**The PALB2 DNA binding domain is an intrinsically disordered recombinase.**

Yevhenii Kyriukha*[1], Maxwell B. Watkins*[2], Jennifer M. Redington*[1], Reza Dastvan[1], Jesse Hopkins[2], Nicola Pozzi[1], Sergey Korolev[1].

*[1]Edward A. Doisy Department of Biochemistry and Molecular Biology, Saint Louis University School of Medicine, St Louis, MO; [2]BioCat, Advanced Photon Source, Argonne National Lab, Argonne, IL;*

**these authors contributed equally*

sergey.korolev@health.slu.edu

**Abstract**

Tumor suppressor Partner and Localizer of BRCA2 (PALB2) is a scaffold protein that links BRCA1, BRCA2 and other DNA repair factors and plays a crucial role in homologous recombination (HR). The interaction of PALB2 with DNA strongly enhances the efficiency of HR. Importantly, the PALB2 DNA-binding domain (PALB2-DBD) supports strand exchange of DNA or RNA, a complex multistep reaction carried by only a few protein families like RecA-like recombinases or Rad52. The mechanisms of PALB2 DNA binding and strand exchange are unknown. Through circular dichroism, electron paramagnetic spectroscopy and small-angle X-ray scattering, we found that the PALB2-DBD is intrinsically disordered, even when bound to DNA. Intrinsically disordered proteins (IDPs) are prevalent in the human proteome and have many important biological functions. The advanced level of complexity of the strand exchange reaction significantly expands the functional repertoire of IDPs. Using confocal single-molecule FRET, we found that PALB2-DBD binding leads to oligomerization-dependent compaction of DNA. We hypothesize that PALB2-DBD uses a “chaperone-like” mechanism to aid formation and resolution of complex DNA and RNA multichain intermediates during DNA replication and repair. Similar DNA-binding IDRs may represent a novel class of functional domains evolved in eukaryotic nucleic acids metabolism complexes.

**Introduction.**

Homologous recombination (HR) is essential for the exchange of genetic information between maternal and paternal alleles and for proper chromosome segregation during meiosis, for non-mutagenic repair of large chromosome aberrations such as double-stranded (ds) DNA breaks (DSB), for repair of stalled replication forks, and for a plethora of other DNA transactions (Cox, 1999, 2001; Krogh and Symington, 2004; San Filippo et al., 2008). HR is almost exclusively supported by a single family of RecA-like recombinases, such as phage UvsX, bacterial RecA, archaea RadA and eukaryotic Rad51 and DMC1 proteins ((Kowalczykowski, 2015) and references within). Their sequences and structures are highly conserved, due to the complexity of a multistep HR reaction involving formation of an ATP-dependent protein filament on single-stranded (ss)DNA that is capable of searching for a homologous dsDNA and catalyzing strand displacement and strand exchange (Cox, 2003; Morrical, 2015; Roca and Cox, 1990; San Filippo et al., 2008). The complexity of this mechanism limits the evolution of more specialized recombinases. Instead, there are numerous partner proteins regulating organism- and pathway-specific activities of RecA-like recombinases at multiple levels (Beernink and Morrical, 1999; Cox, 2007; Liu and Heyer, 2011; Sung and Klein, 2006; Symington, 2002).

While RecA-like recombinases recombine long DNA regions with high fidelity, they are not efficient in case of strand exchange between short fragments or between DNA and RNA to aid complex multichain transactions during such processes as replication fork reversal or transcription/replication collision. Several other protein families have been reported to support DNA-pairing and strand exchange. Those include exonuclease complexes with DNA-annealing proteins, such as *E. coli* RecE/RecT, bacteriophage λ Redα/ Redβ, and viral UL12/ICP8 complexes(Hall and Kolodner, 1994; Morrical, 2015; Murphy, 2012; Reuven et al., 2003). Another strand exchange protein is eukaryotic Rad52 (Bi et al., 2004; Lee et al., 2014; Mazina et al., 2017), which is also a major positive regulator of Rad51 in yeast (Mortensen et al., 2009; Sung, 1997). Both systems include ring-shaped DNA-binding oligomers (Passy et al., 1999; Zhang et al., 2009c). Rad52 forms an oligomeric toroidal ring that binds ss- and dsDNA to promote promiscuous strand exchange (Kagawa et al., 2002; Saotome et al., 2018; Singleton et al., 2002). Rad52 is more efficient in strand exchanges between ssRNA and dsDNA than Rad51 and is proposed to repair breaks occurring during collision of replication with transcription using newly transcribed mRNA as a homologous template (Mazina et al., 2017; Yasuhara et al., 2018). The necessity to process complex multichain DNA and RNA intermediates dynamically present in chromatin during all major nucleic acid metabolism reactions is further underscored by reports of other protein families with strand annealing and exchange activities. Those include FANCA

protein of the Fanconi anemia (FA) complex critical for DNA interstrand crosslink repair(Benitez et al., 2018) and members of RNA processing FET proteins: pro-oncoprotein TLS/FUS and the human splicing factor PSF/hPOMp100 (Akhmedov and Lopez, 2000; Baechtold et al., 1999; Schwartz et al., 2015).

We discovered that the Partner And Localizer of BRCA2 (PALB2) protein is capable of DNA strand exchange (Deveryshetty et al., 2019). Human PALB2 is a 1186 amino acid-long scaffold protein that links BRCA1 and BRCA2 during initiation of HR, where BRCA2 stimulates RAD51 filament formation (Davies et al., 2001; Ducy et al., 2019; Moynahan et al., 2001; Sharan et al., 1997; Venkitaraman, 2000, 2014; Xia et al., 2006). PALB2 interacts with numerous other chromatin and transcription factors, as well as with RAD51 itself and its paralogs (Ducy et al., 2019). The N-terminal domain encompassing amino acids 1-195 interacts with DNA (referred to hereafter as PALB2-DBD) (Fig. 1A)(Buisson et al., 2010; Dray et al., 2010). Alanine substitution of major DNA-binding amino acids 146-RRKK-149 decreases radiation-induced RAD51 foci formation and HR in cells by 50% (Deveryshetty et al., 2019). Surprisingly, we discovered that this domain alone promotes RAD51-independent strand exchange *in vitro* (Deveryshetty et al., 2019). PALB2 lacks any sequence or predicted structural similarities with Rad51 and Rad52 and likely uses a novel mechanism for the strand exchange.

The PALB2 amino acid sequence has a low complexity sequence composition, characteristic for scaffold IDPs (van der Lee et al., 2014), with some regions characterized by a low $\alpha$-helical propensity (Fig. 1A) (https://MobiDB.bio.uniprot.it/Q86YC2 and https://alphafold.ebi.ac.uk/entry/Q86YC2) (Jumper et al., 2021; Piovesan et al., 2023) except two documented folded regions that include the N-terminal $\alpha$-helix (aa 10-40) (Song et al., 2018) and the C-terminal WD40 domain (aa 835-1186) (Oliver et al., 2009). The N-terminal $\alpha$-helix forms an antiparallel coiled-coil either with itself (referred to hereafter as PALB2-cc), forming a homodimer, or with BRCA1 $\alpha$-helix (Buisson and Masson, 2012; Sy et al., 2009; Zhang et al., 2009b). The remaining region of PALB2-DBD (aa 41-195, referred to hereafter as $\Delta$40-DBD) is a neutral peptide with a balanced content of hydrophobic and both positively and negatively charged amino acids characteristic for a context-dependent structural organization (see Discussion below). PALB2-DBD's solution properties, oligomerization, a localized DNA binding site and, most importantly, a complex strand exchange activity, are not typical for IDRs and suggest that oligomerization and/or DNA binding can stimulate formation of additional $\alpha$-helixes and folding of PALB2-DBD.

Here we present evidence that PALB2-DBD is an intrinsically disordered domain or region (IDR) with and without DNA. We characterize its interaction with ssDNA and suggest a plausible

“chaperone-like” novel strand exchange mechanism. PALB2-DBD is the first example of an IDR capable of a multistep reaction involving two complex substrates. This significantly expands the structural and functional repertoire of IDRs, which represent a third of human proteome (van der Lee et al., 2014). Such a mechanism can have a broad implications for protein and RNA chaperones, ~30% and 50% of which are estimated to be IDRs with only a few mechanistically characterized (Holmstrom et al., 2019; Koculi et al., 2006; Vo et al., 2009). Many eukaryotic DNA repair proteins, e.g. BRCA1 and BRCA2, include DNA-binding IDRs of unknown function. We hypothesize that such regions represent a novel functional class of DNA repair effectors that can aid transactions between DNA and RNA strands to form or resolve local transient multistrand intermediates occurring near the replication fork, DNA breaks and transcription sites.

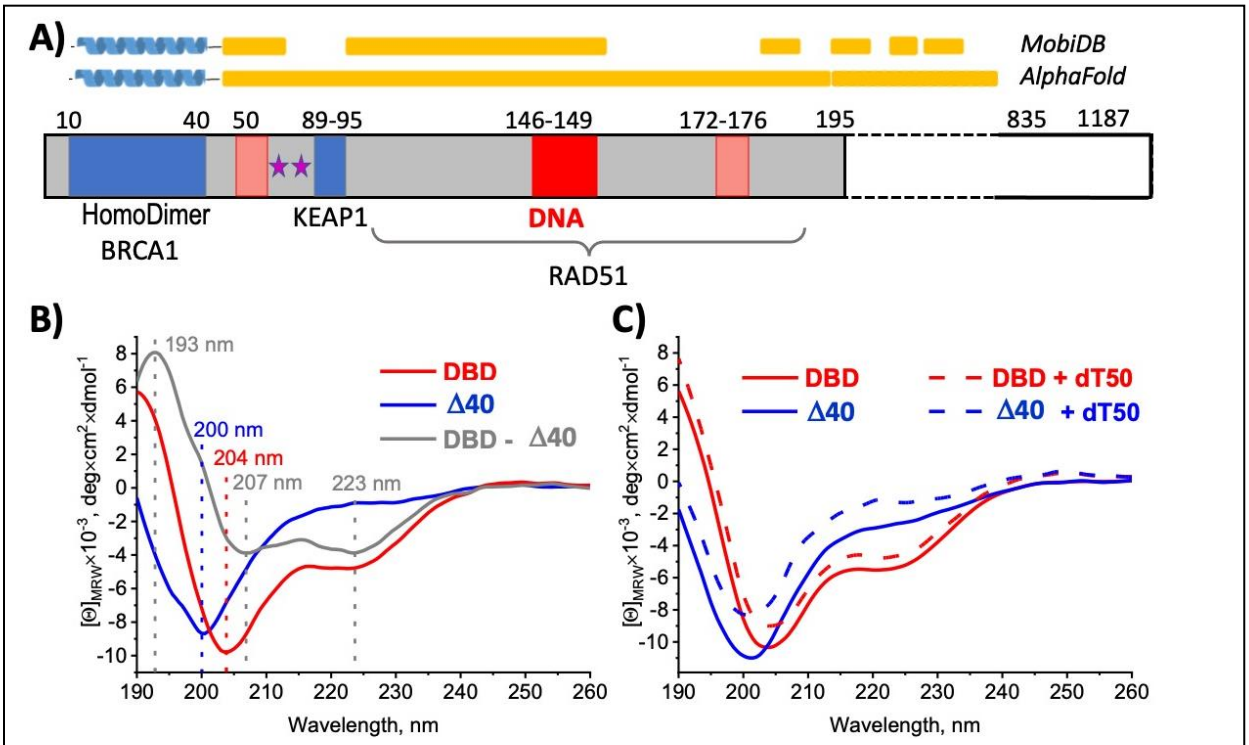


***Figure 1. Amino acid sequence features and CD spectroscopy of PALB2-DBD A)*** *Schematic representation of PALB2 sequence domains. Numbers on the top correspond to the amino acid sequence number. Grey area shows the PALB2-DBD with blue regions depicting known protein interaction sites, red region depicting the major DNA-binding site, and pink depicting minor DNA-binding sites. Stars show phosphorylation sites. Yellow bars on top shows disordered regions accordingly to MobiDB or AlphaFold prediction.* ***B)*** *CD spectra of PALB2-DBD is shown in red, of Δ40-DBD in blue and the subtraction of two spectra in grey. Peaks for each spectra are marked by dashed lines and numbers of the corresponding color.* ***C)*** *CD spectra of PALB2-DBD are shown in red and of Δ40-DBD in blue with and without ssDNA shown by dashed and solid lines, respectively.*

