSS1, in an orientation rotated  $\sim 90^\circ$  about the vertical axis relative to (B). Hydrophobic interactions include the DSS1 Phe<sup>19</sup>, Trp<sup>39</sup>, Trp<sup>43</sup>, which pack with Trp<sup>2646</sup>, Phe<sup>2722</sup>, and with an opening in the hydrophobic core of OB1, respectively. In addition, the majority of the acidic DSS1 residues are near clusters of basic BRCA2 residues. These include Glu<sup>40</sup>, Asp<sup>41</sup>, and Asp<sup>44</sup>, which are near Arg<sup>2708</sup>, Lys<sup>2898</sup>, and Lys<sup>2655</sup>; Asp<sup>50</sup> and Asp<sup>51</sup>, which are near Arg<sup>2708</sup> and Arg<sup>2451</sup>; and Glu<sup>20</sup> and Glu<sup>21</sup> [from the NH<sub>2</sub>-terminal segment in (B)] which are near Lys<sup>2712</sup>, Arg<sup>2717</sup>, Arg<sup>2590</sup> and Lys<sup>2595</sup>.

oligo(dT)<sub>8</sub> (T8 in Fig. 5B), which exhibited unstable binding characterized by smeared migration of DNA. With T36 and T60, SC was observable at 0.3 μM protein concentration (Fig. 5B). FC1 first appeared as a minor band with T24. Its stability also increased with longer oligonucleotides, reaching a plateau with T36 (Fig. 5B). Based on this, we estimate that FC1 requires approximately 32- to 36-nt oligo(dT) for optimal binding. FC2 first appeared with T36, but it was not as stable as FC1 until T60 (Fig. 5B). On the basis of its relative mobility, we presume that FC2 represents two FC1-like complexes on a single oligonucleotide. The apparent dissociation constant for the BRCA2DBD-DSS1 bound T60 is ~0.25 μM. All of the shifted bands disappeared when the reactions were deproteinized with Proteinase K and SDS, confirming that the shifted bands are not the result of changes in secondary structure or covalent modification of the DNA.

The nature of the SC species is not clear. Its mobility did not vary with the length of oligo(dT) and, in fact, was indistinguishable from that of the free protein as seen with Coomassie staining (63). In addition, most of the FC complexes were converted to the SC complex at high protein concentrations (compare lanes 3 and 8 of T36 and of T60 in Fig. 5B).

We next explored the binding of the BRCA2DBD to DNA containing both single-stranded and double-stranded regions, as these DNA structures share common aspects with the substrates and intermediates of homologous recombination (19). We found that 20-nt 3' or 5' overhanging ssDNA, and 10- or 20-nt ssDNA gaps produced both the SC and FC1 and, to a lesser extent, FC2 complexes (Fig. 5C). This finding, together with the fact that the ssDNA regions in these fragments are too short to produce the FC complexes by themselves (Fig. 5B), suggests that dsDNA, in the context of ssDNA, may contribute to the formation of the FC complexes. The same pattern of bands was obtained irrespective of which strand was labeled, ruling out the possibility that the BRCA2DBD dissociates the DNA duplex (63).

**Tower is necessary for the formation of the FC complexes.** To investigate whether the Tower contributes to the DNA binding activity, we constructed a BRCA2DBD mutant lacking the 3HB and the 3HB-proximal portions of the Tα1 and Tα5 helices (BRCA2DBD<sup>ΔTower</sup>). The crystal structure of BRCA2DBD<sup>ΔTower</sup>-DSS1 complex confirmed that the deletion did not affect the rest of the BRCA2DBD-DSS1 structure (Fig. 1E). Figure 5D shows that BRCA2DBD<sup>ΔTower</sup>-DSS1 binds T20, T36, and T60 with overall affinities comparable to those of the intact BRCA2DBD-DSS1. However, the BRCA2DBD<sup>ΔTower</sup>-DSS1 produces only a single shifted band, even with T60. The shifted band migrates to the

same position as the free protein, irrespective of the size of the ssDNA; it has mobility similar to that of the SC species produced by BRCA2DBD-DSS1, and we conclude that it represents a complex qualitatively similar to the SC complex. These findings indicate that the formation of the FC complexes requires the 3HB.

**The isolated helical domain-OB1 fragment binds DNA weakly.** We tested a truncated BRCA2DBD containing only the helical and OB1 domains (BRCA2DBD<sup>ΔOB2-OB3</sup>), coexpressed with DSS1, and found that it produced a single shifted band with ssDNA (T20 and T60, Fig. 5E). The shifted band was evident at 6 μM BRCA2DBD<sup>ΔOB2-OB3</sup>-DSS1 with T60, and at 10 μM with T20. We estimate that it forms with an apparent affinity about 1/100th that of the intact BRCA2DBD-DSS1. We did not determine whether this weak DNA binding activity is due to OB1 or the helical domain, or both, as out of a series of truncated BRCA2DBD constructs, this was the only one that could be expressed in a soluble form.

**Structure of the BRCA2-ssDNA complex.** The best crystals, diffracting to 3.5 Å, were obtained with the BRCA2DBD<sup>ΔTower</sup>-DSS1 complex bound to an oligo(dT)<sub>9</sub> ssDNA. We also obtained crystals of intact BRCA2DBD-DSS1-ssDNA complexes, but these crystals diffracted only to ~4.0 Å (44, 63). In initial model-phased difference Fourier maps, both crystal forms showed clear density for five complete nucleotides (fig. S1B).

The ssDNA binds in the OB2-OB3 channel in a uniform conformation (Figs. 1E and 6). The edges of the bases bind at the bottom of the channel, and the phosphodiester backbone forms a ridge that extends along the surface of the channel. The first three nucleotides (Thy<sup>1</sup> through Thy<sup>3</sup>) interact with OB2, the fourth (Thy<sup>4</sup>) with both OB2 and OB3 and the fifth (Thy<sup>5</sup>) with OB3. The 5' phosphate group of a sixth base (Thy<sup>6</sup>) interacts with OB3, but its sugar and base groups are disordered. Like other SSBs (52, 58, 60), the BRCA2 OB-ssDNA contacts involve stacking interactions between the bases and aromatic residues, hydrogen bond interactions between the edges of the bases and polar residues lining the bottom and sides of the channel, and basic residues or backbone amide groups interacting with phosphate groups (Fig. 6).

