Functional Evaluation of the C-terminal Region of Bacteriophage T4 Rad50

Haley Streff, [1,2], *Yang Gao*[1,3], *and Scott W. Nelson*[1]*

[1]Department of Biochemistry, Biophysics, and Molecular Biology, Iowa State University, Ames, Iowa 50011

[2]Undergraduate Research Program at Iowa State University, Ames, Iowa 50011

[3]Current Address: Department of BioSciences, Rice University, Houston, Texas 77005

*To whom correspondence may be addressed. Email: swn@iastate.edu, Phone: 515-294-3434.

Abstract

Bacteriophage T4 encodes orthologs of the proteins Rad50 (gp46) and Mre11 (gp47), which form a heterotetrameric complex (MR) that participates in the processing of DNA ends for recombination-dependent DNA repair. Crystal and high-resolution cryo-EM structures of Rad50 have revealed DNA binding sites near the dimer interface of Rad50 opposite of Mre11, and near the base of the coiled-coils that extend out from the globular head domain. An analysis of T4-Rad50 using sequenced-based algorithms to identify DNA binding residues predicts that a conserved region of positively charged residues near the C-terminus, distal to the observed binding sites, interacts with DNA. Mutant proteins were generated to test this prediction and their enzymatic and DNA binding activities were evaluated. Consistent with the predictions, the Rad50 C-terminal mutants had reduced affinity for DNA as measured by Rad50 equilibrium DNA binding assays and an increased Km-DNA as determined in MR complex nuclease assays. Moreover, the allosteric activation of ATP hydrolysis activity due to DNA binding was substantially reduced, suggesting that these residues may be involved in the communication between the DNA and ATP binding sites.

Keywords: Protein-DNA binding residues, exonuclease, ATPase, MR complex, Bacteriophage T4,

Abbreviations: DSB, DNA double-strand break; ssDNA, single-stranded DNA; SMC, structural maintenance of chromosome; dsDNA, double-stranded DNA; 2AP, 2-aminopurine deoxyribonucleotide

Introduction

Double-strand breaks (DSBs) are among the most destructive forms of DNA damage and if not properly repaired they can result in chromosomal rearrangements that can lead to cellular dysfunction and death [1,2]. DSBs are caused by internal and external DNA damaging agents and are primarily repaired through two pathways, non-homologous end joining and homologous recombination [3]. In most bacteria and bacteriophages, homologous recombination dominates and is considered to be an error-free process [4].

The first step of DSB repair is to process the end of the DNA into a 3' single-stranded DNA (ssDNA) extension [5]. In prokaryotes, this is accomplished by the helicase/nuclease RecBCD complex [6]. However, bacteriophage T4 does not encode a RecBCD homolog, but instead uses proteins that are orthologs of the Rad50 (gp46) and Mre11 (gp47) proteins that are involved in eukaryotic and archaeal DSB repair [7]. The role of the MR complex in DSB repair is both architectural and enzymatic [8,9]. The MR complex binds double-stranded DNA (dsDNA) and has the ability to tether DNA strands together, presumably through the coiled-coil and zinc-hook domains of Rad50 [8,10–12]. In T4 phage, this tethering appears to be responsible for the coordination of the DSBs ends in the recombination reaction [13]. *In vitro*, the dominant enzymatic activity of the MR complex is a 3' to 5' exonuclease, which is

seemingly incompatible with the production of 3' ssDNA overhangs [14,15]. The nuclease activity *in vivo* has been the subject of intense investigation [16]. The presence of T4 phage recombination proteins UvsY and gp32 enhance an endonuclease activity in Mre11, which when directed to the correct strand produces 3' ssDNA [15]. Alternatively, in the eukaryotic system, Mre11 performs an endonuclease cut on the 5' terminating strand and the exonuclease activity of Mre11 proceeds to remove nucleotides in a 3' to 5' direction towards the DSB [17,18].

Rad50 is a member of the structural maintenance of chromosome (SMC) family, which is part of the ABC protein superfamily [19,20]. The ABC superfamily contains several highly conserved motifs, including Walker A and B motifs, a Signature motif, and H-, D-, and Q-loops [21]. Unique to Rad50 is a zinc-hook motif that lies at the apex of an intramolecular coiled-coiled domain [22,23]. In the absence of dsDNA, the T4 MR complex slowly hydrolyzes ATP (k_{cat} = 0.15 sec^{-1}), but upon binding to the end of dsDNA, the ATP hydrolysis rate of the MR complex increases 30-fold, along with an increase in cooperativity between ATP active sites (Hill coefficient of 2.4) [15,24]. This increase occurs even in the absence of nuclease activity, as nuclease deficient mutants or omission of Mn^{2+} still activates ATPase activity [14,15]. We have previously shown that Rad50 alone has a high affinity for both single-stranded (ssDNA) and dsDNA, whereas the MR complex has reduced affinity unless ATP is present [11,15].