**Results.**

*I.* PALB2-DBD *is disordered.* There are several sequence regions with a low α-helical propensity inside Δ40-DBD that can form a globular structure upon oligomerization of DNA binding. We probed the presence of secondary structure elements using circular dichroism (CD) spectroscopy (Fig. 1B). Spectrum from the PALB2-DBD fragment is characterized by peaks corresponding to α-helical and random coil (RC) structures (Fig. 1B). The spectrum of the Δ40-DBD fragment lacking the N-terminal α-helix does not show the peak corresponding to α-helix, while subtraction of PALB2-DBD and Δ40-DBD spectra resulted in a spectrum typical for an α-helical-only structure. Therefore, CD spectroscopy results demonstrate a lack of secondary structure elements beyond the N-terminal coiled-coil region in PALB2-DBD. To investigate whether additional secondary structure elements can be stabilized by interactions with DNA, we measured CD spectra from PALB2-DBD and Δ40-DBD in the presence of excess ssDNA (Fig. 1C). Neither the spectrum of PLAB2-DBD nor of Δ40-DBD were changed upon addition of excess ssDNA. Therefore, CD spectroscopy results suggest that Δ40-DBD is an intrinsically disordered region (IDR). Importantly, despite a localized major DNA binding site, a feature typical for globular proteins, it interacts with DNA as an IDR.

These results do not completely rule out potential formation of additional α-helixes upon coiled-coil-mediated oligomerization of PALB2-DBD. Since IDR are characterized by high flexibility, we probed for the local flexibility of different regions using continuous wave electron paramagnetic spectroscopy (cwEPR), particularly informative for structural characterization of IDPs and their interactions (Fig. 2)(Drescher, 2012; Longhi et al., 2011). Four cysteines present in human PALB2-DBD were alternatively mutated to create single cysteine variants at positions 11, 57, 77, 162 (Fig. 2D). Two additional cysteines were introduced in the vicinity of the major DNA binding site (S136C, S145C). The nitroxide-containing spin label MTSL was covalently attached to the cysteine thiol group for EPR analysis. The EPR spectrum was measured for each mutant alone and for positions 11, 136 and 145 in the presence of ssDNA (Fig. 2B). Most spectra correspond to a highly flexible conformation of local peptide in the label vicinity. Values of h(+1)/h(0) ratio correspond to a typical value for other studied IDPs (Fig. 2C)(Longhi et al., 2011). The exception is position 11 at the beginning of the PALB2-cc. A lower value corresponds to a more rigid structure near the coiled-coil region. The addition of $dT_{50}$ causes a notable change in flexibility at position 145 immediately next to the major DNA binding site (146-149). This region is still more flexible even in the presence of DNA than at position 11. There was no change in flexibility of the region around position 11 and 136. Thus, DNA binding does not affect the mobility

of the complex nor the flexibility of the structure beyond the immediate vicinity of the DNA-binding site. High flexibility of the probed regions further confirms a lack of any secondary structures formed inside the Δ40-DBD sequence, even within the oligomeric PALB2-DBD and in DNA-bound state.

The described experiments revealed an intrinsically disordered and highly flexible structure of PALB2-DBD beyond the N-terminal α-helix.

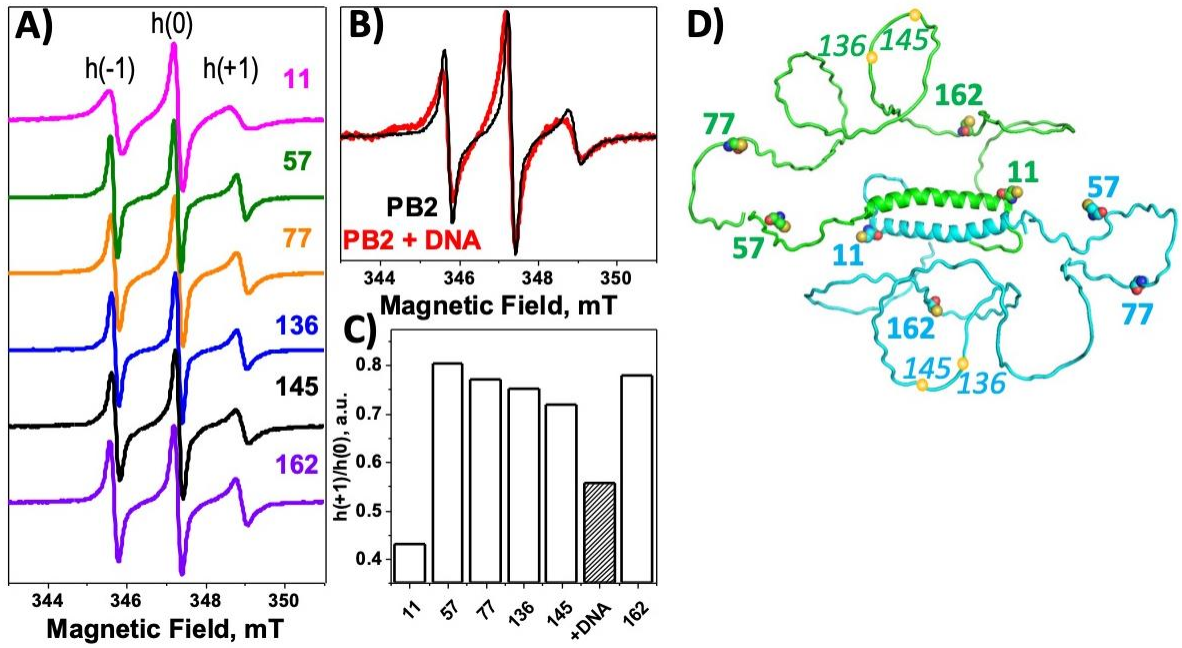


***Figure 2. cwEPR spectroscopy of PALB2-DBD. A)*** *EPR spectra of PALB2-DBD labeled at different positions indicated by the number on the right,* ***B)*** *Overlay of spectra of PALB2-DBD labeled at position 145 in free (black) and DNA-bound (red) states.* ***C)*** *h(+1)/h(0) values for each position of the attached label. Texture-filled bar corresponds to the signal from position 145 in the presence of DNA. No changes of the ratio were observed upon DNA binding for other positions.* ***D)*** *A hypothetical conformation of PALB2-DBD dimer generated with AlphaFold program. Monomers are shown in cyan and green. Locations of cysteine positions used for labeling are shown by spheres and numbers.*

*II. PALB2-DBD dimer forms a compact molten globule.* IDPs differ by functionally relevant structural organization ranging from compact molten globule (MG) or preMG structures to extended random coil (RC) conformations. MGs are characterized by a relatively small hydrodynamic radius ($R_h$) only two-three times larger than the radius of a folded globular structure of the same MW (Habchi et al., 2014). RC structures have up to ten times larger $R_h$ than that of folded protein. The structural organization of PALB2-DBD is further complicated by dimerization with two IDRs attached to opposite ends of a 40 Å-long antiparallel coiled-coil.

We evaluated the compaction of a monomeric Δ40-DBD by a size exclusion chromatography (SEC) to compare elution time of the protein in a DNA-binding buffer with that of

in the presence of the strong denaturant Guanidinium hydrochloride (GdmHCl), which forces RC formation (Supplementary Fig. S1A). The protein in 6M GdmHCl elutes significantly earlier than the protein in a non-denaturing buffer, suggesting a preMG or MG compact conformation of Δ40-DBD. In fact, the elution volume under nondenaturing conditions corresponds to a globular protein with a molecular weight only twice that of Δ40-DBD (Supplementary Fig. S1B). From this, we conclude that the Δ40-DBD region lacks secondary structure elements but forms a relatively compact MG structure.

Previously, we reported an apparent tetrameric form of PALB2-DBD based on SEC and DNA binding stoichiometry (Deveryshetty et al., 2019). Considering a disordered structure of PALB2-DBD that leads to overestimation of the molecular weight by SEC, we analyzed PALB2-DBD using analytical centrifugation with a sedimentation velocity method. PALB2-DBD alone sediments as a dimer even at relatively high protein concentration (Supplementary Fig. S2A). Additionally, the molecular weight was directly measured with multiangle light scattering (MALS) combined with SEC at two different solution conditions with high (0.5M) and low (0.16 M) NaCl concentrations (Supplementary Figs. 2B, C). Two salt conditions were used since PALB2-DBD is more soluble and less prone to aggregation in high salt buffer (0.5-1.0M NaCl), while the DNA-binding is strongly inhibited by elevated NaCl concentration (Deveryshetty et al., 2019). MALS measurements unequivocally confirmed a dimeric form of PALB2-DBD under both conditions.

Small angle X-ray scattering (SAXS) analysis provides important structural characteristics of proteins in solution, particularly valuable in case of IDPs not amenable to other structural methods. We analyzed PALB2-DBD in two different buffer conditions using SEC-MALS-SAXS (Table S1, Supplementary Figs S3, S4). SAXS measurements of PALB2-DBD samples in two different buffers with high- and low-salt content resulted in a solution scattering typical for IDPs (Fig. 3A).

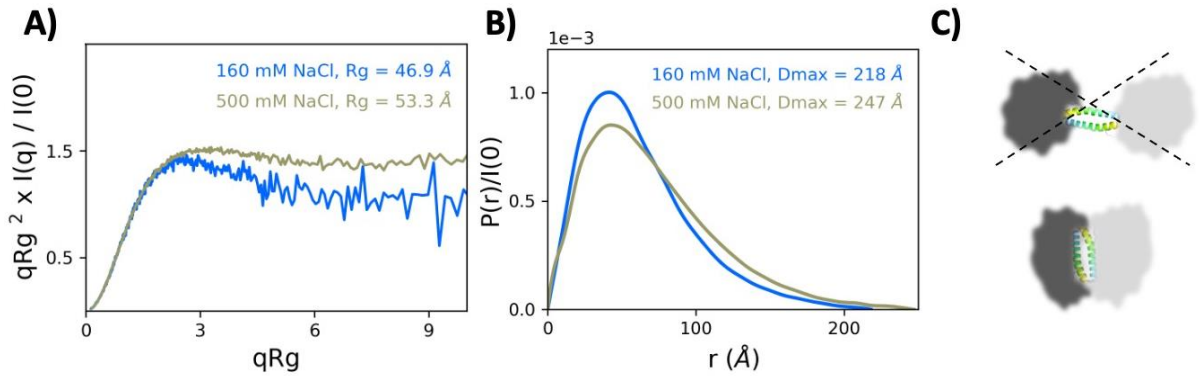


***Figure 3. SAXS/MALS analysis of PALB2-DBD. A)*** *SAXS data for high (green) and low (blue) salt buffers, visualized as a dimensionless Kratky plot. Data were binned logarithmically by factor of 4 for visual clarity.* ***B)*** *Pairwise distance distribution functions (P(r)) from SAXS measurements of PALB2-DBD in low salt (blue) and high salt (green) buffers.* ***C)*** *Two hypothetical models of IDRs attached to opposite ends of coiled-coil linker. A single domain length in P(r) distribution (B) rules out a "dumbbell" model (top) and suggests a compact dimer conformation (bottom).*

As described above, two IDRs are located at opposite ends of 40 Å-long antiparallel coiled-coil structure and can form two independent or spatially distinct MGs in the PALB2-DBD dimer. However, only one dominant length scale is observed in the pairwise distance distribution (*P(r)*) measured in SAXS (Fig. 3B), ruling out the dumbbell model with two distinct MGs (Fig. 3C). The small radius of gyration ($R_g$) values obtained in SAXS experiments also evidence a rather compact organization of the entire dimer (Fig. 3A). Taking these observations together, our data suggest that each monomer forms a compact MG stabilized by intramolecular interactions. Similar interactions between two IDRs are likely to contribute to dimerization synergistically with a coiled-coil interaction, thus leading to a compact structure of the entire dimer.