OB2 makes the most extensive contacts, and these are qualitatively similar to those in the RPA-ssDNA complex (52). For example, Trp<sup>2909</sup> stacks with the Thy<sup>1</sup> base in parallel and with the Thy<sup>2</sup> base edge-on, being sandwiched between the two (Fig. 6), and this is very similar to the Cyt<sup>1</sup>-Phe<sup>238</sup>-Cyt<sup>2</sup> stacking in RPA. The majority of DNA binding OB folds contain an aromatic residue at this position, and BRCA2 Trp<sup>2909</sup> is invariant across species (Fig. 1B).

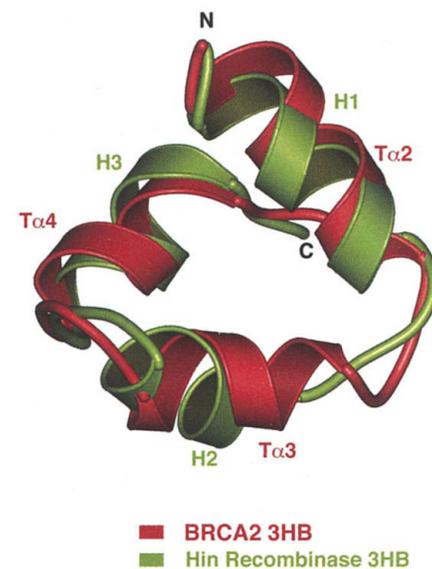
The rest of the contacts also follow the paradigm of RPA (Fig. 6 and legend for list

of contacts). One difference between OB2 and RPA is the involvement of the residues from the base of the Tower (Tβ1), where Phe<sup>2762</sup> stacks with the base of Thy<sup>3</sup> and Tyr<sup>2760</sup> packs with the sugar group of Thy<sup>2</sup>.

The phosphate group of Thy<sup>6</sup> is clearly visible in our electron density maps, interacting with the backbone amide of Ala<sup>3007</sup> from the OB3L12 loop, but the base is disordered. The rest of the OB3 groove is not available for DNA binding, as it is involved in crystal packing contacts in both crystal forms. We do not know whether OB3 has the capacity to interact with additional bases.

Three of the OB2 residues that contact DNA (Lys<sup>2754</sup>, Gln<sup>2945</sup>, and Ser<sup>2907</sup>) but have no structural roles have been found mutated in cancer; this underscores the importance ssDNA binding for BRCA2 function.

**The BRCA2DBD stimulates RAD51-mediated recombination.** We used an *in vitro* assay involving incubation of circular φX174 ssDNA with RAD51, followed by the addition of RPA, and then by the addition of linearized φX174 RFI (replication form I) dsDNA (65). Strand pairing and exchange result in the formation of joint molecule intermediates (JM in Fig. 7), nicked circular dsDNA (NC in Fig. 7), and displaced linear ssDNA products. In this assay, strand pairing and exchange are strongly dependent on RPA (31, 65-67), and it has been suggested that this is due to the ability of RPA to remove secondary structure from ssDNA, helping RAD51 form a more continuous ssDNA-



**Fig. 4.** BRCA2 contains a three-helix bundle (3HB) similar to the helix-turn-helix dsDNA binding motif. The 35-residue BRCA2 3HB is superimposed on the 36-residue 3HB domain of the Hin recombinase DNA binding domain (48). The 35 residues correspond to the entire 3HB, starting with the NH<sub>2</sub>-terminus of Tα2 and ending at the COOH-terminus of Tα4. There is only one residue insertion between H2 and H3 of the Hin recombinase 3HB in the alignment.