Until recently, the location of the DNA binding site(s) on Rad50 has been uncertain. Following the first x-ray crystal structure of *Pfu*-Rad50, it was suggested that dsDNA may bind to a positively charged groove that is formed at the dimer interface of Rad50 [25]. Mutation of residues within this groove and an adjacent β-sheet were shown to decrease DNA affinity. The first co-crystal structure of Rad50 bound to DNA revealed an unexpected DNA binding site made up of the base of the coiled-coil domain and a nearby portion of the nucleotide binding

domain [26]. Mutagenesis of key residues in these two regions reduced DNA binding measured via EMSA, confirming the importance of the DNA-protein interactions [26]. Subsequent co-crystal structures from both the Hopfner and Cho labs revealed another DNA binding site, near the Rad50 dimer interface as predicted by Tainer [27,28]. Again, mutation of select residues in the interaction site reduced DNA affinity. A recent high-resolution cryo-EM structure of the DNA-bound MR complex has revealed a substantial rotation of the Mre11 dimer from the bottom of Rad50 to its side, which creates a channel for the dsDNA end to enter the Mre11 nuclease site from its position in the central groove of Rad50 [29]. DNA binding also triggers the coiled-coils of Rad50 to transition from a ring-like structure to a rod, which clamps around the dsDNA.

The recent structures of Rad50 and the MR complex bound to DNA have greatly clarified the potential DNA binding sites of Rad50. However, it is unclear if all the DNA binding modes have been revealed and the mechanism of the allosteric communication between the DNA binding, ATPase, and exonuclease activity remains uncertain. For this reason, we sought to identify additional DNA binding sites on Rad50 by employing sequenced-based DNA binding site prediction methods [30]. We analyzed T4-Rad50 using several prediction programs and each identified portions of the known DNA and ATP binding sites, along with a group of amino acids at the C-terminus of the protein [29,31–33]. To test these predictions, we deleted ten residues at the C-terminus of the protein and made charge reversal mutations of key positively charged residues. Consistent with the predictions, the mutations weakened Rad50's affinity for dsDNA and increased the Km-DNA for the nuclease activity of the MR complex. Additionally, exonuclease processivity and the activation of ATP hydrolysis upon binding of DNA was

severely reduced, suggesting that the C-terminus of Rad50 participates in DNA-dependent communication between Mre11 and Rad50.

Material and Methods

DNA binding site prediction– Several amino acid sequence-based DNA binding site prediction servers were used: DRNApred [34], BindN+ [33], DP-Bind [35], DNAbind [32], DNApred [36], and TargetDNA [31]. In general, the output of these prediction programs provides an overall score for each amino acid, which when combined with a cut-off value provides a binary answer for residue. In most cases, the number of predicted residues was high, so we set our cut-off at approximately 10% of the protein (i.e., best scoring 56 residues). The exception to this is TargetDNA, which only predicted 17 residues using the default values. For this output, we reduced the stringency of the cut-off until ~10% of the total residues were included in the prediction.

Mutagenesis, Protein Expression, and Purification– Site-directed mutations of Rad50 were generated following Stratagene QuikChange mutagenesis protocol [37]. The sequences of the DNA primers used are available upon request. Bacterial expression and purification of Mre11 and wild-type (WT) and mutated Rad50 proteins were carried out as described previously [14,15].

Steady-state ATPase Kinetics– ATPase activity of Rad50 or MR complex was measured with a fluorometric coupling assay [38]. Assays were performed at 30 °C using excitation and emission

wavelengths of 340 and 460 nm, respectively. The standard reaction buffer contained 50 mM Tris-HCl, pH 7.6, 50 mM KCl, and 5 mM $MgCl_2$, and 0.1 mg/ml BSA. The coupling components were 50 μM NADH, 150 μM phosphoenolpyruvate, 6.67 units/ml pyruvate kinase, and 10 units/ml lactate dehydrogenase. The reactions were initiated via addition of Rad50 or preassembled MR complex. Initial velocities (v) were measured at different substrate concentrations ($[S]$) and the maximum velocity (V_{max}), Michaelis constant (K_m), and Hill coefficient (n) were estimated according to equation 1 with SigmaPlot software.

$$v = \frac{V_{max} \cdot [S]^n}{{K_m}^n + [S]^n} \qquad \text{Eq. (1)}$$

The sequence of the oligonucleotides used to create the dsDNA substrate are identical to those used previously [15].

Determination of Nuclease Activity– Nuclease activity was determined fluorometrically using DNA substrates with a fluorescent 2-aminopurine deoxyribonucleotide (2-AP) at the 1st or 17th position relative to the 3'end of the substrate [39]. The assay was performed at 30 °C using the standard buffer with the addition of 0.3mM $MnCl_2$ and excitation and emission wavelengths of 310 and 375 nm, respectively. Various DNA substrate concentrations were used (0.5 to 20 μM) in determining the *k*cat and Michaelis constant (K_m) of DNA using the Eq. 1 with n set to 1.0.

Determination of Equilibrium Constants for DNA Binding– The DNA binding affinity of Rad50 and the MR complex was determined using a fluorescence polarization assay. Assays were performed at 22ºC using the standard buffer with 10 nM hexachlorofluorescein labeled DNA.

When the MR complex was used, Mre11 was held in excess over Rad50 (1.2:1 ratio, Mre11:Rad50). The ATP concentration 5-fold greater than the K_m-ATP for each protein. The reactions were started by the sequential addition of the protein and the DNA substrate to a well in a black 96-well microplate. Samples were excited at 520 nm and polarization data was collected at 560 nM. Data was analyzed and K_d values were determined using SigmaPlot.