To get further insight into the structural organization of PALB2-DBD using experimental SAXS results, we performed molecular structure modeling using the SASSIE program (Curtis et al., 2012). Random structures were generated using the Complex Monte-Carlo algorithm in SASSIE with fixed N-terminal $\alpha$-helixes and a dimeric coiled-coil interface (amino acids 11-40) as determined by NMR (Song et al., 2018). The starting model of a dimer was generated by the AlphaFold program using the ColabFold Google server (Mirdita et al., 2022) (Fig. 2D). The resulting trajectory yielded 17,353 structures, from which a sub-ensemble of structures (selected using GAJOE (Bernado and Svergun, 2012)) provided a good fit ($\chi^2$=1.14) to the experimental data of the system in 0.5M NaCl buffer (Fig. 4A). The poll of the best-fit ensemble is bimodal with

the majority with even lower $R_g$ = 45 Å and a second population with $R_g$ = 62 Å. Importantly, the experimental Guinier $R_g$ of 54.9 Å is significantly lower than the average $R_g$ = 75 Å of the entire pool of all randomly generated structures. Additionally, we analyzed data using a molecular form factor (MFF) fitting protocol developed specifically for IDP's (http://sosnick.uchicago.edu/SAXSonIDPs)(Clark et al., 2020; Riback et al., 2017). Result yields a Flory exponent $\nu$ = 0.52, which describes a polymer dimension dependence on its length ($R_g \sim N^{\nu}$). This value of $\nu$ is lower than typical reported values for IDP's in good solvent (~0.54-0.55)(Clark et al., 2020; Riback et al., 2017) and is suggestive of a relatively more compact overall conformation, assuming that due to the almost end-to-end dimerization this system behaves relatively similarly to a single-chain IDP with a total length of 400 aa.

Fitting data obtained at a low salt concentration (0.16M NaCl) was more difficult. Initial unguided SASSIE modeling resulted in relatively poor fit with $\chi^2$=2.5 (Fig. 4C), seemingly due to the propensity of the trajectory to more extended structures. A guided fit with an imposed restriction on average $R_g$ during modeling was improved with $\chi^2$=1.83 but was still worse than that for a high salt buffer. A similar bimodal distribution of theoretical $R_g$ was obtained even in the case of the guided modeling (Fig. 4E). Alternatively, we used the AllosMod-FoXS program (Schneidman-Duhovny et al., 2010; Weinkam et al., 2012) with concurrent energy landscape modeling (i.e. multiple different starting points) in an attempt to perturb a structure to fit the experimental data (Supplementary Fig. S5). Using several different starting models, we generated 100 concurrent trajectories, each generating 100 models. Pooling all the structures from the two trajectories and fitting a sub-ensemble using GAJOE (Bernado et al., 2007; Tria et al., 2015) resulted in a better fit for both experimental datasets. Although this method does not allow for fixing of the dimeric interface, the resulting structures retained an overall similar orientation of the $\alpha$-helices. This modeling also resulted in bimodal $R_g$ distribution of best-fit structural ensembles.

These data further support a highly compacted dimeric PALB2-DBD structure of two IDRs. Notably, compaction parameters at 0.5M NaCl, $R_g$ = 58.4 Å and Flory scaling exponent $\nu$ = 0.52 (Supplementary Table S2), are similar to corresponding values of those reported for an unstructured PNt domain with slightly smaller size of a single chain (335 aa) than that of PALB2-DBD dimer (Riback et al., 2017). In the case of the low salt condition, both the $R_g$ (47.5 Å) and $\nu$ (0.47) values are even smaller than in the high salt condition, suggesting a role for polar interactions in structure compaction. It is worth noting that the $R_g$ values calculated from fitting the MFF (55.0 Å and 49.1 Å for the high and low salt buffer conditions, respectively) are slightly larger than the Guinier $R_g$ values referenced previously but are relatively close in both cases and as such would not affect our interpretation of the analysis described here.

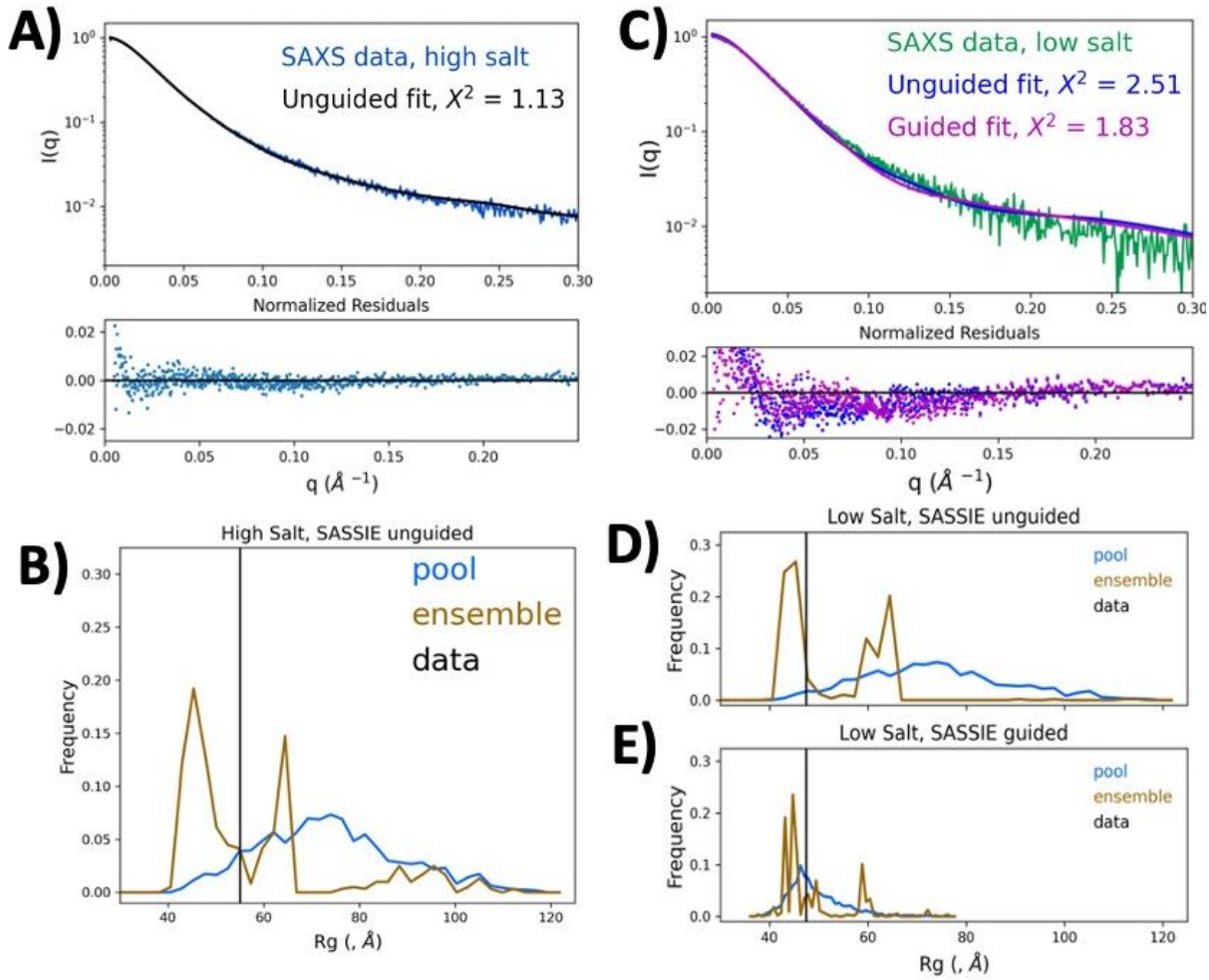


***Figure 4. Modeling of the PALB2-DBD structural ensemble. A)*** *Fit of experimental distance distribution from SAXS experiment in 0.5M NaCl buffer with theoretical distribution calculated from unguided SASSIE models. Normalized residual errors are shown at the bottom.* ***B)*** *$R_g$ distribution of the entire pool of 17,000 structures is shown in blue, experimental $R_g$ by black bar and the distribution of $R_g$ of the best fit pool structures in brown.* ***C)*** *Similar to (A) but for data obtained from samples in 0.16M NaCl buffer fit to unguided models (blue) and guided models with restricted $R_g$ (magenta).* ***D)*** *$R_g$ distributions for unguided model pools.* ***E)*** *same as (D) for guided models.*