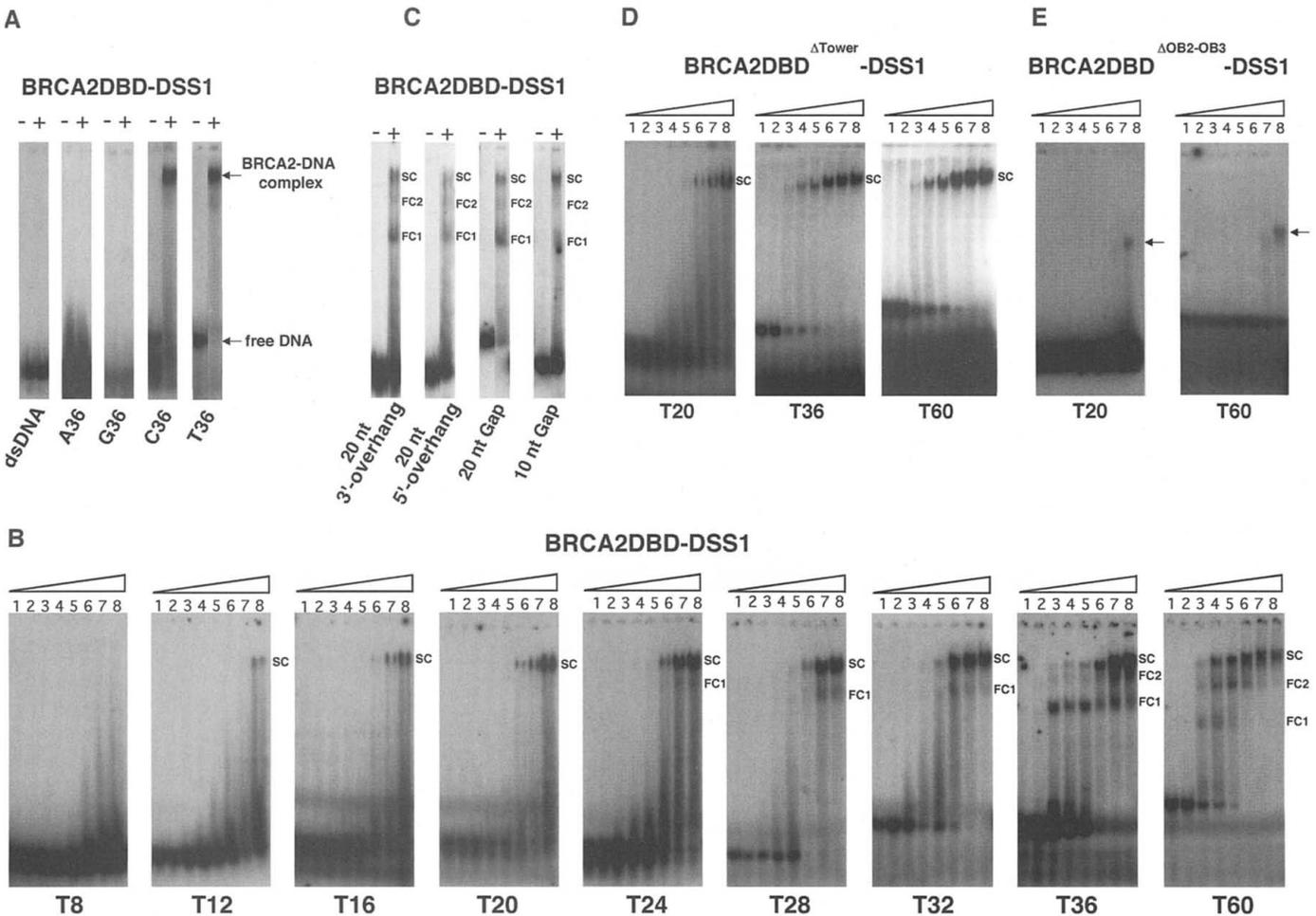
RAD51 nucleoprotein filament (68). Because of the similarities between the BRCA2DBD and RPA, we first tested if the BRCA2DBD can substitute for RPA in this assay. In the absence of RPA, we were unable to detect any appreciable stimulation of RAD51 over a 60-fold range of BRCA2DBD-DSS1 concentration (44), but we cannot rule out the possibilities that full-length BRCA2 or different experimental conditions may be needed.

Although RPA is required for this RAD51-mediated in vitro recombination reaction, RPA's stimulatory effects are mini-

mal if it is given access to ssDNA before RAD51 (66, 69). This is thought to be due to RPA competing with RAD51 for the ssDNA substrate, as it binds ssDNA ~100 times as tightly as RAD51 (53, 70). It has been suggested that accessory factors, termed mediators, may be required to remove RPA from ssDNA in vivo (71). Several proteins of the RAD52 epistasis group have been shown to have mediator activity in vitro (72–75).

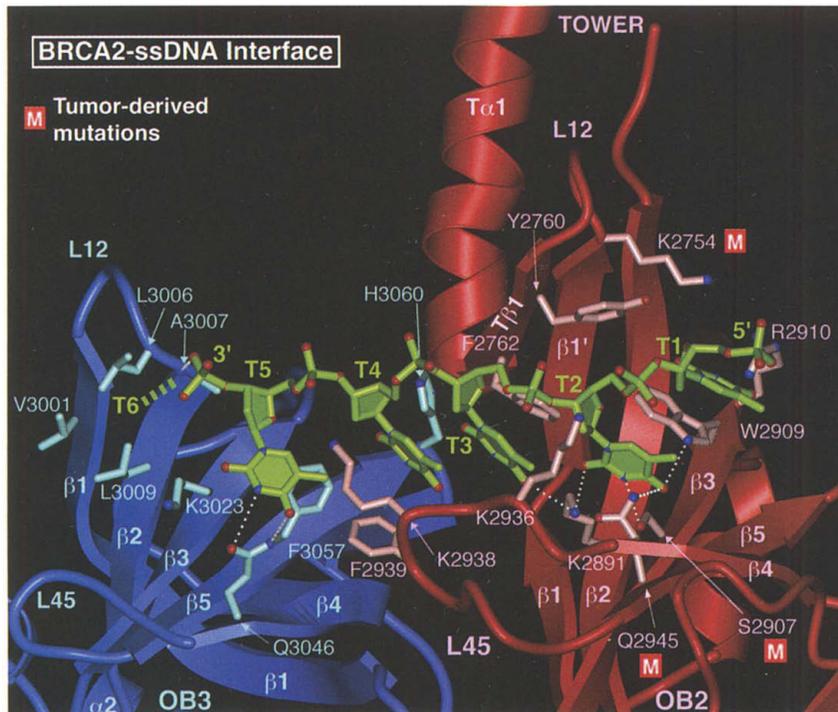
We thus tested whether the BRCA2DBD-DSS1 complex can stimulate the RPA-dependent recombination activity of RAD51 using

assay conditions similar to those used to demonstrate a mediator activity for RAD52 (72–74, 76). As described previously, we found that the reaction efficiency, measured as the amount of JM and NC produced, was less when RPA and RAD51 were added to ssDNA simultaneously ( $7.4 \pm 2.0\%$ , lane 3 in Fig. 7) compared with when RPA was added after RAD51 ( $12.8 \pm 3.0\%$ , lane 9). When we added BRCA2DBD-DSS1 at the same time as RPA and RAD51, we observed a stimulation of the reaction as a function of the amount of BRCA2DBD-DSS1 added.