Results and Discussion

DNA binding site prediction– Sequence-based DNA binding site prediction algorithms produce numerous false positives and negatives [30]. For this reason, we utilized several different algorithms and compared their outputs against each other. In total, 164 out of 560 residues were predicted to interact with DNA by at least one of the six algorithms, 55 residues were predicted by at least three out of the six algorithms, and just six residues were chosen by all six algorithms (Fig. 1A). The majority of the residues with at least three algorithms predicting a DNA interaction are found in the N-terminal half of the globular head domain, whereas the coiled-coil domain is largely devoid of predicted interactions (only 1 residue with at least 3 positive predictions) (Fig 1B). It is notable that almost the entire Walker A motif, which makes up portion of the ATP recognition site, was predicted to interact with DNA. While technically an incorrect prediction, it is not surprising that a highly conserved nucleoside-phosphate binding motif would be selected. Another set of highly predicted residues (59-70) are part of or in proximity to the dsDNA binding site revealed by x-ray crystal structures of Rad50 and the high-resolution cryo-EM structure of the MR complex bound to dsDNA that spans much of the Rad50 dimer interface [27,29,40].

We focused on our efforts on the grouping of residues at the extreme C-terminus that are strongly predicted to interact with DNA (551 through 557). In fact, the TargetDNA prediction software [31], which is the most recently developed and best performing of the six algorithms used here [31], ranks five of these C-terminal residues amongst the ten residues with the highest probability scores (Fig. 1C).

Role of the C-terminal residues in DNA binding– To test the predictions made by TargetDNA and the other algorithms, we constructed a C-terminal deletion mutant (Δ10Ct) and a triple charge reversal mutant (K551E/K552E/R555E). The mutants expressed to a similar degree as WT-Rad50 and behaved identically during the subsequent purification steps (data not shown). To determine the equilibrium binding constant for WT and each mutant protein binding to DNA, we employed a DNA anisotropy assay, which uses a 50 bp hexachlorofluorescein-labeled dsDNA as a substrate. Using this assay, the Kd-DNA for WT-Rad50 was determined to be 0.05 ± 0.02 μM, similar to our previously published value (Table I) [14]. Unlike other systems, the binding of T4-Rad50 to dsDNA is ATP-independent [14]. Δ10Ct-Rad50 binds dsDNA with 16-fold lower affinity that the WT enzyme, with a Kd-DNA of 0.82 ± 0.34 μM. The triple charge reversal mutant has an even lower DNA affinity than Δ10Ct, with a 37-fold increase in Kd-DNA (1.94 ± 0.29 μM), consistent with charge repulsion between the glutamates and the phosphate backbone of the DNA. These results confirm the computational predictions and suggest there may be more DNA binding surface on Rad50 than structures have indicated.

ATP affinity and hydrolysis activity of WT and mutant Rad50 proteins– In the absence of Mre11 and dsDNA, WT-Rad50 is a poor ATPase with a k_{cat} of 0.15 sec^{-1} and a K_M-ATP of 16 μM [15].

The Δ10Ct-Rad50 mutant has a reduced ATP hydrolysis rate with a *k*cat of 0.0013 sec^{-1}, and a much larger Km-ATP of 408 μM (Table I). The triple charge reversal mutant has a similar effect, with a Km-ATP of 849 μM and a *k*cat of 0.0064 sec^{-1}. Similar to the Km-ATP, the equilibrium binding constant for ATP (Kd-ATP) is increased by 6- and 15-fold for Δ10Ct and the triple mutant, respectively (Table I). These large reductions in ATP binding and hydrolysis are likely to be an allosteric effect, as these C-terminal residues are predicted to be at least 10 Å away from the ATP binding site on the basis of a structural homology model using the *M. thermatoga* Rad50 structure (Fig. 2).

When bound to DNA, the MR complex increases its ATP hydrolysis rate by approximately 20-fold, with a k_{cat} of 1.8 sec^{-1} and a Km of 49 μM (Table II). This activation is absent or reduced in both of the DNA binding mutants. These assays were carried-out at increased DNA concentrations to insure that DNA is saturating the mutant MR complexes. The *k*cat for Δ10Ct-Rad50 is reduced by 433-fold compared to the WT MR complex (k_{cat} = 0.003 sec^{-1}), but has a normal Km-ATP of 48 μM. The ATPase activity of the charge reversal mutant is less affected than the deletion, with a *k*cat value of 0.036 sec^{-1} and a Km-ATP of 129 μM. These results suggest that the C-terminal region of Rad50 plays an important role in coupling DNA binding to ATP hydrolysis activity.

DNA nuclease activity of WT and mutant MR complexes– While Rad50 provides most of the DNA binding activity for the MR complex, it also allosterically activates Mre11 for DNA binding [41]. T4-Mre11 contains a negatively charged flexible linker that binds to its own DNA binding site, thereby inhibiting itself. Upon complex formation with Rad50, the linker is sequestered by Rad50, exposing the DNA binding site on Mre11 [42]. We determined the Km-

DNA and the *k*cat-exo for WT and mutant MR complexes using a 50 bp DNA substrate with a 2-aminopurine residue at the 1 position (relative to the 3' end). With the 2AP residue in the 1 position, the nuclease reaction is ATP-independent [14,41]. The WT-MR complex has a Km-DNA of 3.8 μM and a k_{cat} of 41 min^{-1}. Similar to the increase in Kd-DNA, both the C-terminal deletion mutant and the charge reversal mutant have increased Km-DNA values of 10 and 16 μM, respectively (Table III). The nuclease turnover rates are relatively unchanged, however, with *k*cat values of 77 and 46 $minute^{-1}$ for Δ10Ct-Rad50 and the charge reversal triple mutant, respectively. The unchanged *k*cat values suggests that the positioning of the DNA in the exonuclease site of Mre11 is not affected by the mutations in the Rad50 subunit but that they are affecting the DNA binding mode that directly contributes to the exonuclease reaction.