*III. smFRET analysis revealed that PALB2-DBD dynamically condenses DNA*. The recombination mechanism depends on conformational changes of DNA bound to recombinase. RecA-like recombinases form nucleoprotein filaments stretching out ssDNA (Chen et al., 2008; Ogawa et al., 1993; Yang et al., 2020). Rad52 wraps ssDNA around a toroidal oligomeric structure (Grimme et al., 2010; Saotome et al., 2018; Singleton et al., 2002). Previously, we suggested a Rad52-like mechanism based on an intensity-based Förster Resonance Energy Transfer (FRET) experiment

where protein interaction led to an increased FRET between Cy3 and Cy5 dyes located at the ends of doubly labeled $dT_{40}$ or $dT_{70}$ oligonucleotides (Deveryshetty et al., 2019). Titration by PALB2-DBD led to a higher FRET efficiency resembling the results of a similar experiment with Rad52 (Grimme et al., 2010). However, these measurements were performed on an ensemble averaged in a fluorimeter, making it difficult to pinpoint the species and sequence of events responsible for such a phenomenon, particularly considering the disordered nature of PALB2-DBD and oligomerization. To address the limitation of our previous studies, we performed comparable experiments utilizing single-molecule (sm)FRET. Through a confocal setup, we determined the FRET efficiency of freely diffusing single DNA molecules labeled with the FRET pair Cy3/Cy5 in the absence and presence of PALB2-DBD. Moreover, we used Pulsed Interleaved Excitation (PIE) to calculate stoichiometry for each molecule, as this quantity enables the identification of molecules containing the correct 1:1 ratio of donor and acceptor (Fig. 5A) and the detection of other photophysical phenomena such as bleaching, blinking, and protein-induced fluorescent enhancement (PIFE), as documented earlier (Chinnaraj et al., 2022; Kapanidis et al., 2004; Kudryavtsev et al., 2012; Pozzi et al., 2016). Measurements were performed with several ssDNA substrates with fluorescent dyes placed at different distances. Those include $dT_{50}$ with Cy3 at 5' end and Cy5 at positions 25 or 50, and $dT_{70}$ labeled at 5' and 3' ends. Cy3-$dT_{70}$-Cy5 and Cy3-$dT_{50}$-Cy5 substrates in solution show a single, symmetric distribution with mean FRET of 0.11 (Fig. 5A) and 0.18 (Fig. 5B), respectively, in agreement with published results (Murphy et al., 2004). The addition of PALB2-DBD led to a significant increase in mean FRET. This confirms observations from our previous studies (Deveryshetty et al., 2019). Notably, two distinct FRET populations were observed in the presence of the protein, corresponding to two different states of DNA bound to PALB2-DBD, one significantly more condensed than the other with mean FRET of 0.34 and 0.71 in case of Cy3-$dT_{70}$-Cy5 (Fig. 5A). The mean FRET value for each peak is invariant for a given DNA fragment, while the population of each state changes. A medium FRET state is a predominant specie at 1 μM PALB2-DBD, while at 3 μM, high FRET is a major peak (Fig. 7A). Titration of $dT_{50}$ by PALB2-DBD (Fig. 5B) revealed similar distributions of two populations of invariant mean FRET peaks of 0.45 and 0.79 with a high FRET peak becoming more predominant at high protein concentration. These results imply that PALB2-DBD binds to multiple sites within the same DNA and, by doing so, dynamically controls its conformation. They also imply that the PALB2-DBD concentration is an important factor in controlling DNA conformation.

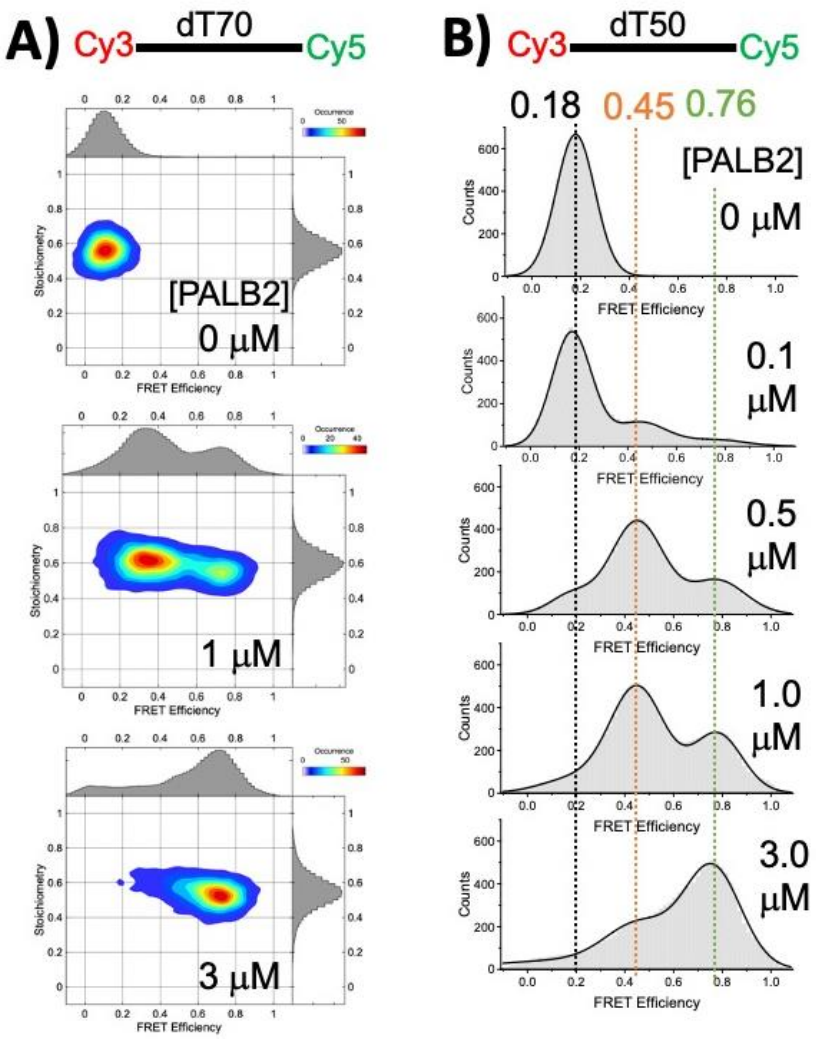


***Figure 5. Bimodal DNA compaction revealed by confocal smFRET. A)*** *2D plots of stoichiometry versus FRET efficiency for 100 pM Cy3-dT$_{70}$-Cy5 alone and in the presence of 1.0 μM and 3.0 μM PALB2-DBD;* ***B)*** *FRET histograms (grey) of 100 pM Cy3-dT$_{50}$-Cy5 titrated by PALB2-DBD. Black lines show fitting of data using three species as gaussian distributions. Protein concentrations are show on each graph. Mean FRET values of major peaks are shown with color corresponding to each peak.*

A similar trend was observed with a dT$_{50}$ substrate with Cy3 and Cy5 labels placed 25 nucleotides apart (Fig. 6A). Mean FRET for free DNA is 0.45, the major peak at 1 μM PALB2-DBD is at FRET of 0.71 and at 3 μM PALB2-DBD at FRET of 0.89. While it is difficult to deconvolute two separate peaks at each protein concentration due to the initially high FRET for free DNA, the compaction of DNA is also significant. A similar trend of compaction proportional to the number of nucleotides between labels is observed in all three cases. For example, changes between a peak of free DNA and that of the first peak in complex with PALB2-DBD were from 0.11 to 0.34, from 0.18 to

0.45 and from 0.5 to 0.7 for distances of 70, 50 and 25 nt, respectively. This ruled out a Rad52-like wrapping model of DNA binding, in which the FRET change between remote ends should be greater than that between labels situated closer in a sequence.

The broad distributions of each peak suggest a random coil model of ssDNA conformation where an average distance between ends is restricted due to interactions with two major DNA-binding sites in a compact PALB2-DBD dimer. In such a model, two random parts of ssDNA should interact with two DNA binding sites situated within an average $R_g$ distance from each other. This model explains an invariant position of the peak with an intermediate mean FRET. The mechanism of an additional compaction corresponding to a high FRET peak remains to be further investigated. One of the hypotheses is that it corresponds to a higher oligomeric state of PALB2-DBD. To evaluate the role of PALB2-DBD oligomerization for DNA compaction, we tested a Δ40-DBD variant lacking the coiled-coil dimerization interface using Cy3-$dT_{50}$-Cy5 (Fig. 6B). DNA compaction was also observed in this case, albeit at higher protein concentrations, reflecting a significantly reduced DNA binding affinity (Deveryshetty et al., 2019). Even at 10 μM protein, there was a fraction of free DNA in solution. Importantly, FRET distributions were significantly broader

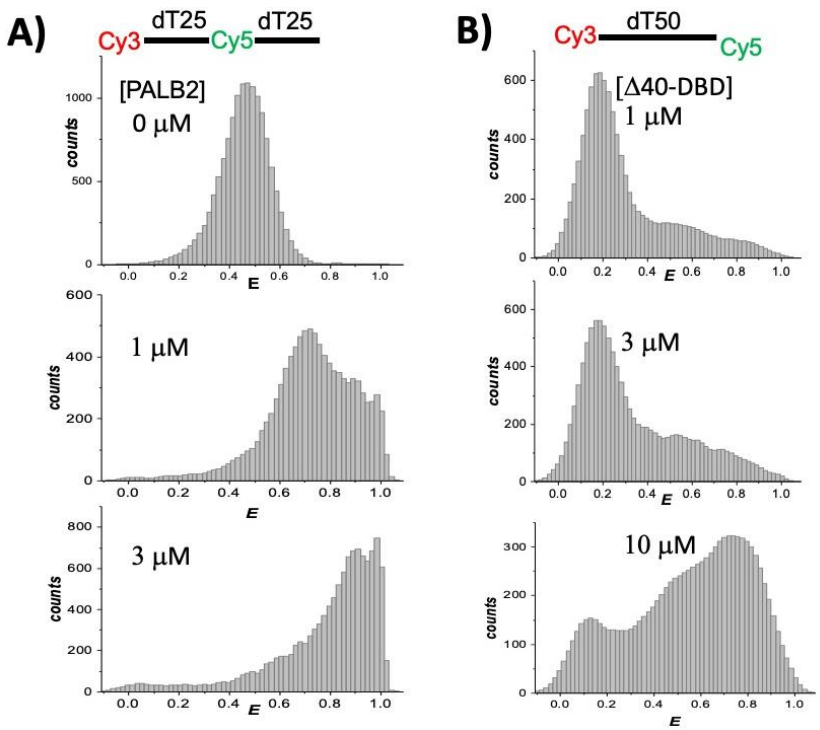


***Figure 6. smFRET measurements of ssDNA compaction. A)*** *FRET histograms of 100 pM $dT_{50}$ with Cy3 placed at 5' end and Cy5 attached at position 25 of $dT_{50}$ substrate alone and in the presence of 1 μM and 3 μM PALB2-DBD;* ***B)*** *FRET histograms of Cy3-$dT_{50}$-Cy5 titrated by Δ40-DBD. Protein concentrations are shown on each graph.*

and more heterogeneous corresponding to multiple species of DNA conformations. Lack of two distinct peaks in the case of Δ40-DBD indirectly supports oligomerization-dependent DNA compaction. In such a case, the medium FRET peak corresponds to interaction with PALB2-DBD dimers and the high FRET peak to interaction with a tetramer. Monomeric Δ40-DBD will form dimers, trimers and tetramers, resulting in a broad distribution of DNA conformations specific for each oligomeric state, as observed experimentally.

The tetramerization hypothesis is supported by previously observed 1:4 stoichiometry of DNA binding in fluorescence polarization and ensemble FRET assays (Deveryshetty et al., 2019), but is contradicted by a dimeric state of PALB2-DBD measured by AUC and MALS. We repeated sedimentation velocity experiments with PALB2-DBD in the presence of various ssDNA substrates (Supplementary Fig. S6A, B). DNA was labeled with Cy3 to distinguish its position from that of protein-only. Measurements were repeated at two different protein concentrations corresponding to 1:2 and 1:4 DNA:protein ratios. In the case of Cy3-$dT_{50}$, two DNA peaks were observed. A larger peak overlapped with that of the PALB2-DBD dimer and a smaller peak with a molecular weight comparable to that of the tetrameric PALB2-DBD. Although free Cy3-$dT_{50}$ sediments similar to that of PALB2-DBD dimer, most DNA should be in protein-bound form, as the concentration of components is several times above the DNA binding $K_d$. To distinguish free from bound DNA, we repeated the experiment with shorter Cy3-$dT_{25}$ (Suppl. Fig. S6B). There was no peak corresponding to free DNA in solution at both stoichiometries. Interestingly, we observed a small peak corresponding to a tetrameric structure as well. These data suggest that DNA readily interacts with PALB2-DBD dimers and stimulates partial tetramerization of PALB2-DBD. Results also rule out the independent binding of two dimers to neighboring parts of a long DNA molecule, since similar tetrameric structures were formed with 25- and 50-mer ssDNA. These results only partially support the hypothesis that the high FRET state corresponds to DNA bound to tetramers. Indeed, there is only a minor population of tetramers in AUC, while it is a major peak in smFRET experiments at similar protein concentrations (3 μM in smFRET and 4 and 8 μM in AUC). Importantly, DNA:protein stoichiometries differ in two experiments with a large protein excess versus DNA with only 100 pM DNA in smFRET and 1 μM in AUC. We repeated smFRET measurements with an addition of 0.5 μM unlabeled DNA (Supplementary Fig. S6D). This significantly reduced the high FRET efficiency fraction, approximately recapitulating the ratio between dimeric and tetrameric species in AUC experiments. Therefore, tetramerization requires a large excess of free protein over DNA in solution.