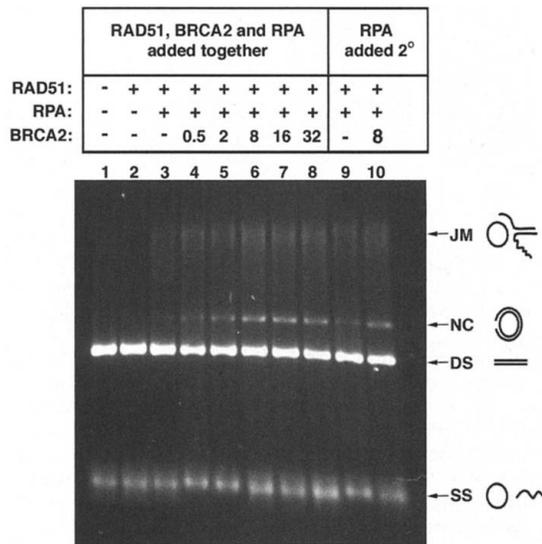
**Fig. 5.** BRCA2DBD binds ssDNA in vitro, forming two distinct protein-DNA complexes. (A) Native gel electrophoretic mobility-shift assay with  $^{32}\text{P}$  end-labeled DNA probes (5 nM) incubated with 10  $\mu\text{M}$  of mouse BRCA2DBD-DSS1 where indicated with a (+). dsDNA is a 40-bp random-sequence duplex DNA, A36 is oligo(dA) $_{36}$ , G36 is oligo(dG) $_{36}$ , C36 is oligo(dC) $_{36}$ , and T36 is oligo(dT) $_{36}$ . (B) Oligo(dT) ranging in size from 8 to 36 nt (T8 to T36) and oligo(dT) $_{60}$  (T60), each at 5 nM concentration, were incubated with the following mouse BRCA2DBD-DSS1 concentrations: 0.0 (lane 1), 0.1  $\mu\text{M}$  (lane 2), 0.3  $\mu\text{M}$  (lane 3), 0.6  $\mu\text{M}$  (lane 4), 1.0  $\mu\text{M}$  (lane 5), 3.0  $\mu\text{M}$  (lane 6), 6.0  $\mu\text{M}$  (lane 7), and 10.0  $\mu\text{M}$  (lane 8). All gels were run at 200 V for 220 min. The three bands with different shifted mobilities are labeled SC, FC1, and FC2. The identification of the fastest mobility band in T36 as FC1 is based on extrapolating from the mobility of FC1 in T24, T28, and T32, which increased linearly as a function of oligo(dT) size. (C) DNA probes containing both single-stranded and double-stranded DNA regions can produce the FC1 complex even with a single-stranded region as short as 10 nt. DNA probes, each at 5 nM concentration, were incubated with 10  $\mu\text{M}$  mouse

BRCA2DBD-DSS1 complex where indicated with a (+). The 5'- and 3'-overhang probes both contain a 20-nt mixed-sequence, single-stranded region, and a 20-bp mixed-sequence duplex DNA. Gap probes contain either a 10- or 20-nt oligo(dT) flanked by 20 bp of mixed-sequence duplex DNA (44). (D) Deletion of the Tower abolishes the formation of the FC complexes. Oligo(dT) $_{20}$ , oligo(dT) $_{36}$ , and oligo(dT) $_{60}$  probes, each at 5 nM concentration, were incubated with the mouse BRCA2DBD $^{\Delta\text{Tower}}$ -DSS1 complex at the same protein concentrations and separated under the same electrophoretic conditions as in (B). Only a single shifted band, with the properties of the SC complex in (B) is observable, even after prolonged exposure of the gels. (E) Deletion of the OB2, OB3, and Tower domains eliminates most but not all of BRCA2's affinity for ssDNA. Oligo(dT) $_{20}$  and oligo(dT) $_{60}$  probes, each at 5 nM concentration, were incubated with the mouse BRCA2DBD $^{\Delta\text{OB2-OB3}}$ -DSS1 complex at the same protein concentrations and separated under the same electrophoretic conditions as in (B). About half of the oligo(dT) $_{60}$  probe is bound at 10  $\mu\text{M}$  of mouse BRCA2DBD $^{\Delta\text{OB2-OB3}}$ -DSS1 complex. Oligo(dT) $_{20}$  probe has a substantially lower affinity.



**Fig. 6.** The OB2 and OB3 domains interact with five nucleotides in the structures of the mouse BRCA2DBD-DSS1-oligo(dT) ternary complexes. The best crystals were obtained with BRCA2DBD<sup>ΔTower</sup>-DSS1 bound to oligo(dT)<sub>9</sub> containing either four 5-bromo-uracil or four 5-iodo-uracil substitutions (at positions 2, 4, 6, and 8). Crystals were also obtained with the intact mouse BRCA2DBD-DSS1 complex bound to oligo(dT) ranging in size from 9 to 24 nt, but these typically diffracted to worse than ~4.0 Å. Irrespective of the length of oligonucleotide and BRCA2DBD construct used in the cocrystallization, most data sets showed clear difference electron density for five complete nucleotides and the phosphate group of a sixth one. The anomalous dispersion signal from the data set containing 5-iodo-uracil was used to confirm the building of the DNA (44). This figure shows the interactions between OB2 (red), OB3 (blue), and oligo(dT) (green) from the 3.5 Å refined structure of the BRCA2DBD<sup>ΔTower</sup>-DSS1 bound to 5-bromo-uracil containing oligo(dT)<sub>9</sub>. The green dashed line indicates the disordered thymidine of Thy<sup>6</sup>. The side chain of Lys<sup>2754</sup> has high temperature factors, and its interaction with the Thy<sup>1</sup> phosphate group is not shown. The rest of the BRCA2 side chains that are shown are all involved in DNA contacts, except for Val<sup>3001</sup>, Leu<sup>3009</sup>, and Lys<sup>3023</sup> of OB3, which are located in positions used by other OB folds to bind DNA, but are involved in crystal-packing contacts instead.

