Molecular Modelling and Biochemical Characterization of a DNA Repair Enzyme Binding to DNA

Brandon Manary, Nina Bernstein*†, Jorge Llano*‡

† *Department of Biological Sciences, MacEwan University, PO Box 1796, Alberta T5J 2P2, Canada*
‡ *Department of Physical Sciences, MacEwan University, PO Box 1796, Alberta T5J 2P2, Canada*

Introduction

Polynucleotide kinase/phosphatase (PNKP) is an enzyme that plays an essential role in various DNA strand-break repair pathways, such as single-strand break repair (SSBR), nonhomologous end-joining (NHEJ) and base-excision repair (BER).[1] After DNA is damaged by agents such as reactive oxygen species, ionizing radiation or mutagenic chemicals, the backbone may break, leaving incorrect functional groups on what are normally the 5′-phosphate or 3′-hydroxyl termini. PNKP repairs this damage before the backbone is ligated again, by restoring the ends to their proper terminal groups using the enzyme's bifunctional 5′-kinase and 3′-phosphatase activities.

Mutations that disrupt the catalytic activity of PNKP give rise to neurological diseases, such as microcephaly with seizures (MCSZ) and ataxia with oculomotor apraxia 4 (AOA4).[2] Because it is a DNA repair enzyme, PNKP is being investigated as a potential drug target of sensitizing agents for chemotherapy and radiation therapy in cancer treatment.[3]

C. elegans PNKP (CePNKP) is a useful model for studying human PNKP (hPNKP), but, intriguingly, CePNKP shows a more exclusive preference for DNA substrates with recessed 5′-hydroxyl ends than hPNKP. Notably, the kinase domain of mammalian PNKP (mPNKP) prefers phosphorylating DNA substrates with recessed 5′-hydroxyl ends over DNA substrates with blunt ends.[4]

To elucidate the reason for the unique substrate specificity of CePNKP, the structure of CePNKP in complex with its DNA substrate must be determined in atomistic detail by molecular modelling. The generated structural model is compared with our experimental results of kinase activity assays of CePNKP.

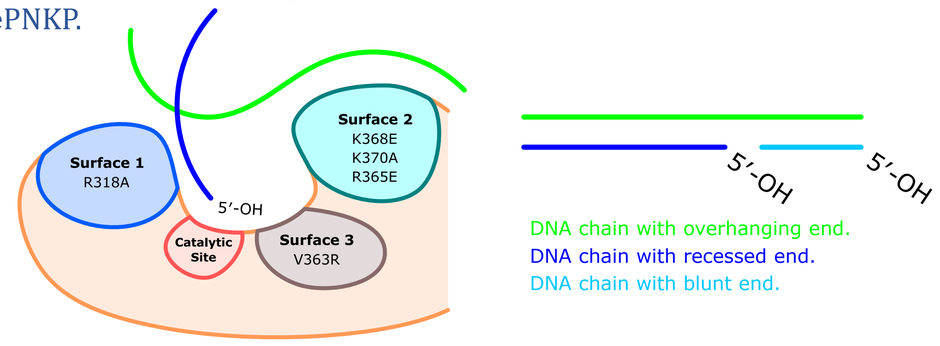

Figure 1. The three DNA-binding surfaces on PNKP. Amino-acid site mutations that disrupt the binding of DNA to PNKP are listed for each surface.[5]

Experimental Section

Wild-type CePNKP (WT Ce), and a mutant for each surface (**Fig. 1**) with amino-acid substitutions in positions likely to be binding sites for the DNA molecule were analyzed.

Quantification of proteins. In order find and compare the concentrations of protein in the stock solutions, the A280 assay was applied to measure the absorbances of 2 μL drops of each protein solution. This data was used to determine protein concentrations by the Warburg–Christian method.

Then, each protein stock solution was diluted to the same concentration of 18.20 mg/mL. This was done to ensure that the results of the kinase activity assay could be comparable.