**Discussion.**

We present evidence that PALB2-DBD is an intrinsically disordered domain or region that forms a dimer stabilized by both a coiled-coil interface between N-terminal $\alpha$-helices and interactions between IDRs. While PALB2-DBD is comprised of a sequence of low complexity, it was unexpected that an unstructured protein can support an elaborate multi-substrate and multistep mechanism of strand exchange. Indeed, $\Delta$40-DBD domain has a balanced composition of hydrophobic and both positively and negatively charged amino acids (Supplementary Fig. S7), that can undergo "context dependent" folding affected by protein-protein or DNA protein interactions (Uversky, 2002; Uversky et al., 2000). Analysis using the CIDER server (Holehouse et al., 2017) classified it at the boundary between "*Janus sequences: collapsed or expanded – context dependent*" and "*Weak polyampholytes & polyelectrolytes: Globule & tadpoles*". CD and EPR spectroscopy results confirm the lack of stable structural organization of the PALB2-DBD beyond the N-terminal $\alpha$-helix for both monomeric and dimeric forms as well as upon ssDNA binding. Therefore, the PALB2-DBD represents a novel class of IDP that can perform a complex multistep reaction while simultaneously interacting with two substrates. The strand exchange mechanism by PALB2-DBD is likely to differ drastically from those supported by known strand exchange proteins and is defined by structural features of the protein.

*The PALB2-DBD dimer forms a compact MG*. SEC and SAXS results revealed a compact structure of each monomer ($\Delta$40-DBD) and of the entire PALB2-DBD dimer. Such compaction is likely achieved through favored inter- and intramolecular interactions between two domains. This does not represent a collapsed state of disordered protein chains. The Flory scaling exponent describing a polymer dimension dependence on its length $\nu$ is around 0.5, which is expected for a non-self-avoiding random coil and corresponds to a so-called $\Theta$ solvent where intermolecular and solvent interactions counterbalance each other (Clark et al., 2020). This can explain the relatively good solubility and stability of the PALB2-DBD in aqueous solution. A noticeable reduction of $R_g$ and $\nu$ in a low salt buffer suggests a strong contribution of polar intermolecular interactions to compaction of the PALB2-DBD. The same intramolecular interactions are likely to stabilize and compact the entire dimer. The interactions of two IDRs are weak, since the $\Delta$40-DBD is monomeric in solution. However, they are synergistic with the coiled-coil interactions. The dimerization of isolated coiled-coil helices is rather weak, with $K_d$=80 $\mu$M as reported by an NMR titration experiment (Song et al., 2018), while in our experiments the dimer remains intact even during non-equilibrium gel filtration at 100 times lower concentrations. Moreover, a single dominant length scale observed in the pairwise distance distribution (*P(r)*) measured in SAXS

(Fig. 3B) further supports interactions of entire domains and not only coiled-coil helixes. It is tempting to hypothesize that protein-protein interactions by IDRs can contribute to heterooligomerization within DNA repair complexes. For example, PALB2 binds BRCA1 through the same N-terminal $\alpha$-helix that forms a coiled-coil interface with BRCA1 $\alpha$-helix (Zhang et al., 2009a; Zhang et al., 2009b) and these interactions can be strengthened by attractions between IDRs of two proteins.

*Compaction of ssDNA by PALB2-DBD*. The conformation of DNA bound to protein defines the mechanism of strand exchange reaction. RecA-like recombinases stretch ssDNA, keeping nucleotide triplets with exposed bases in a B-form-like geometry while bending the DNA helix and transiently separating the two strands (Yang and Pavletich, 2021). The mechanism of Rad52 recombination is less well understood. It wraps ssDNA around the toroidal oligomer backbone, keeping bases exposed for base pairing (Grimme et al., 2010; Rothenberg et al., 2008). Binding of dsDNA along the toroidal oligomer can potentially distort helical structure, leading to transient strand separation and strand exchange. We observed significant compaction of ssDNA upon PALB2-DBD binding, ruling out a RecA-like mechanism. Changes of FRET efficiency upon PALB2-DBD binding are proportional to the number of nucleotides between labels, thus ruling out a wrapping model. Therefore, the mechanism of PALB2-DBD interaction with ssDNA differs from those of RecA and Rad52.

Wide FRET distributions of PALB2-DBD-bound ssDNA and the invariant positions of observed peaks suggest random coil folding with additional distance restrictions. Interaction of ssDNA at two random positions with two DNA binding sites of the dimer (and potentially, with minor DNA binding sites as well) situated within an average $R_g$ distance from each other should lead to the observed distribution, as illustrated in Figure 7. A DNA compaction is approximately two-fold with a medium FRET value of Cy3-d$T_{50}$-Cy5 bound to PALB2-DBD equal to that of unbound DNA with Cy3 and Cy5 25 nucleotides apart.

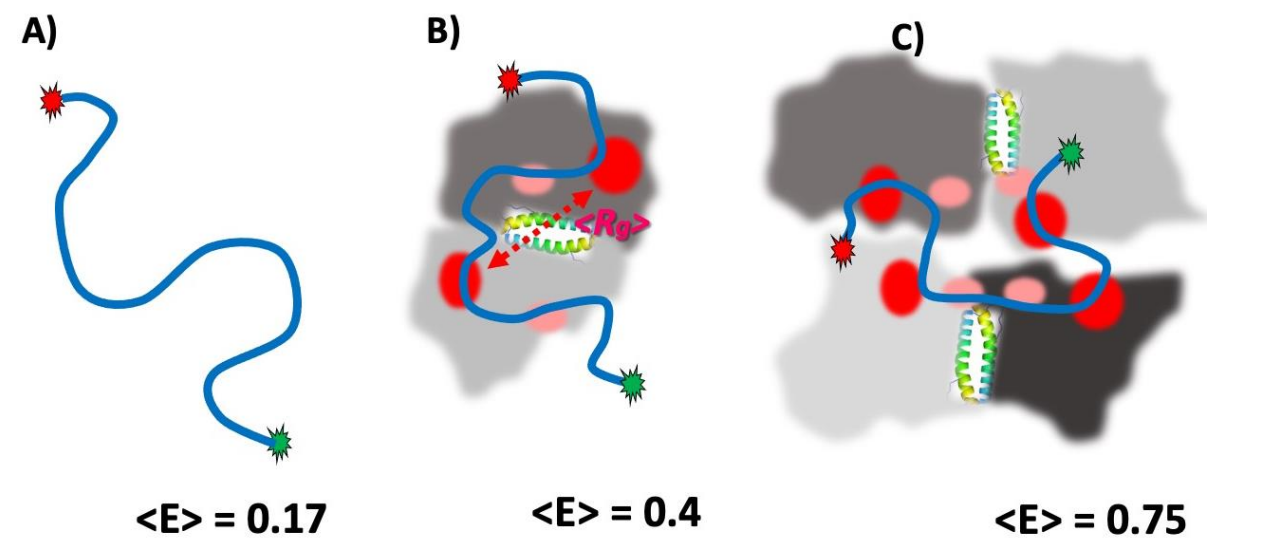


***Figure 7. Model of ssDNA compaction by PALB2-DBD dimer and tetramer**. **A)** Cy3-Cy5-ssDNA (blue line with red and green marks for position of fluorophores) in solution; **B)** ssDNA bound to PALB2-DBD dimer. Two monomers are shown by dark and light grey, with major DNA-binding sites shown in red and minor sites in pink; **C)** ssDNA bound to PALB2-DBD tetramer. Each dimer is formed around coiled-coil interface, while the interaction of two dimers is mediated by bound DNA.*

The mechanism behind the formation of high FRET peaks for each substrate is less clear. We hypothesize that this conformation can be formed by interaction with tetrameric PALB2-DBD. Similar attraction forces between unstructured Δ40-DBD, which stabilizes the dimer, can stimulate interactions between Δ40-DBDs of two dimers synergistically with DNA binding. Additional compaction of DNA bound to a larger complex can be achieved by condensation of DNA-binding sites around a ssDNA molecule. The physiological role of PALB2-DBD tetramerization remains unclear, considering a required stoichiometric excess of the protein suggested by our results, and the large size of the full-length protein as well as its functional complexes with even larger BRCA2 and BRCA1. However, the observed tetramerization further supports protein-protein interaction properties of the unstructured Δ40-DBD, which can contribute to forming heterooligomeric complexes with disordered parts of BRCA1 and BRCA2 as well as with other protein partners or to the formation of DNA-repair condensates through liquid-liquid phase separation.

*The mechanism of PALB2-DBD-mediated strand exchange.* Based on the DNA condensation results, the most attractive mechanistic hypothesis is that the PALB2 dimer uses a “chaperone-like” mechanism to destabilize a DNA helix by imposing a similar compaction force on dsDNA as observed for ssDNA. After dissociation from the protein, two DNA chains can

reanneal either with each other or with a homologous ssDNA bound to the same protein complex. We previously demonstrated that the PALB2-DBD does not unwind DNA in ensemble solution experiments and, in contrast, stimulates the annealing of complementary strands (Deveryshetty et al., 2019). Therefore, distortion of the double helix should be transient while bound to the protein, and concurrent with binding of an additional ssDNA. Disordered structural organization of the protein may be advantageous for transient interactions. Such a mechanism could explain a slow equilibrium process undergoing multiple association/dissociation cycles and lack of directionality, in contrast to RecA-like recombinases or Rad52 (Deveryshetty et al., 2019; Mazina et al., 2017).

A few cases of strand exchange activity were reported for RNA chaperones (Bhaskaran and Russell, 2007; Todd et al., 1996). *Escherichia coli* StpA protein promotes both RNA strand annealing and strand displacement (Rajkowitsch and Schroeder, 2007), similar to those reported for PALB2-DBD (Deveryshetty et al., 2019). SptA is a small globular protein that forms dimers required for these activities. Other examples include structured globular FANCA protein of the Fanconi anemia (FA) complex (Benitez et al., 2018) and limited reports of D-loop formation by FET proteins critical for RNA metabolism, specifically, pro-oncoprotein TLS/FUS and the human splicing factor PSF/hPOMp100 (Akhmedov and Lopez, 2000; Baechtold et al., 1999; Schwartz et al., 2015). The latter proteins include structured domains (RNA recognition motif (RRM) and zinc-finger) and sequences of low complexity and prion-like RGG domains implicated in higher-order self-assemblies. Neither the mechanism nor physiological relevance of DNA binding and D-loop formation by these proteins are known.

There are examples of disordered RNA chaperones composed of positively charged peptides that promote formation of compact nucleic acid conformations by acting as polypeptide counterions to overcome repulsive electrostatic interactions of RNA or DNA chains (Holmstrom et al., 2019; Koculi et al., 2006; Vo et al., 2009).