**Fig. 7.** The BRCA2DBD-DSS1 complex (BRCA2 on figure) stimulates RAD51-mediated homologous DNA pairing and strand exchange in vitro. The proteins included in each reaction are indicated above the lanes. Reactions contained φX174 ssDNA (30 μM nt), linearized dsDNA (30 μM nt), human RAD51 (7.5 μM), heterotrimeric human RPA (4.0 μM), and the BRCA2DBD-DSS1 complex (BRCA2 in figure) at the micromolar concentrations indicated. In lanes 1 to 8, RAD51, RPA, and BRCA2DBD-DSS1 were mixed together before they were added to ssDNA. In lanes 9 and 10, RAD51 and BRCA2DBD-DSS1 were first incubated together, then with ssDNA, before RPA was added. In these assays, the RPA concentration is 2 times (4 μM) what has been described as the optimum (2 μM), which still stimulates the reaction, but not as much, presumably because of a partial competition with RAD51 for ssDNA. The DNA substrates (DS for dsDNA, SS for ssDNA), intermediates (JM for joint molecules), and products (NC for nicked circular) are shown schematically and their positions on the gel are indicated by arrows.



Adding as little as 0.5 μM BRCA2DBD-DSS1, which corresponds to 1/15th of the RAD51 concentration, resulted in a 2.4-fold stimulation of the reaction (17.8 ± 5.4% compared with 7.4 ± 2.0%, lanes 4 and 3, respectively). The extent of stimulation reached a plateau at 4.5-fold with 8.0 μM BRCA2DBD-DSS1 (33.6 ± 3.7%, lane 6). The BRCA2DBD-DSS1 complex stimulated the reaction when RPA was added after RAD51 and ssDNA, which corresponds to the optimal order of addition, although the amount of stimulation was less (by a factor of 1.9, compare lanes 9 and 10).

**Role of DSS1 in the BRCA2 pathway.** The function of DSS1 and its role in the BRCA2 pathway are unknown. In HeLa cells, approximately half of the endogenous BRCA2 co-immunoprecipitates with DSS1, and the majority of DSS1 coimmunoprecipitates with BRCA2 (42). A cellular role for the DSS1-BRCA2 interaction is supported by the following observations: (i) The portions of DSS1 that are involved in BRCA2 binding are highly conserved, whereas the regions not involved in BRCA2-binding are not (Fig. 1C). (ii) The DSS1-interacting BRCA2 residues are conserved (Fig. 1B and fig. S3). And (iii), two DSS1-interacting BRCA2 residues (Ala<sup>2564</sup> and Arg<sup>2580</sup>) are mutated in cancer (Fig. 1B). One model about the function of DSS1, suggested by its 37% acidic-residue (-21 charge at pH 7.0) and 13% aromatic-residue content (Fig. 1C), is that it mimics oligonucleotides, possibly regulating the accessibility of a subset of the putative DNA binding sites on the helical and OB1 domains.

**DNA binding activities of BRCA2.** Our biochemical data show that the BRCA2DBD-DSS1 can produce two distinct types of protein-DNA complexes. The FC complex requires the Tower, forms avidly with 0.3 μM BRCA2DBD-DSS1, and requires a 32- to 36-nt oligo(dT). In contrast, the SC complex can form even with 8 to 12 nt, it is not affected by the deletion of the Tower, and it forms stably only at ≥ 3.0 μM.

Two binding modes have also been demonstrated for RPA (52, 53, 55–57, 77). One mode requires only the DBD-A and DBD-B OB folds, as seen in the RPA-ssDNA crystal structure; it has low affinity but high cooperativity, and it has a binding site of 8 nt [8-nt mode (55–57)]. It is thought that the crystal structure; of the RPA-ssDNA complex represents this mode. The other mode requires all four DBDs (A/B/C/D) of RPA, as demonstrated by deletion, mutagenesis and cross-linking studies, it has high affinity but low cooperativity, and it involves an occluded size of ~30 nt (30-nt mode).

On the basis of the analogies between the 8-nt binding mode of RPA and the SC com-

plex, it is likely that the binding of OB2 and OB3 to the 5 nt that we observe in our BRCA2DBD<sup>ΔTower</sup>-DSS1-oligo(dT)<sub>6</sub> crystal structure represents the SC or a closely related complex. The FC-type complexes could then involve the engagement of additional regions of BRCA2 in DNA binding. Our observation that a BRCA2 fragment containing only the helical domain and OB1 still has DNA binding activity, albeit much weaker than the intact DBD, indicates that this region contains at least some of the additional DNA binding domains. The weak DNA-affinity of the isolated helical domain-OB1 is again reminiscent of RPA, where the isolated DBD-C/D complex has a DNA affinity about 1/100th that of the DBD-A/B fragment (78–80).

In addition, our data implicating the Tower 3HB in the formation of the FC complexes point to a fourth DNA-interacting region within the BRCA2DBD. It is not yet clear whether the Tower is recognizing ssDNA or dsDNA, but most of the helix-turn-helix containing three helix bundle domains recognize dsDNA. The FC complexes can be produced by DNA substrates that contain short ssDNA in the context of dsDNA (Fig. 5D); this suggests that the 3HB interacts with dsDNA or dsDNA-like structures [see SOM text for additional discussion]. Although in our binding experiments in Fig. 5B we used ssDNA, the adoption of secondary structure by the ssDNA, which would be more avid with longer oligonucleotides, could result in dsDNA-like structures.

Finally, the highly basic nature and high aromatic-residue content of the DSS1-binding BRCA2 groove raise the possibility that it may interact with ssDNA in the absence of DSS1. Dissociation of all or part of DSS1 could happen as a result of competition with DNA or after a modification in either DSS1 or BRCA2. It is interesting that RPA's DBD-D domain has cryptic DNA binding activity that is observable only when a 40-residue NH<sub>2</sub>-terminal region that contains the majority of the regulatory phosphorylation sites is deleted (79).