When the position of the 2-aminopurine probe is moved away from the 3' end of the DNA substrate, the observed activity decreases and the reaction becomes activated by ATP hydrolysis. Previous result have suggested that the primary role of ATP in the exonuclease reaction is to increase the MR complex's processivity [14]. The WT-MR complex has a specific activity at saturating dsDNA (apparent-k_{cat}) of 0.019 min^{-1} in the absence of ATP and is activated 10-fold in the presence of at (0.18 min^{-1}). In the absence of ATP, the nuclease activity of the Δ10Ct-Rad50 and triple mutant are 17- and 111-fold reduced compared to the WT enzyme, respectively. This sharp reduction in exonuclease activity as the 2AP probe is moved from the end of the DNA suggests that the mutant MR complexes are less processive than the WT enzyme. Additionally, consistent the decrease in the ATPase activity of the mutant MR complexes, the ATP activation of the nuclease activity is also reduced. The nuclease activity of Δ10Ct-Rad50 is activated only 1.7-fold and the charge reversal mutant is activated 3.4-fold, as compared to ~10-fold for the WT enzyme.

Conclusions

The mutations generated here confirm that the C-terminus of Rad50 is somehow involved in dsDNA binding. The existing structures of Rad50 bound to DNA have revealed binding sites near the coiled-coil domain and at the dimer interface, whereas the C-terminus of Rad50 is distal from both these sites and thus represents a potentially novel site of interaction. Alternatively, it is possible that the effect of the mutations are allosteric, which would suggest that these C-terminal residues are a hotspot for the communication between the ATPase and DNA binding sites. An analysis of Rad50 sequences from bacteria (*Thermotoga maritima*), archaea (*Methanocaldococcus jannaschii*), and eukaryotes (*Chaetomium thermophilum*) using TargetDNA [31] predicted DNA binding residues near the C-terminus of each of these proteins. This suggests that this DNA binding site (or allosteric hotspot) may be conserved amongst Rad50 proteins and could represent a common route between Rad50 and the Mre11 nuclease site.

Acknowledgements:

This work was supported by the National Science Foundation [award number 1716269]

References

[1] R.D. Jachimowicz, J. Goergens, H.C. Reinhardt, DNA double-strand break repair pathway choice - from basic biology to clinical exploitation, Cell Cycle. 18 (2019) 1423–1434. https://doi.org/10.1080/15384101.2019.1618542.

[2] L.F. Povirk, Biochemical mechanisms of chromosomal translocations resulting from DNA double-strand breaks, DNA Repair (Amst). 5 (2006) 1199–1212.

https://doi.org/10.1016/j.dnarep.2006.05.016.

[3] M. Shrivastav, L.P. De Haro, J.A. Nickoloff, Regulation of {DNA} double-strand break repair pathway choice, Cell Res. 18 (2008) 134–147. https://doi.org/10.1038/cr.2007.111.

[4] T. Helleday, J. Lo, D.C. van Gent, B.P. Engelward, DNA double-strand break repair: from mechanistic understanding to cancer treatment, DNA Repair (Amst). 6 (2007) 923–935. https://doi.org/10.1016/j.dnarep.2007.02.006.

[5] J.M. Daley, H. Niu, A.S. Miller, P. Sung, Biochemical mechanism of DSB end resection and its regulation, DNA Repair (Amst). 32 (2015) 66–74. https://doi.org/10.1016/j.dnarep.2015.04.015.

[6] M.S. Dillingham, S.C. Kowalczykowski, {RecBCD} enzyme and the repair of double-stranded {DNA} breaks, Microbiol. Mol. Biol. Rev. MMBR. 72 (2008) 642–671. https://doi.org/10.1128/MMBR.00020-08.

[7] J. Liu, S.W. Morrical, Assembly and dynamics of the bacteriophage {T}4 homologous recombination machinery, Virol. J. 7 (2010) 357. https://doi.org/10.1186/1743-422X-7-357.

[8] J.C. Connelly, D.R.F. Leach, Tethering on the brink: The evolutionarily conserved Mre11-Rad50 complex, Trends Biochem. Sci. 27 (2002) 410–418. https://doi.org/10.1016/S0968-0004(02)02144-8.

[9] R.A. Deshpande, G.J. Williams, O. Limbo, R.S. Williams, J. Kuhnlein, J.-H. Lee, S. Classen, G. Guenther, P. Russell, J.A. Tainer, T.T. Paull, ATP-driven Rad50 conformations regulate DNA tethering, end resection, and ATM checkpoint signaling, EMBO J. 33 (2014) 482–500. https://doi.org/10.1002/embj.201386100.

[10] A.N. Lupas, M. Gruber, The Structure of α-Helical Coiled Coils, in: B.T.-A. in P.

Chemistry (Ed.), Academic Press, 2005: pp. 37–38. http://www.sciencedirect.com/science/article/pii/S0065323305700036 (accessed November 4, 2015).

[11] T. Barfoot, T.J. Herdendorf, B.R. Behning, B.A. Stohr, Y. Gao, K.N. Kreuzer, S.W. Nelson, Functional analysis of the bacteriophage T4 Rad50 Homolog (gp46) Coiled-coil Domain, J. Biol. Chem. 290 (2015) 23905–23915. https://doi.org/10.1074/jbc.M115.675132.