To our knowledge, PALB2-DBD represents the first known example of a compact intrinsically disordered structure with a balanced content of hydrophobic and charged residues of both polarities that support nucleic acid strand exchange or strand displacement. Numerous diverse RNA chaperones ranging from ATP-dependent helicases to small IDPs play a critical role in RNA metabolism due to the complexity, diversity and the dynamic nature of RNA structures and their complexes. PALB2-DBD-like domains may represent a novel class of RNA chaperones. Chaperone activities are not implicated in DNA metabolism due to significantly more restricted conformational space. However, multiple DNA transactions, such as repair of replication forks collided with transcription complexes and replication fork reversal require formation of transient

multichain DNA or RNA-DNA structures and other higher-order nucleic acid interactions(Michelini et al., 2017; Storici et al., 2007; Wei et al., 2012) that may benefit from strand exchange activity of the PALB2-DBD or similar unstructured domains of DNA repair proteins. The fact that PALB2 resides at the promoter regions of highly transcribed genes under nondamaging conditions (Bleuyard et al., 2017) further supports this hypothetical function.

DNA-binding proteins are enriched in disordered domains and have an increased disordered content in eukaryotes (Wang et al., 2016). Many such IDRs serve as scaffold or linker regions that coordinate functions of large multidomain proteins, e.g. RPA (Pokhrel et al., 2019; Yates et al., 2018) and large complexes. However, some IDRs coincide with DNA-binding domains. They may increase chromatin affinity by transient DNA binding without forming a strong roadblock on DNA, except cases of structure-specific interactions like nucleolin-G-quadruplex recognition (Masuzawa and Oyoshi, 2020). The “DNA chaperone” mechanism of PALB2-DBD suggests that DNA-binding IDRs may evolved to perform more complex reactions using an “inexpensive”, from the evolutionary point of view, functional unit that facilitates formation and resolution of transient multichain DNA/RNA intermediates during chromatin repair. Indeed, while this reaction is more complex than other reported for IDRs so far, it depends on rather simple structural requirements and similar IDRs can be incorporated into larger proteins, like scaffold BRCA1 and BRCA2. Such domains can function synergistically with RAD51 (Buisson et al., 2010; Deveryshetty et al., 2019; Zhao et al., 2017), or they can perform a required reaction “on the fly” without recruitment of additional specialized factors, like RAD52, and without formation of stable presynaptic RAD51 filaments.

## Materials and Methods

*Cloning and Purification.* PALB2 N-terminal fragments 1-195 I195W and 1-295 K295W were cloned into a pSMT-MBP plasmid containing an N-terminal 6xHis-SUMO and a C-terminal MBP tag using Gibson Assembly. The pSMT-MBP plasmid was created from a pET28b+ based pSMT3 plasmid given to us by Dr. R. A. Kovall, University of Cincinnati, by insertion of a TEV cleavage site and MBP at the C-terminus of the ORF. Mutations were introduced by a Stratagene QuikChange™ protocol. All constructs were fully sequenced. Constructs were transformed into BL21* cells for expression. The Cell Cultures were grown in TB to $OD_{600}$ = 1.6, protein expression was induced by adding 0.2 mM IPTG and carried out at 16°C overnight. Cell suspension was centrifuged at 4,000 rpm for 15 min. Cell pellets were resuspended in a lysis buffer (25 mM HEPES pH 8.0, 1 M NaCl, 10% glycerol, 2 mM CHAPS, 1 mM TCEP, 0.3% Brij35, 2 mM EDTA, 1 mM PMSF), frozen in liquid nitrogen and stored at -80°C. Frozen cell suspension was thawed

and lysed with lysozyme at 0.25 mg/ml at RT for 20 min, followed by three rounds of sonication (50% output and 50% pulsar setting for 4 min). Cell debris was removed by centrifugation at 30,600 g for 45 min. Supernatant was loaded onto an Amylose resin column (New England Biolabs®) equilibrated with the binding buffer (25 mM HEPES pH 8.0, 1 M NaCl, 10% glycerol, 2 mM CHAPS, 1 mM TCEP). Resin was washed with binding buffer with added 0.05% NP40 and 1 mM EDTA, followed by a second wash with low salt buffer (50 mM NaCl, 25mM HEPES pH 8.0, 1 mM TCEP, 2 mM CHAPS, 0.2% NP40, 1 mM EDTA). Protein was eluted in a binding buffer containing 20 mM Maltose. The SUMO and MBP tags were cleaved using Ulp1 and TEV proteases respectively. Protein was diluted five-fold with binding buffer without NaCl to bring the NaCl concentration to 200 mM, loaded onto a Hi-Trap heparin affinity column (2x5 ml, GE Life Sciences) and eluted with a gradient of NaCl (200 mM – 800 mM). Protein eluted from the column at ~600 mM NaCl concentration. Gel Filtration was performed using a Superdex-200 10/300GL column (GE Life Sciences).

*Circular Dichroism (CD).* The protein buffer was exchanged using a Zeba Spin desalting column to a buffer containing 10 mM $NaPO_4$, 150 mM NaF, 2 mM CHAPS. Samples were centrifuged at 16,000 g to remove aggregates and the final concentration was verified. Measurements were performed using a Jasco J-715. The following samples were used for measurements: 1) 30 µM PALB2-DBD, 2) 30 µM ∆40-DBD, and 3), 4) same as 1), 2) with addition of 30 µM $dT_{50}$. Each spectrum is an average of 10 scans.

*Electron Paramagnetic Resonance (EPR).* Monocysteine PALB2 mutants were labelled with MTSL (S-(1-oxyl-2,2,5,5-tetramethyl-2,5-dihydro-1H-pyrrol-3-yl)methyl methanesulfonothioate) in a protein sample free of reducing agents. The unbound label was separated on a Superdex-200 10/300 gel filtration column. Spectra were measured using 20 µM protein in 20 mM Tris-Acetate pH 7, 100 mM NaCl, 10% DMSO, and 5% glycerol with or without 80 µM $dT_{40}$. CW EPR spectroscopy was performed on an EMX spectrometer operating at X-band frequency (9.77 GHz) at room temperature. A dielectric resonator (ER4123D) and 25 µl borosilicate glass capillaries were used. The center field was set to 3474 gauss (G), while the sweep width was 150 G. The modulation amplitude, microwave power, and resolution were set to 1.6 G, 2 mW, and 1125 points, respectively.

*SEC-MALS-SAXS.* SAXS was performed at BioCAT (beamline 18ID at the Advanced Photon Source, Chicago) with in-line size exclusion chromatography (SEC) to separate sample from aggregates and other contaminants, thus ensuring optimal sample quality. Multiangle light scattering (MALS), dynamic light scattering (DLS) and refractive index measurement (RI)) were used for additional biophysical characterization (SEC-MALS-SAXS). The samples were loaded

on a Superdex 200 Increase 10/300 GL column (Cytiva) run by a 1260 Infinity II HPLC (Agilent Technologies) at 0.6 ml/min. The flow passed through (in order) the Agilent UV detector, a MALS detector and a DLS detector (DAWN Helios II, Wyatt Technologies), and an RI detector (Optilab T-rEX, Wyatt). The flow then went through the SAXS flow cell. The flow cell consists of a 1.0 mm ID quartz capillary with ~20 µm walls. A coflowing buffer sheath is used to separate samples from the capillary walls, helping prevent radiation damage (Kirby et al., 2016). X-ray data was collected using a 150 (h) x 25 (v) (µm) 12 keV beam, and scattering intensity was recorded using an Eiger2 XE 9M (Dectris) photon-counting detector. The detector was placed 3.6 m from the sample giving access to a *q*-range of 0.0029 $Å^{-1}$ to 0.417 $Å^{-1}$, with the momentum transfer vector *q* defined as $q = (4\pi/\lambda)\sin(\theta)$. 0.5 s exposures were acquired every 1 s during elution and data was reduced using BioXTAS RAW 2.1.1 (Hopkins et al., 2017). Buffer blanks were created by averaging regions flanking the elution peak and subtracted from exposures selected from the elution peak to create the I(q) vs q curves used for subsequent analyses. Molecular weights and hydrodynamic radii were calculated from the MALS and DLS data, respectively, using the ASTRA 7 software (Wyatt).

*Confocal smFRET.* All samples were prepared in buffer with 0.16 M NaCl, 20 mM HEPES pH 8.9, 1 mM TCEP, 0.003% Tween20 and centrifuged at 40,000 rpm for 20 min before experiment. FRET measurements of freely diffusing single molecules were performed with a confocal microscope MicroTime 200 (PicoQuant) as published(Pozzi et al., 2016; Ruben et al., 2020). The donor and acceptor were excited by laser light from 532 nm and 638 nm lasers, respectively. A pulsed interleaved excitation (PIE) setup was used with a pulse rate of 40 MHz to alternate the donor and acceptor excitation. PIE reports the status of both donor and acceptor fluorophores by sorting molecules based on relative donor:acceptor stoichiometry (S) and apparent FRET efficiency (E), as described before(Ferreon et al., 2010; Kapanidis et al., 2004; Kudryavtsev et al., 2012). Measurements were performed 25 µm deep in the solution using a laser power of ~15 µW for 30 to 60 min per sample. Data recording was performed using the SymPhoTime Software 64, version 2.4 (PicoQuant). Data analysis was carried out with the Matlab-based software PAM (https://pam.readthedocs.io/en/latest/)(Schrimpf et al., 2018) using a customized profile optimized for our microscope. Signals from single molecules were observed as bursts of fluorescence. Bursts with more than 50 counts were searched with the Dual Channel Burst Search (DCBS) algorithm. Integration time was set to 0.5 ms.

Appropriate corrections for direct excitation of the acceptor at the donor excitation wavelength (DE), leakage of the donor in the acceptor channel (Lk), and the instrumental factor (g) were determined experimentally using a mixture of dsDNA models with known FRET efficiency

and stoichiometry labeled with dyes Cy3 and Cy5. These are the following: DE=0.05, Lk=0.08, g = 0.85.

*Analytical Ultracentrifugation (AUC).* PALB2-DBD samples were dialyzed against 1M NaCl, 10% glycerol, 1 mM TCEP, 2 mM CHAPS, 25 mM HEPES pH 7.5 consecutively three times for at least 2 hours each. For complexes with DNA, protein was first dialyzed against 500 mM NaCl, 10% glycerol, 1 mM TCEP, 2 mM CHAPS, 25 mM HEPES pH 7.5 for at least 6 hours and then mixed with the corresponding DNA substrate while concomitantly reducing NaCl concentration by dilution with buffer without NaCl. Alternatively, DNA was mixed with protein at 0.5M NaCl buffer and samples were dialyzed against several buffers with a stepwise decreasing concentration of NaCl (350 mM, 250 mM, 100 mM NaCl or 400 mM, 300 mM, 200 mM, 100 mM and 10% glycerol, 1 mM TCEP, 2 mM CHAPS, 25 mM HEPES pH 7.5).

Experiments were performed on Beckman Coulter Optima AUC. Samples were loaded into cells and centrifuged at 30,000 rpm for 20 hours at 20°C. Absorbance was measured at 280 nm for protein samples and at 550 nm for samples with Cy3-labeled DNA every 4 min for a total of 250 scans. Buffer density and viscosity were calculated by Sednterp: 1.069 ρ and 0.01486 η for 1M NaCl buffer, 0.01436 η and 1.04913 ρ for 0.5M NaCl buffer and 1.033 ρ and 0.014 η for DNA complexes at 0.1M NaCl. These values were used to fit the data using the Lamm equation for the continuous c(s) distribution model in Sedfit (Schuck, 2000). Two N-terminal fragments were used, 195W and 295W, which consisted of amino acids 1-195 and 1-295 respectively. The last amino acid was mutated to tryptophan to increase the absorption at 280 nm. DNA-only samples were analyzed using a spatial specific volume of 0.55. The sequence for the 25mer substrate is 5Cy3/TGG CGA CGG CAG CGA GGC TCT CTA C; all other substrates are poly(dT).