**Role of BRCA2 in RAD51-mediated homologous recombination.** The DNA binding activities of the BRCA2DBD, in conjunction with the RAD51-binding activities of the eight BRC repeats can, in principle, have a direct role in facilitating several steps of RAD51-mediated homologous recombination. These steps could include the recruitment of RAD51 to sites of processed DSBs, the displacement of RPA, the formation of the RAD51-ssDNA nucleoprotein filament, or the pairing and exchange of the ssDNA with dsDNA. The model of BRCA2 facilitating the recruitment of RAD51 to DSBs, in particular, could explain several observations. In the cell, DSBs are typically resected to produce long stretches of ssDNA (19). A role for BRCA2 in recognizing aspects of the ssDNA/dsDNA hy-

brid structures of processed DSBs could explain our observations implicating the Tower in interactions with dsDNA-like structures in the context of the OB folds recognizing ssDNA, possibly with contributions from the DSS1-binding regions. A role for BRCA2 in recruiting RAD51 to DSBs could also explain RAD51's lack of specificity for ssDNA over dsDNA (27, 31, 32, 81), which is one of the features that distinguish it from the prokaryotic RecA homolog. RecA has a strong preference for ssDNA (82), and this property allows it to overcome the inhibitory effects of dsDNA and may be part of the mechanism of properly localizing RecA to DSBs (83). With the more complex eukaryotic genomes, RAD51 might have evolved to rely on BRCA2 for its localization to processed DSBs or to specific DNA structures within them, and its lack of preference for ssDNA over dsDNA may reflect this.

The model of BRCA2 recruiting RAD51 to processed DSBs could also be consistent with the observation that isolated BRC-repeat peptides interfere with RAD51's ability to bind both ssDNA and dsDNA in vitro (84). In principle, a mechanism that confers specificity for DSBs through a recruitment component could be complemented by an inhibitory component that prevents RAD51 from interacting with inappropriate DNA substrates, especially as RAD51 has similar affinities for dsDNA and ssDNA. After it is recruited to sites within processed DSBs, the recruited RAD51 may dissociate from the BRC repeats and partition onto ssDNA. This aspect of the model may also explain why some of the eight BRC repeats have widely different affinities for RAD51 (26, 39). The weakly bound RAD51 molecules could be first to partition onto the DNA, whereupon the cooperativity of nucleoprotein filament formation could drive even the relatively tightly bound RAD51 molecules to partition onto the RAD51-ssDNA nucleoprotein filament.

This model does not exclude a need for RPA to bind the ssDNA ahead of time to remove secondary structure and to facilitate ssDNA-RAD51 nucleoprotein filament formation. The full-length BRCA2-RAD51 complex may have a high enough cooperative affinity for ssDNA to displace RPA, or BRCA2 may displace RPA through the mediator mechanism proposed for RAD52 and several other proteins [(72–76), and see additional discussion in SOM text]. A mediator-like activity is supported by our in vitro data, which show stimulation of the RPA-dependent recombination activity of RAD51 by the BRCA2DBD.

Our crystallographic, biochemical, and mutational data thus establish a new molecular function for the BRCA2 tumor suppressor in binding DNA and provide a framework for understanding the loss of homologous recom-

bination-mediated DSB repair in BRCA2-associated cancers.

#### References and Notes

1. R. Wooster et al., *Nature* **378**, 789 (1995).
2. S. V. Tavtigian et al., *Nature Genet.* **12**, 333 (1996).
3. K. L. Nathanson, R. Wooster, B. L. Weber, K. N. Nathanson, *Nature Med.* **7**, 552 (2001).
4. The Breast Cancer Linkage Consortium, *J. Natl. Cancer Inst.* **91**, 1310 (1999).
5. N. G. Howlett et al., *Science* **297**, 606 (2002).
6. R. Scully, D. M. Livingston, *Nature* **408**, 429 (2000).
7. A. R. Venkitaraman, *Cell* **108**, 171 (2002).
8. S. K. Sharan et al., *Nature* **386**, 804 (1997).
9. F. Connor et al., *Nature Genet.* **17**, 423 (1997).
10. K. J. Patel et al., *Mol. Cell* **1**, 347 (1998).
11. M. Morimatsu, G. Donoho, P. Hasty, *Cancer Res.* **58**, 3441 (1998).
12. A. Tutt et al., *Curr. Biol.* **9**, 1107 (1999).
13. V. P. Yu et al., *Genes Dev.* **14**, 1400 (2000).
14. S. S. Yuan et al., *Cancer Res.* **59**, 3547 (1999).
15. M. E. Moynahan, A. J. Pierce, M. Jasin, *Mol. Cell* **7**, 263 (2001).
16. A. Tutt et al., *EMBO J.* **20**, 4704 (2001).
17. A. N. Tutt, C. T. van Oostrom, G. M. Ross, H. van Steeg, A. Ashworth, *EMBO Rep.* **3**, 255 (2002).
18. M. Kraakman-van der Zwet et al., *Mol. Cell Biol.* **22**, 669 (2002).
19. K. K. Khanna, S. P. Jackson, *Nature Genet.* **27**, 247 (2001).
20. T. Ludwig, D. L. Chapman, V. E. Papaioannou, A. Efstratiadis, *Genes Dev.* **11**, 1226 (1997).
21. A. Suzuki et al., *Genes Dev.* **11**, 1242 (1997).
22. A. J. Pierce et al., *Trends Cell Biol.* **11**, S52 (2001).
23. F. Xia et al., *Proc. Natl. Acad. Sci. U.S.A.* **98**, 8644 (2001).
24. R. Mizuta et al., *Proc. Natl. Acad. Sci. U.S.A.* **94**, 6927 (1997).
25. A. K. Wong, R. Pero, P. A. Ormonde, S. V. Tavtigian, P. L. Bartel, *J. Biol. Chem.* **272**, 31941 (1997).
26. P. L. Chen et al., *Proc. Natl. Acad. Sci. U.S.A.* **95**, 5287 (1998).
27. P. Baumann, S. C. West, *Trends Biochem. Sci.* **23**, 247 (1998).
28. S. C. West, *Annu. Rev. Biochem.* **61**, 603 (1992).
29. R. M. Story, I. T. Weber, T. A. Steitz, *Nature* **355**, 318 (1992).
30. T. Ogawa, X. Yu, A. Shinohara, E. H. Egelman, *Science* **259**, 1896 (1993).
31. P. Sung, *Science* **265**, 1241 (1994).
32. P. Baumann, F. E. Benson, S. C. West, *Cell* **87**, 757 (1996).
33. J. Chen et al., *Mol. Cell* **2**, 317 (1998).
34. J. J. Chen, D. Silver, S. Cantor, D. M. Livingston, R. Scully, *Cancer Res.* **59**, 1752s (1999).
35. D. S. Lim, P. Hasty, *Mol. Cell Biol.* **16**, 7133 (1996).
36. T. Tsuzuki et al., *Proc. Natl. Acad. Sci. U.S.A.* **93**, 6236 (1996).
37. P. Bork, N. Blomberg, M. Nilges, *Nature Genet.* **13**, 22 (1996).
38. J. M. Stark et al., *J. Biol. Chem.* **277**, 20185 (2002).
39. C. F. Chen, P. L. Chen, Q. Zhong, Z. D. Sharp, W. H. Lee, *J. Biol. Chem.* **274**, 32931 (1999).
40. M. Warren et al., *Hum. Mol. Genet.* **11**, 841 (2002).
41. C. Szabo, A. Masiello, J. F. Ryan, L. C. Brody, *Hum. Mutat.* **16**, 123 (2000).
42. N. J. Marston et al., *Mol. Cell Biol.* **19**, 4633 (1999).
43. M. A. Crakower et al., *Hum. Mol. Genet.* **5**, 571 (1996).
44. Material and Methods are available as supporting online material on Science Online.
45. A. G. Murzin, *EMBO J.* **12**, 861 (1993).
46. D. Suck, *Nature Struct. Biol.* **4**, 161 (1997).
47. G. Webster et al., *FEBS Lett.* **411**, 313 (1997).
48. J. A. Feng, R. C. Johnson, R. E. Dickerson, *Science* **263**, 348 (1994).
49. W. Yang, T. A. Steitz, *Cell* **82**, 193 (1995).
50. K. Ogata et al., *Cell* **79**, 639 (1994).
51. C. R. Kissinger, B. S. Liu, E. Martin-Blanco, T. B. Kornberg, C. O. Pabo, *Cell* **63**, 579 (1990).
52. A. Bochkarev, R. A. Pfuetzner, A. M. Edwards, L. Frappier, *Nature* **385**, 176 (1997).
53. M. S. Wold, *Annu. Rev. Biochem.* **66**, 61 (1997).
54. L. Holm, C. Sander, *J. Mol. Biol.* **218**, 183 (1991).
55. A. Bochkarev, E. Bochkareva, L. Frappier, A. M. Edwards, *EMBO J.* **18**, 4498 (1999).
56. E. Bochkareva, S. Korolev, S. P. Lees-Miller, A. Bochkarev, *EMBO J.* **21**, 1855 (2002).