[12] M. Hohl, Y. Kwon, S.M. Galván, X. Xue, C. Tous, A. Aguilera, P. Sung, J.H.J. Petrini, The Rad50 coiled-coil domain is indispensable for Mre11 complex functions, Nat. Struct. Mol. Biol. 18 (2010) 1124–1131. https://doi.org/10.1038/nsmb.2116.

[13] J.R. Almond, B.A. Stohr, A.K. Panigrahi, D.W. Albrecht, S.W. Nelson, K.N. Kreuzer, Coordination and processing of DNA ends during double-strand break repair: The role of the bacteriophage T4 Mre11/Rad50 (MR) complex, Genetics. 195 (2013) 739–755. https://doi.org/10.1534/genetics.113.154872.

[14] D.W. Albrecht, T.J. Herdendorf, S.W. Nelson, Disruption of the bacteriophage T4 Mre11 dimer interface reveals a two-state mechanism for exonuclease activity, J. Biol. Chem. 287 (2012) 31371–31381. https://doi.org/10.1074/jbc.M112.392316.

[15] T.J. Herdendorf, D.W. Albrecht, S.J. Benkovic, S.W. Nelson, Biochemical characterization of bacteriophage T4 Mre11-Rad50 complex, J. Biol. Chem. 286 (2011) 2382–2392. https://doi.org/10.1074/jbc.M110.178871.

[16] T.H. Stracker, J.H.J. Petrini, The {MRE}11 complex: starting from the ends, Nat. Rev. Mol. Cell Biol. 12 (2011) 90–103. https://doi.org/10.1038/nrm3047.

[17] R.A. Deshpande, J.-H. Lee, S. Arora, T.T. Paull, Nbs1 Converts the Human Mre11 Rad50

Nuclease Complex into an Endo/Exonuclease Machine Specific for Protein-DNA Adducts, Mol. Cell. 64 (2016) 593–606. https://doi.org/10.1016/j.molcel.2016.10.010.

[18] M.L. Nicolette, K. Lee, Z. Guo, M. Rani, J.M. Chow, S.E. Lee, T.T. Paull, Mre11gRad50gXrs2 and Sae2 promote 5′ strand resection of DNA double-strand breaks, Nat. Struct. Mol. Biol. 17 (2010) 1478–1485. https://doi.org/10.1038/nsmb.1957.

[19] K.P. Hopfner, J.A. Tainer, Rad50/SMC proteins and ABC transporters: Unifying concepts from high-resolution structures, Curr. Opin. Struct. Biol. 13 (2003) 249–255. https://doi.org/10.1016/S0959-440X(03)00037-X.

[20] S.H. Harvey, M.J.E. Krien, M.J. O'Connell, Structural maintenance of chromosomes (SMC) proteins, a family of conserved ATPases, Genome Biol. 3 (2002) REVIEWS3003.

[21] A.L. Davidson, E. Dassa, C. Orelle, J. Chen, Structure, function, and evolution of bacterial {ATP}-binding cassette systems, Microbiol. Mol. Biol. Rev. MMBR. 72 (2008) 317–364. https://doi.org/10.1128/MMBR.00031-07.

[22] K.-P. Hopfner, L. Craig, G. Moncalian, R.A. Zinkel, T. Usui, B.A.L. Owen, A. Karcher, B. Henderson, J.-L. Bodmer, C.T. McMurray, J.P. Carney, J.H.J. Petrini, J.A. Tainer, The Rad50 zinc-hook is a structure joining Mre11 complexes in DNA recombination and repair, Nature. 418 (2002) 562–566. https://doi.org/10.1038/nature00922.

[23] J. He, L.Z. Shi, L.N. Truong, C.-S. Lu, N. Razavian, Y. Li, A. Negrete, J. Shiloach, M.W. Berns, X. Wu, Rad50 zinc hook is important for the Mre11 complex to bind chromosomal DNA double-stranded breaks and initiate various DNA damage responses, J. Biol. Chem. 287 (2012) 31747–31756. https://doi.org/10.1074/jbc.M112.384750.

[24] T.J. Herdendorf, S.W. Nelson, Catalytic mechanism of bacteriophage T4 Rad50 ATP hydrolysis, Biochemistry. 53 (2014) 5647–5660. https://doi.org/10.1021/bi500558d.

[25] K.P. Hopfner, A. Karcher, L. Craig, T.T. Woo, J.P. Carney, J.A. Tainer, Structural biochemistry and interaction architecture of the DNA double-strand break repair Mre11 nuclease and Rad50-ATPase, Cell. 105 (2001) 473–485. https://doi.org/10.1016/S0092-8674(01)00335-X.

[26] A. Rojowska, K. Lammens, F.U. Seifert, C. Direnberger, H. Feldmann, K.-P. Hopfner, Structure of the Rad50 DNA double-strand break repair protein in complex with DNA, EMBO J. 33 (2014) 2847–2859. https://doi.org/10.15252/embj.201488889.

[27] F.U. Seifert, K. Lammens, G. Stoehr, B. Kessler, K.-P. Hopfner, Structural mechanism of ATP-dependent DNA binding and DNA end bridging by eukaryotic Rad50., EMBO J. 35 (2016) 759–72. https://doi.org/10.15252/embj.201592934.

[28] Y. Liu, S. Sung, Y. Kim, F. Li, G. Gwon, A. Jo, A.A.-K. Kim, T. Kim, O.-K.O. Song, S.E. Lee, Y. Cho, ATP-dependent DNA binding, unwinding, and resection by the Mre11/Rad50 complex., 35 (2016) 743–758. https://doi.org/10.15252/embj.201592462.