**Acknowledgements**. We are grateful to Drs. Gregory DeKostner and Carl Frieden for help with CD spectroscopy, to Dr. Sahiti Kupa for help with the preliminary CD spectroscopy experiments, to Dr. Samantha Gies for help with EPR spectroscopy, to Ian Miller for help with mutagenesis, to Laasyapriya Sarva for help with protein purification and activity validation, to Andrea Soranno for helpful discussions, and to Dr. Joel Eissenberg for critical reading and help with the manuscript preparation.

**Funding.** The research was supported by Saint Louis University Institute for Drug and Biotherapeutic Innovation and E.A. Doisy Department of Biochemistry and Molecular Biology; by Washington University Siteman Cancer Center (SCC), the Foundation for Barnes-Jewish Hospital Siteman Investment Program (SIP) and Institute of Clinical and Translational Sciences JIT Core

Usage Funding program; by NIH S10 OD030343. SAXS experiments used resources of the Advanced Photon Source, a U.S. Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Argonne National Laboratory under Contract No. DE-AC02-06CH11357. BioCAT was supported by grant P30 GM138395 from the National Institute of General Medical Sciences of the National Institutes of Health. The content is solely the responsibility of the authors and does not necessarily reflect the official views of the National Institute of General Medical Sciences or the National Institutes of Health.

**Author contributions**

S.K. conceptualized the study; Y. K., M.B.W, J.M.R., R.D., J.H., N. P., S.K. designed, performed and analyzed experiments; M.B.W performed computational modeling of SAXS data; R.D. designed and analyzed cwEPR; N.P. helped to design and analysis of confocal smFRET, S.K. wrote the manuscript with input from all authors, supervised the study and acquired funding.

**Conflict of interest.** None declared.

**Data availability**.

All constructs are freely available upon request. SAXS scattering data are deposited to SASDBD (https://www.sasbdb.org/) with accession ID's SAS5082 and SAS5083 for high salt and low salt data sets, respectively. smFRET data in PTU file format (time-correlated single-photon counting data) are available upon request.

***References***

Akhmedov, A.T., and Lopez, B.S. (2000). Human 100-kDa homologous DNA-pairing protein is the splicing factor PSF and promotes DNA strand invasion. Nucleic Acids Res *28*, 3022-3030.

Baechtold, H., Kuroda, M., Sok, J., Ron, D., Lopez, B.S., and Akhmedov, A.T. (1999). Human 75-kDa DNA-pairing protein is identical to the pro-oncoprotein TLS/FUS and is able to promote D-loop formation. J Biol Chem *274*, 34337-34342.

Beernink, H.T., and Morrical, S.W. (1999). RMPs: recombination/replication mediator proteins. Trends Biochem Sci *24*, 385-389.

Benitez, A., Liu, W., Palovcak, A., Wang, G., Moon, J., An, K., Kim, A., Zheng, K., Zhang, Y., Bai, F.*, et al.* (2018). FANCA Promotes DNA Double-Strand Break Repair by Catalyzing Single-Strand Annealing and Strand Exchange. Mol Cell *71*, 621-628 e624.

Bernado, P., Mylonas, E., Petoukhov, M.V., Blackledge, M., and Svergun, D.I. (2007). Structural characterization of flexible proteins using small-angle X-ray scattering. J Am Chem Soc *129*, 5656-5664.

Bernado, P., and Svergun, D.I. (2012). Analysis of intrinsically disordered proteins by small-angle X-ray scattering. Methods Mol Biol *896*, 107-122.

Bhaskaran, H., and Russell, R. (2007). Kinetic redistribution of native and misfolded RNAs by a DEAD-box chaperone. Nature *449*, 1014-1018.

Bi, B., Rybalchenko, N., Golub, E.I., and Radding, C.M. (2004). Human and yeast Rad52 proteins promote DNA strand exchange. Proc Natl Acad Sci U S A *101*, 9568-9572.

Bleuyard, J.Y., Fournier, M., Nakato, R., Couturier, A.M., Katou, Y., Ralf, C., Hester, S.S., Dominguez, D., Rhodes, D., Humphrey, T.C.*, et al.* (2017). MRG15-mediated tethering of PALB2 to unperturbed chromatin protects active genes from genotoxic stress. Proc Natl Acad Sci U S A *114*, 7671-7676.

Buisson, R., Dion-Cote, A.M., Coulombe, Y., Launay, H., Cai, H., Stasiak, A.Z., Stasiak, A., Xia, B., and Masson, J.Y. (2010). Cooperation of breast cancer proteins PALB2 and piccolo BRCA2 in stimulating homologous recombination. Nat Struct Mol Biol *17*, 1247-1254.

Buisson, R., and Masson, J.Y. (2012). PALB2 self-interaction controls homologous recombination. Nucleic Acids Res *40*, 10312-10323.

Chen, Z., Yang, H., and Pavletich, N.P. (2008). Mechanism of homologous recombination from the RecA-ssDNA/dsDNA structures. Nature *453*, 489-484.

Chinnaraj, M., Flaumenhaft, R., and Pozzi, N. (2022). Reduction of protein disulfide isomerase results in open conformations and stimulates dynamic exchange between structural ensembles. J Biol Chem *298*, 102217.

Clark, P.L., Plaxco, K.W., and Sosnick, T.R. (2020). Water as a Good Solvent for Unfolded Proteins: Folding and Collapse are Fundamentally Different. J Mol Biol *432*, 2882-2889.

Cox, M.M. (1999). Recombinational DNA repair in bacteria and the RecA protein. Prog Nucleic Acid Res Mol Biol *63*, 311-366.

Cox, M.M. (2001). Historical overview: searching for replication help in all of the rec places. Proc Natl Acad Sci U S A *98*, 8173-8180.

Cox, M.M. (2003). The bacterial RecA protein as a motor protein. Annu Rev Microbiol *57*, 551-577.

Cox, M.M. (2007). Regulation of bacterial RecA protein function. Crit Rev Biochem Mol Biol *42*, 41-63.

Curtis, J.E., Raghunandan, S., Nanda, H., and Krueger, S. (2012). SASSIE: A program to study intrinsically disordered biological molecules and macromolecular ensembles using experimental scattering restraints. Comput Phys Commun *183*, 382-389.

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.

Deveryshetty, J., Peterlini, T., Ryzhikov, M., Brahiti, N., Dellaire, G., Masson, J.Y., and Korolev, S. (2019). Novel RNA and DNA strand exchange activity of the PALB2 DNA binding domain and its critical role for DNA repair in cells. Elife *8*:e44063

Dray, E., Etchin, J., Wiese, C., Saro, D., Williams, G.J., Hammel, M., Yu, X., Galkin, V.E., Liu, D., Tsai, M.S.*, et al.* (2010). Enhancement of RAD51 recombinase activity by the tumor suppressor PALB2. Nat Struct Mol Biol *17*, 1255-1259.

Drescher, M. (2012). EPR in protein science : intrinsically disordered proteins. Top Curr Chem *321*, 91-119.

Ducy, M., Sesma-Sanz, L., Guitton-Sert, L., Lashgari, A., Gao, Y., Brahiti, N., Rodrigue, A., Margaillan, G., Caron, M.C., Cote, J.*, et al.* (2019). The Tumor Suppressor PALB2: Inside Out. Trends Biochem Sci *44*, 226-240.

Ferreon, A.C., Moran, C.R., Gambin, Y., and Deniz, A.A. (2010). Single-molecule fluorescence studies of intrinsically disordered proteins. Methods Enzymol *472*, 179-204.

Grimme, J.M., Honda, M., Wright, R., Okuno, Y., Rothenberg, E., Mazin, A.V., Ha, T., and Spies, M. (2010). Human Rad52 binds and wraps single-stranded DNA and mediates annealing via two hRad52-ssDNA complexes. Nucleic Acids Res *38*, 2917-2930.

Habchi, J., Tompa, P., Longhi, S., and Uversky, V.N. (2014). Introducing protein intrinsic disorder. Chem Rev *114*, 6561-6588.

Hall, S.D., and Kolodner, R.D. (1994). Homologous pairing and strand exchange promoted by the Escherichia coli RecT protein. Proc Natl Acad Sci U S A *91*, 3205-3209.

Holehouse, A.S., Das, R.K., Ahad, J.N., Richardson, M.O., and Pappu, R.V. (2017). CIDER: Resources to Analyze Sequence-Ensemble Relationships of Intrinsically Disordered Proteins. Biophys J *112*, 16-21.

Holmstrom, E.D., Liu, Z., Nettels, D., Best, R.B., and Schuler, B. (2019). Disordered RNA chaperones can enhance nucleic acid folding via local charge screening. Nat Commun *10*, 2453.

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Zidek, A., Potapenko, A.*, et al.* (2021). Highly accurate protein structure prediction with AlphaFold. Nature *596*, 583-589.

Kagawa, W., Kurumizaka, H., Ishitani, R., Fukai, S., Nureki, O., Shibata, T., and Yokoyama, S. (2002). Crystal structure of the homologous-pairing domain from the human Rad52 recombinase in the undecameric form. Mol Cell *10*, 359-371.

Kapanidis, A.N., Lee, N.K., Laurence, T.A., Doose, S., Margeat, E., and Weiss, S. (2004). Fluorescence-aided molecule sorting: analysis of structure and interactions by alternating-laser excitation of single molecules. Proc Natl Acad Sci U S A *101*, 8936-8941.

Koculi, E., Thirumalai, D., and Woodson, S.A. (2006). Counterion charge density determines the position and plasticity of RNA folding transition states. J Mol Biol *359*, 446-454.

Kowalczykowski, S.C. (2015). An Overview of the Molecular Mechanisms of Recombinational DNA Repair. Cold Spring Harb Perspect Biol *7*.

Krogh, B.O., and Symington, L.S. (2004). Recombination proteins in yeast. Annu Rev Genet *38*, 233-271.

Kudryavtsev, V., Sikor, M., Kalinin, S., Mokranjac, D., Seidel, C.A., and Lamb, D.C. (2012). Combining MFD and PIE for accurate single-pair Forster resonance energy transfer measurements. Chemphyschem *13*, 1060-1078.

Lee, M., Lee, C.H., Demin, A.A., Munashingha, P.R., Amangyeld, T., Kwon, B., Formosa, T., and Seo, Y.S. (2014). Rad52/Rad59-dependent recombination as a means to rectify faulty Okazaki fragment processing. J Biol Chem *289*, 15064-15079.

Liu, J., and Heyer, W.D. (2011). Who's who in human recombination: BRCA2 and RAD52. Proc Natl Acad Sci U S A *108*, 441-442.

Longhi, S., Belle, V., Fournel, A., Guigliarelli, B., and Carriere, F. (2011). Probing structural transitions in both structured and disordered proteins using site-directed spin-labeling EPR spectroscopy. J Pept Sci *17*, 315-328.

Masuzawa, T., and Oyoshi, T. (2020). Roles of the RGG Domain and RNA Recognition Motif of Nucleolin in G-Quadruplex Stabilization. ACS Omega *5*, 5202-5208.