57. S. A. Bastin-Shanower, S. J. Brill, *J. Biol. Chem.* **276**, 36446 (2001).
58. Y. Shamoo, A. M. Friedman, M. R. Parsons, W. H. Konigsberg, T. A. Steitz, *Nature* **376**, 362 (1995).
59. E. Bochkareva, V. Belegu, S. Korolev, A. Bochkarev, *EMBO J.* **20**, 612 (2001).
60. S. Raghunathan, A. G. Kozlov, T. M. Lohman, G. Waksman, *Nature Struct. Biol.* **7**, 648 (2000).
61. T. J. Hubbard, A. G. Murzin, S. E. Brenner, C. Chothia, *Nucleic Acids Res.* **25**, 236 (1997).
62. R. Aasland, A. F. Stewart, T. Gibson, *Trends Biochem. Sci.* **21**, 87 (1996).
63. H. Yang et al., unpublished observations.
64. T. M. Lohman, M. E. Ferrari, *Annu. Rev. Biochem.* **63**, 527 (1994).
65. S. Sigurdsson, K. Trujillo, B. Song, S. Stratton, P. Sung, *J. Biol. Chem.* **276**, 8798 (2001).
66. T. Sugiyama, E. M. Zaitseva, S. C. Kowalczykowski, *J. Biol. Chem.* **272**, 7940 (1997).
67. P. Baumann, S. C. West, *EMBO J.* **16**, 5198 (1997).
68. P. Sung, D. L. Robberson, *Cell* **82**, 453 (1995).
69. P. Sung, *Genes Dev.* **11**, 1111 (1997).
70. A. Shinohara, H. Ogawa, T. Ogawa, *Cell* **69**, 457 (1992).
71. S. L. Gasior et al., *Proc. Natl. Acad. Sci. U.S.A.* **98**, 8411 (2001).
72. F. E. Benson, P. Baumann, S. C. West, *Nature* **391**, 401 (1998).
73. A. Shinohara, T. Ogawa, *Nature* **391**, 404 (1998).
74. J. H. New, T. Sugiyama, E. Zaitseva, S. C. Kowalczykowski, *Nature* **391**, 407 (1998).
75. S. Sigurdsson et al., *Genes Dev.* **15**, 3308 (2001).
76. P. Sung, *J. Biol. Chem.* **272**, 28194 (1997).
77. L. J. Blackwell, J. A. Borowiec, *Mol. Cell. Biol.* **14**, 3993 (1994).
78. R. A. Pfuetzner, A. Bochkarev, L. Frappier, A. M. Edwards, *J. Biol. Chem.* **272**, 430 (1997).
79. E. Bochkareva, L. Frappier, A. M. Edwards, A. Bochkarev, *J. Biol. Chem.* **273**, 3932 (1998).
80. S. J. Brill, S. Bastin-Shanower, *Mol. Cell. Biol.* **18**, 7225 (1998).
81. F. E. Benson, A. Stasiak, S. C. West, *EMBO J.* **13**, 5764 (1994).
82. P. Howard-Flanders, S. C. West, A. Stasiak, *Nature* **309**, 215 (1984).
83. R. Kanaar, J. H. Hoelijmakers, *Nature* **391**, 335 (1998).
84. A. A. Davies et al., *Mol. Cell* **7**, 273 (2001).
85. R. M. Esnouf, *Acta Crystallogr. D. Biol. Crystallogr.* **55**, 938 (1999).
86. R. H. Folmer, M. Nilges, R. N. Konings, C. W. Hilbers, *EMBO J.* **14**, 4132 (1995).
87. P. A. Tucker et al., *EMBO J.* **13**, 2994 (1994).
88. M. P. Horvath, V. L. Schweiker, J. M. Bevilacqua, J. A. Ruggles, S. C. Schultz, *Cell* **95**, 963 (1998).
89. M. Ruff et al., *Science* **252**, 1682 (1991).
90. M. R. Singleton, S. Scaife, D. B. Wigley, *Cell* **107**, 79 (2001).
91. P. J. Loll, A. K. Meeker, D. Shortle, M. Pease, E. E. Lattman, *J. Biol. Chem.* **263**, 18190 (1988).
92. M. Nayal, B. C. Hitz, B. Honig, *Protein Sci.* **8**, 676 (1999).
93. We thank H. Erdjument-Bromage of the Sloan-Kettering Microchemistry Facility for NH<sub>2</sub>-terminal sequence and mass spectroscopic analyses, K.A. McAllister and R. Wiseman for the rat Brca2 cDNA clone, M. Wold for the RPA expression plasmid, the staff of the Cornell High Energy Synchrotron Source MacChess, of the National Synchrotron Light Source X9A and X4A beamlines, and of the Argonne National Laboratory Advanced Photon Source ID19 and COMCAT beamlines for help with data collection. This work was supported by the NIH, the Howard Hughes Medical Institute, the Dewitt Wallace Foundation, the Samuel and May Rudin Foundation, and the Arthur and Rochelle Belfer Foundation. Coordinates have been deposited in the RCSB Protein Data Bank (accession codes 1MIU, 1IYJ, and 1MJE).