[29] L. Käshammer, J.H. Saathoff, K. Lammens, F. Gut, J. Bartho, A. Alt, B. Kessler, K.P. Hopfner, Mechanism of DNA End Sensing and Processing by the Mre11-Rad50 Complex, Mol. Cell. 76 (2019) 382-394.e6. https://doi.org/10.1016/j.molcel.2019.07.035.

[30] J. Yan, S. Friedrich, L. Kurgan, A comprehensive comparative review of sequencebased predictors of DNA- and RNA-binding residues, Brief. Bioinform. 17 (2016) 88–105. https://doi.org/10.1093/bib/bbv023.

[31] J. Hu, Y. Li, M. Zhang, X. Yang, H. Bin Shen, D.J. Yu, Predicting Protein-DNA Binding Residues by Weightedly Combining Sequence-Based Features and Boosting Multiple SVMs, IEEE/ACM Trans. Comput. Biol. Bioinforma. 14 (2017) 1389–1398. https://doi.org/10.1109/TCBB.2016.2616469.

[32] C. Yan, M. Terribilini, F. Wu, R.L. Jernigan, D. Dobbs, V. Honavar, No Title, 7 (2006) 262. https://doi.org/10.1186/1471-2105-7-262.

[33] L. Wang, C. Huang, M.Q. Yang, J.Y. Yang, BindN+ for accurate prediction of DNA and RNA-binding residues from protein sequence features, BMC Syst. Biol. 4 (2010) S3. https://doi.org/10.1186/1752-0509-4-S1-S3.

[34] J. Yan, L. Kurgan, DRNApred, fast sequence-based method that accurately predicts and discriminates DNA-and RNA-binding residues, Nucleic Acids Res. 45 (2017) gkx059. https://doi.org/10.1093/nar/gkx059.

[35] S. Hwang, Z. Gou, I.B. Kuznetsov, DP-Bind: a web server for sequence-based prediction of DNA-binding residues in DNA-binding proteins, Bioinformatics. 23 (2007) 634–636. https://doi.org/10.1093/bioinformatics/btl672.

[36] Y.H. Zhu, J. Hu, X.N. Song, D.J. Yu, DNAPred: Accurate Identification of DNA-Binding Sites from Protein Sequence by Ensembled Hyperplane-Distance-Based Support Vector Machines, J. Chem. Inf. Model. 59 (2019) 3057–3071. https://doi.org/10.1021/acs.jcim.8b00749.

[37] Manual, QuikChange ™ Site-Directed For in Vitro Use Only, Mutagenesis. 200518 (2010) 1–18.

[38] S.P. Gilbert, A.T. Mackey, Kinetics: a tool to study molecular motors, Methods. 22 (2000) 337–354. https://doi.org/10.1006/meth.2000.1086.

[39] L.B. Bloom, M.R. Otto, R. Eritja, L.J. Reha-Krantz, M.F. Goodman, J.M. Beechem, Pre-steady-state kinetic analysis of sequence-dependent nucleotide excision by the 3'-exonuclease activity of bacteriophage T4 DNA polymerase, Biochemistry. 33 (1994) 7576–7586. http://www.ncbi.nlm.nih.gov/pubmed/8011623 (accessed October 19, 2010).

[40] Y. Liu, S. Sung, Y. Kim, F. Li, G. Gwon, A. Jo, A. Kim, T. Kim, O. Song, S.E. Lee, Y. Cho, ATP -dependent DNA binding, unwinding, and resection by the Mre11/Rad50 complex, EMBO J. 35 (2016) 743–758. https://doi.org/10.15252/embj.201592462.

[41] T.J. Herdendorf, D.W. Albrecht, S.J. Benkovic, S.W. Nelson, Biochemical characterization of bacteriophage T4 Mre11-Rad50 complex, J. Biol. Chem. 286 (2011) 2382–2392. https://doi.org/10.1074/jbc.M110.178871.

[42] Y. Gao, S.W. Nelson, Autoinhibition of bacteriophage T4 mre11 by its C-terminal domain, J. Biol. Chem. 289 (2014) 26505–26513. https://doi.org/10.1074/jbc.M114.583625.

[43] L.A. Kelley, S. Mezulis, C.M. Yates, M.N. Wass, M.J.E. Sternberg, The Phyre2 web portal for protein modeling, prediction and analysis, Nat. Protoc. 10 (2015) 845–858. https://doi.org/10.1038/nprot.2015.053.

Figure Legends

Figure 1. Summary of Rad50/DNA interaction site prediction algorithms. *A)* Distribution of the number of amino acids predicted to interact with DNA. The number above the bar graph represents the exact number of amino acids predicted by a given number of algorithms. *B)* A plot of Rad50 amino acids with at least three algorithms predicting a positive DNA binding predictions. *C)* The TargetDNA prediction for the probability of each Rad50 amino acid interacting with DNA.

Figure 2. Structural homology model of T4-Rad50. The model was generated with the Phyre2 [43] suite using the structure of *Pyrococcus yayanosii* Smc protein (PDBid = 5XEI) as a

homologous template (15% identity, 100 confidence [43]). The positively charged residues near the C-terminus of T4-Rad50 are indicated and colored red. The DNA and ATPγS are modelled from a structural alignment of the T4-Rad50 homology model and the MR complex from *Methanocaldococcus jannaschii* (PDBid = 5F3W).