Mazina, O.M., Keskin, H., Hanamshet, K., Storici, F., and Mazin, A.V. (2017). Rad52 Inverse Strand Exchange Drives RNA-Templated DNA Double-Strand Break Repair. Mol Cell *67*, 19-29 e13.

Michelini, F., Pitchiaya, S., Vitelli, V., Sharma, S., Gioia, U., Pessina, F., Cabrini, M., Wang, Y., Capozzo, I., Iannelli, F.*, et al.* (2017). Damage-induced lncRNAs control the DNA damage response through interaction with DDRNAs at individual double-strand breaks. Nat Cell Biol *19*, 1400-1411.

Mirdita, M., Schutze, K., Moriwaki, Y., Heo, L., Ovchinnikov, S., and Steinegger, M. (2022). ColabFold: making protein folding accessible to all. Nat Methods *19*, 679-682.

Morrical, S.W. (2015). DNA-pairing and annealing processes in homologous recombination and homology-directed repair. Cold Spring Harb Perspect Biol *7*, a016444.

Mortensen, U.H., Lisby, M., and Rothstein, R. (2009). Rad52. Curr Biol *19*, R676-677.

Moynahan, M.E., Pierce, A.J., and Jasin, M. (2001). BRCA2 is required for homology-directed repair of chromosomal breaks. Mol Cell *7*, 263-272.

Murphy, K.C. (2012). Phage recombinases and their applications. Adv Virus Res *83*, 367-414.

Murphy, M.C., Rasnik, I., Cheng, W., Lohman, T.M., and Ha, T. (2004). Probing single-stranded DNA conformational flexibility using fluorescence spectroscopy. Biophys J *86*, 2530-2537.

Ogawa, T., Yu, X., Shinohara, A., and Egelman, E.H. (1993). Similarity of the yeast RAD51 filament to the bacterial RecA filament. Science *259*, 1896-1899.

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.

Passy, S.I., Yu, X., Li, Z., Radding, C.M., and Egelman, E.H. (1999). Rings and filaments of beta protein from bacteriophage lambda suggest a superfamily of recombination proteins. Proc Natl Acad Sci U S A *96*, 4279-4284.

Piovesan, D., Del Conte, A., Clementel, D., Monzon, A.M., Bevilacqua, M., Aspromonte, M.C., Iserte, J.A., Orti, F.E., Marino-Buslje, C., and Tosatto, S.C.E. (2023). MobiDB: 10 years of intrinsically disordered proteins. Nucleic Acids Res *51*, D438-D444.

Pokhrel, N., Caldwell, C.C., Corless, E.I., Tillison, E.A., Tibbs, J., Jocic, N., Tabei, S.M.A., Wold, M.S., Spies, M., and Antony, E. (2019). Dynamics and selective remodeling of the DNA-binding domains of RPA. Nat Struct Mol Biol *26*, 129-136.

Pozzi, N., Bystranowska, D., Zuo, X., and Di Cera, E. (2016). Structural Architecture of Prothrombin in Solution Revealed by Single Molecule Spectroscopy. J Biol Chem *291*, 18107-18116.

Rajkowitsch, L., and Schroeder, R. (2007). Dissecting RNA chaperone activity. RNA *13*, 2053-2060.

Reuven, N.B., Staire, A.E., Myers, R.S., and Weller, S.K. (2003). The herpes simplex virus type 1 alkaline nuclease and single-stranded DNA binding protein mediate strand exchange in vitro. J Virol *77*, 7425-7433.

Riback, J.A., Bowman, M.A., Zmyslowski, A.M., Knoverek, C.R., Jumper, J.M., Hinshaw, J.R., Kaye, E.B., Freed, K.F., Clark, P.L., and Sosnick, T.R. (2017). Innovative scattering analysis shows that hydrophobic disordered proteins are expanded in water. Science *358*, 238-241.

Roca, A.I., and Cox, M.M. (1990). The RecA protein: structure and function. Crit Rev Biochem Mol Biol *25*, 415-456.

Rothenberg, E., Grimme, J.M., Spies, M., and Ha, T. (2008). Human Rad52-mediated homology search and annealing occurs by continuous interactions between overlapping nucleoprotein complexes. Proc Natl Acad Sci U S A *105*, 20274-20279.

Ruben, E., Planer, W., Chinnaraj, M., Chen, Z., Zuo, X., Pengo, V., De Filippis, V., Alluri, R.K., McCrae, K.R., Macor, P.*, et al.* (2020). The J-elongated conformation of beta2-glycoprotein I predominates in solution: implications for our understanding of antiphospholipid syndrome. J Biol Chem *295*, 10794-10806.

San Filippo, J., Sung, P., and Klein, H. (2008). Mechanism of eukaryotic homologous recombination. Annu Rev Biochem *77*, 229-257.

Saotome, M., Saito, K., Yasuda, T., Ohtomo, H., Sugiyama, S., Nishimura, Y., Kurumizaka, H., and Kagawa, W. (2018). Structural Basis of Homology-Directed DNA Repair Mediated by RAD52. iScience *3*, 50-62.

Schneidman-Duhovny, D., Hammel, M., and Sali, A. (2010). FoXS: a web server for rapid computation and fitting of SAXS profiles. Nucleic Acids Res *38*, W540-544.

Schrimpf, W., Barth, A., Hendrix, J., and Lamb, D.C. (2018). PAM: A Framework for Integrated Analysis of Imaging, Single-Molecule, and Ensemble Fluorescence Data. Biophys J *114*, 1518-1528.

Schuck, P. (2000). Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling. Biophys J *78*, 1606-1619.

Schwartz, J.C., Cech, T.R., and Parker, R.R. (2015). Biochemical Properties and Biological Functions of FET Proteins. Annu Rev Biochem *84*, 355-379.

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.

Singleton, M.R., Wentzell, L.M., Liu, Y., West, S.C., and Wigley, D.B. (2002). Structure of the single-strand annealing domain of human RAD52 protein. Proc Natl Acad Sci U S A *99*, 13492-13497.

Song, F., Li, M., Liu, G., Swapna, G.V.T., Daigham, N.S., Xia, B., Montelione, G.T., and Bunting, S.F. (2018). Antiparallel Coiled-Coil Interactions Mediate the Homodimerization of the DNA Damage-Repair Protein PALB2. Biochemistry *57*, 6581-6591.

Storici, F., Bebenek, K., Kunkel, T.A., Gordenin, D.A., and Resnick, M.A. (2007). RNA-templated DNA repair. Nature *447*, 338-341.

Sung, P. (1997). Function of yeast Rad52 protein as a mediator between replication protein A and the Rad51 recombinase. J Biol Chem *272*, 28194-28197.

Sung, P., and Klein, H. (2006). Mechanism of homologous recombination: mediators and helicases take on regulatory functions. Nat Rev Mol Cell Biol *7*, 739-750.

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

Symington, L.S. (2002). Role of RAD52 epistasis group genes in homologous recombination and double-strand break repair. Microbiol Mol Biol Rev *66*, 630-670.

Todd, M.J., Lorimer, G.H., and Thirumalai, D. (1996). Chaperonin-facilitated protein folding: optimization of rate and yield by an iterative annealing mechanism. Proc Natl Acad Sci U S A *93*, 4030-4035.

Tria, G., Mertens, H.D., Kachala, M., and Svergun, D.I. (2015). Advanced ensemble modelling of flexible macromolecules using X-ray solution scattering. IUCrJ *2*, 207-217.

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

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 der Lee, R., Buljan, M., Lang, B., Weatheritt, R.J., Daughdrill, G.W., Dunker, A.K., Fuxreiter, M., Gough, J., Gsponer, J., Jones, D.T.*, et al.* (2014). Classification of intrinsically disordered regions and proteins. Chem Rev *114*, 6589-6631.

Venkitaraman, A.R. (2000). The breast cancer susceptibility gene, BRCA2: at the crossroads between DNA replication and recombination? Philos Trans R Soc Lond B Biol Sci *355*, 191-198.

Venkitaraman, A.R. (2014). Cancer suppression by the chromosome custodians, BRCA1 and BRCA2. Science *343*, 1470-1475.

Vo, M.N., Barany, G., Rouzina, I., and Musier-Forsyth, K. (2009). HIV-1 nucleocapsid protein switches the pathway of transactivation response element RNA/DNA annealing from loop-loop "kissing" to "zipper". J Mol Biol *386*, 789-801.

Wang, C., Uversky, V.N., and Kurgan, L. (2016). Disordered nucleiome: Abundance of intrinsic disorder in the DNA- and RNA-binding proteins in 1121 species from Eukaryota, Bacteria and Archaea. Proteomics *16*, 1486-1498.

Wei, W., Ba, Z., Gao, M., Wu, Y., Ma, Y., Amiard, S., White, C.I., Rendtlew Danielsen, J.M., Yang, Y.G., and Qi, Y. (2012). A role for small RNAs in DNA double-strand break repair. Cell *149*, 101-112.

Weinkam, P., Pons, J., and Sali, A. (2012). Structure-based model of allostery predicts coupling between distant sites. Proc Natl Acad Sci U S A *109*, 4875-4880.

Xia, B., Sheng, Q., Nakanishi, K., Ohashi, A., Wu, J., Christ, N., Liu, X., Jasin, M., Couch, F.J., and Livingston, D.M. (2006). Control of BRCA2 cellular and clinical functions by a nuclear partner, PALB2. Mol Cell *22*, 719-729.

Yang, H., and Pavletich, N.P. (2021). Insights into homology search from cryo-EM structures of RecA-DNA recombination intermediates. Curr Opin Genet Dev *71*, 188-194.

Yang, H., Zhou, C., Dhar, A., and Pavletich, N.P. (2020). Mechanism of strand exchange from RecA-DNA synaptic and D-loop structures. Nature *586*, 801-806.

Yasuhara, T., Kato, R., Hagiwara, Y., Shiotani, B., Yamauchi, M., Nakada, S., Shibata, A., and Miyagawa, K. (2018). Human Rad52 Promotes XPG-Mediated R-loop Processing to Initiate Transcription-Associated Homologous Recombination Repair. Cell *175*, 558-570 e511.

Yates, L.A., Aramayo, R.J., Pokhrel, N., Caldwell, C.C., Kaplan, J.A., Perera, R.L., Spies, M., Antony, E., and Zhang, X. (2018). A structural and dynamic model for the assembly of Replication Protein A on single-stranded DNA. Nat Commun *9*, 5447.

Zhang, F., Fan, Q., Ren, K., and Andreassen, P.R. (2009a). PALB2 functionally connects the breast cancer susceptibility proteins BRCA1 and BRCA2. Mol Cancer Res *7*, 1110-1118.

Zhang, F., Ma, J., Wu, J., Ye, L., Cai, H., Xia, B., and Yu, X. (2009b). PALB2 links BRCA1 and BRCA2 in the DNA-damage response. Curr Biol *19*, 524-529.

Zhang, J., Xing, X., Herr, A.B., and Bell, C.E. (2009c). Crystal structure of E. coli RecE protein reveals a toroidal tetramer for processing double-stranded DNA breaks. Structure *17*, 690-702.

Zhao, W., Steinfeld, J.B., Liang, F., Chen, X., Maranon, D.G., Jian Ma, C., Kwon, Y., Rao, T., Wang, W., Sheng, C.*, et al.* (2017). BRCA1-BARD1 promotes RAD51-mediated homologous DNA pairing. Nature *550*, 360-365.