**Supporting Online Material**

www.sciencemag.org/cgi/content/full/297/5588/1837/DC1

Materials and Methods

SOM Text

Figs. S1 to S3

17 July 2002; accepted 14 August 2002.

## REPORTS

## A Proton Buffering Role for Silica in Diatoms

Allen J. Milligan and François M. M. Morel\*

For 40 million years, diatoms have dominated the reverse weathering of silica on Earth. These photosynthetic protists take up dissolved silicic acid from the water and precipitate opaline silica to form their cell wall. We show that the biosilica of diatoms is an effective pH buffer, enabling the enzymatic conversion of bicarbonate to CO<sub>2</sub>, an important step in inorganic carbon acquisition by these organisms. Because diatoms are responsible for one-quarter of global primary production and for a large fraction of the carbon exported to the deep sea, the global cycles of Si and C may be linked mechanistically.

Aquatic protists of the class Bacillariophyceae (diatoms) have an absolute requirement for the element Si from which the cell wall, known as a frustule, is produced (1). Silicic acid is taken up by a specific transporter (2) and polymerized intracellularly within a specialized vesicle onto an organic matrix consisting of cationic polypeptides (3). After polymerization, the amorphous hydrated silica with a general formula of Si<sub>n</sub>O<sub>2n-x</sub>(OH)<sub>2x</sub> (where n and x are whole numbers) is moved to the exterior of the cell (1). Diatoms often dominate the phytoplankton assemblage in regions of high productivity where algal nutrients, including Si, are available. These

nutrients are brought to surface waters by rivers, strong vertical mixing, or upwelling and allow diatoms to "bloom" because of their high intrinsic growth rates. Thus, it has been hypothesized that the precipitation of silica must somehow provide an ecological advantage to diatoms. Suggested functions for the siliceous cell wall include serving as an ultraviolet (UV) filter (4), as armor to protect against grazing by zooplankton (5), and as ballast to control water column position (6). It has also been suggested that it is energetically cheaper to construct a cell wall with silica rather than with organic carbon (7). But to date there has been no experimental demonstration of an actual physiological function for silica in diatoms or of its supposed advantage in cell wall formation.

Whereas dissolved nutrients in upwelled water are at concentrations that do not limit growth rates, the concentration of CO<sub>2</sub> in

these waters is not necessarily saturating to photosynthesis (8). This is due to the poor affinity and specificity of the main carboxylating enzyme, ribulose-1,5-bisphosphate carboxylase-oxygenase (RubisCO), which, at the CO<sub>2</sub> and O<sub>2</sub> concentrations of seawater, is much undersaturated with CO<sub>2</sub> and catalyzes the oxygenation as well as the carboxylation of ribulose-1,5-bisphosphate (RubiP) (9). Many photosynthetic protists have overcome this limitation by using a carbon concentrating mechanism (CCM) that augments the CO<sub>2</sub> concentration in the vicinity of RubisCO (10). As part of their CCM, diatoms, like several other microalgae, possess an extracellular carbonic anhydrase (CA) that catalyzes the slow reaction (uncatalyzed half-life ~30 s) between bicarbonate (HCO<sub>3</sub><sup>-</sup>) and CO<sub>2</sub> at their surface (11). The expression of this CA is regulated, increasing at low ambient CO<sub>2</sub> concentrations (12). Although the mechanism of the CCM of diatoms is not completely understood, there is general agreement that the activity of the external CA is an important part of it (13).

Forms of CA that have high catalytic rates ( $k_{cat} > 10^4 \text{ s}^{-1}$ ) require a pH buffer to either provide or receive the proton involved in the reaction between HCO<sub>3</sub><sup>-</sup> and CO<sub>2</sub> because proton exchange with water is relatively slow (14, 15). The proton transfer reaction between the pH buffer and the active site has been shown to be the rate-limiting step in catalysis (16). In seawater the principal natural buffers are, in order of importance, bicarbonate, borate, and silicate. The bicarbonate-carbonate system has been shown to be an effective buffer for some

Department of Geosciences, Guyot Hall, Princeton University, Princeton, NJ 08544, USA.

\*To whom correspondence should be addressed. E-mail: morel@princeton.edu