Table I – Rad50 Substrate Binding and Kinetic Constants

Protein	Kd-DNA μM	Km-ATP μM	*k*cat sec-1	Kd-ATP μM
WT	0.13 ± 0.02	16 ± 1	0.15 ± 0.01	47± 10
ΔCtDBD	0.82 ± 0.34	408 ± 72	0.0013 ± 0.0003	190 ± 85
K551E/K552E/ R555E	1.94 ± 0.29	849 ± 172	0.0064 ± 0.0006	142 ± 85

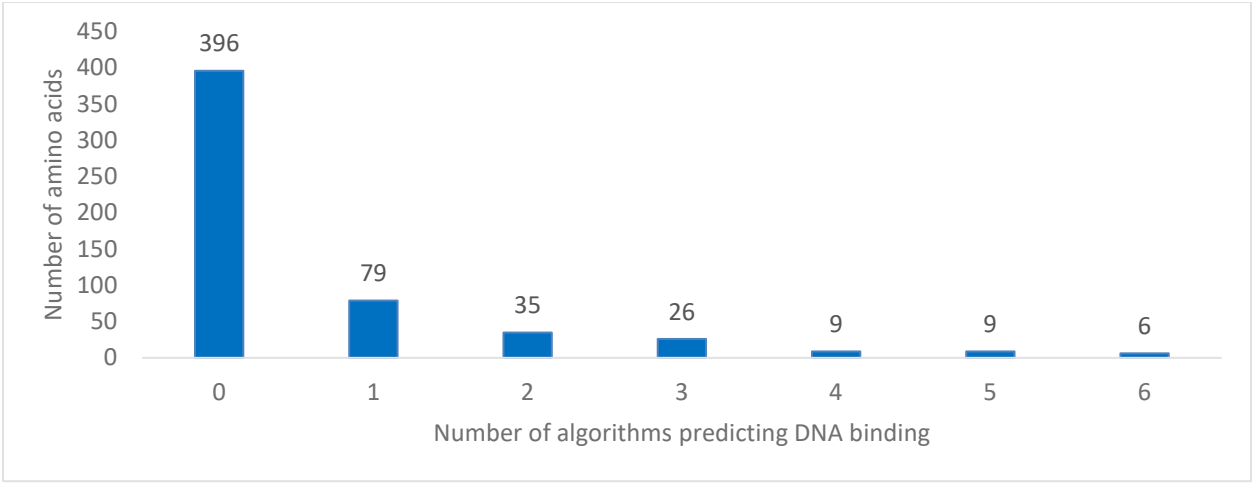
Table II – MR Complex ATPase Activity

Protein	Km-ATP μM	*k*cat sec-1
WT	49 ± 2	1.3 ± 0.1
ΔCtDBD	48 ± 7	0.0030 ± 0.0007
R555E/ K551E/K552E	129 ± 35	0.036 ± 0.005

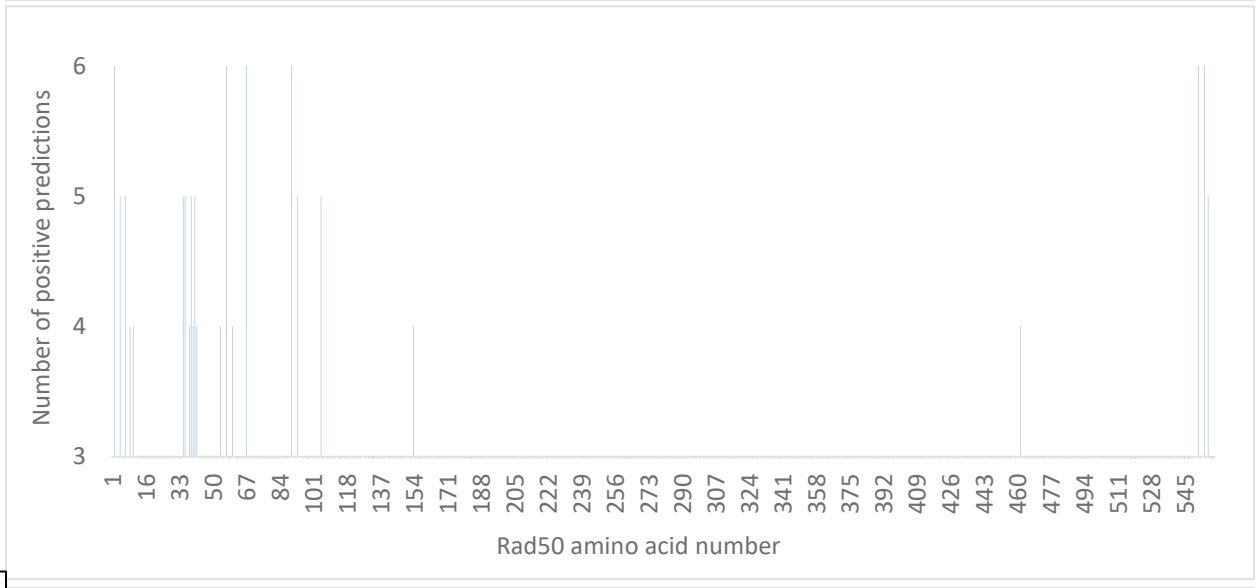
Table III – MR Complex Nuclease Activity

Protein	Km-DNA (1 pos) μM	*kcat*-nuclease (1 pos) min^{-1}	Specific Activity (17 pos) min^{-1}	Specific Activity (17 pos + ATP) min^{-1}
WT	3.8 ± 0.7	41 ± 3	0.019 ± 0.007	0.18 ± 0.01
ΔCtDBD	10 ± 1	77 ± 4	0.0011±0.0001	0.0019 ± 0.0002
R555E/ K551E/K552E	16 ± 2	46 ± 3	0.00017± 0.00002	0.00059 ± 0.00003

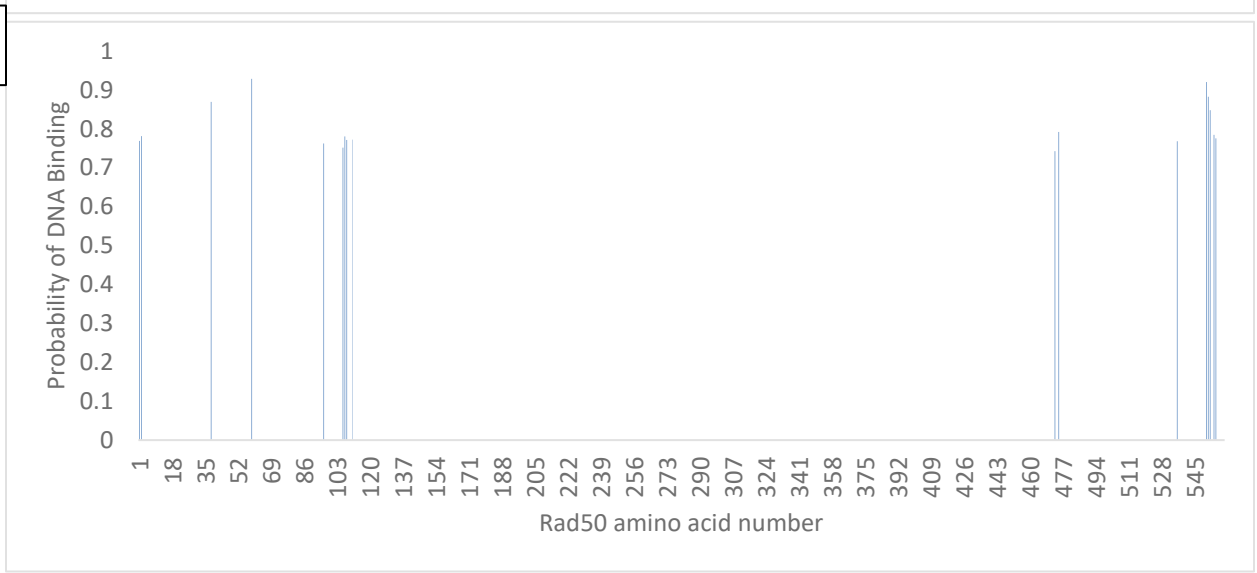
A

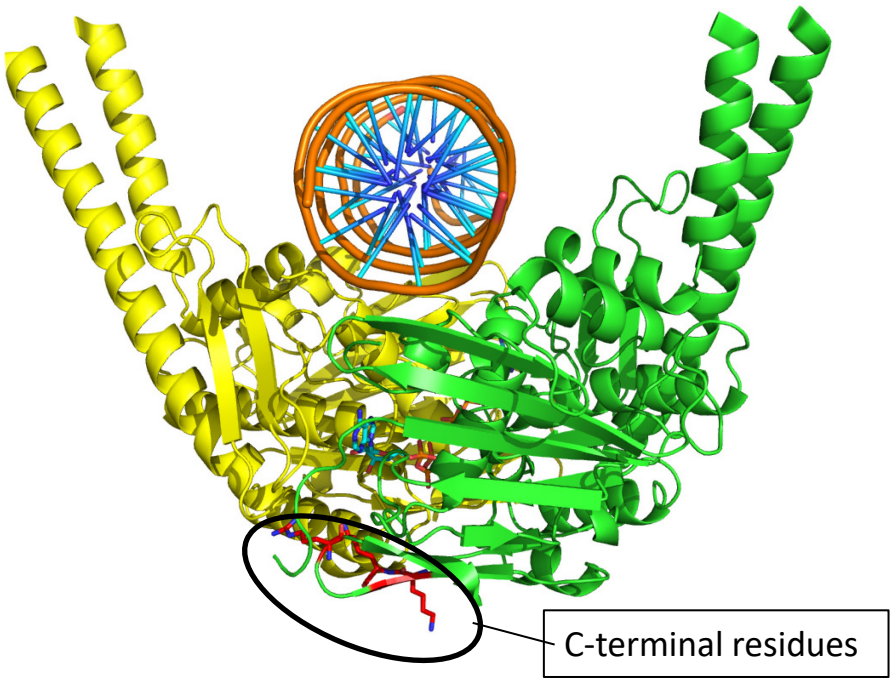

Number of amino acids
450
400
350
300
250
200
150
100
50
0
396
79
35
26
9
9
6
0
1
2
3
4
5
6
Number of algorithms predicting DNA binding

B


Number of positive predictions
6
5
4
3
1
16
33
50
67
84
101
118
137
154
171
188
205
222
239
256
273
290
307
324
341
358
375
392
409
426
443
460
477
494
511
528
545
Rad50 amino acid number

C


Probability of DNA Binding
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1
18
35
52
69
86
103
120
137
154
171
188
205
222
239
256
273
290
307
324
341
358
375
392
409
426
443
460
477
494
511
528
545
Rad50 amino acid number


C-terminal residues