Kinase activity assay. Each 18.20 mg/mL protein solution was serially diluted with dilution factors of 1, 2, 4, 8, 16, 32, 64. Equal amounts of substrate (oligo DNA and ATP) were added to each diluted protein solution. The solutions were incubated at 16 °C for 30 minutes. Then, the reactions were stopped with a urea loading buffer. The solutions were subjected to denaturing polyacrylamide gel electrophoresis (PAGE) to qualitatively assess the catalytic activity of each wild-type and mutant CePNKP on 5′-OH groups at the recessed and blunt ends of DNA.

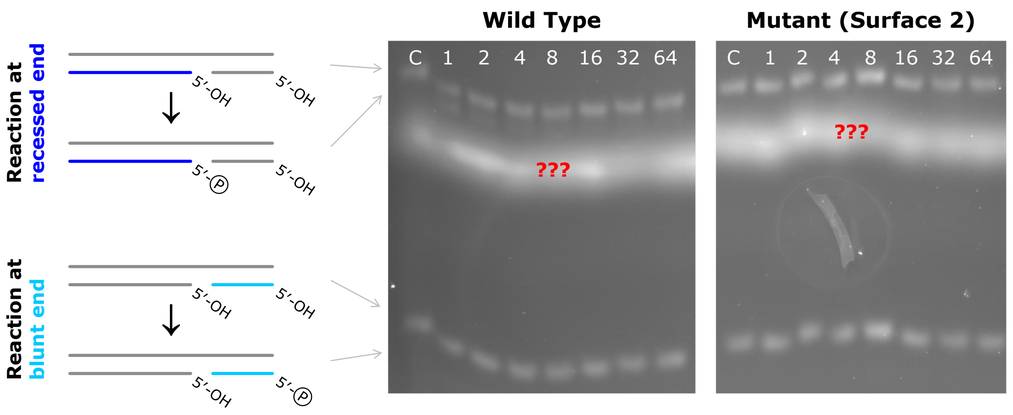

Figure 2. PAGE visualization of the CePNKP phosphorylation activity.

Computational Section

For CePNKP, the protein sequence is available, but its 3D structure has not been determined experimentally. Hence, we performed homology modelling to thread the CePNKP sequence into the mouse PNKP fold found by X-ray crystallography.

Homology modelling. MODELLER 9.17 software[6] was used to create a 3D structural model of CePNKP using mouse PNKP as a template. The procedure for this is summarized in **Fig. 3**:

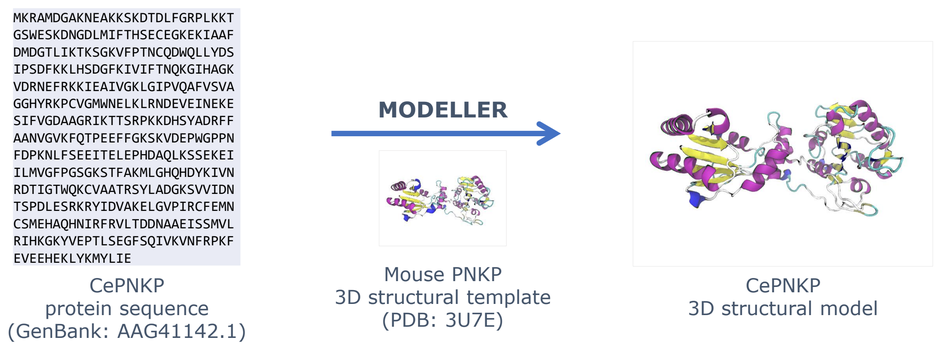

Figure 3. Flowchart of the homology modelling procedure.

GROMACS 2016.1 software[7] was used to set up structural models for solutions and to perform MD simulations with the force field AMBER94 for nucleic acids and AMBER99 for proteins.

Model of Oligo-DNA in solution: The ideal 3D structure of B-DNA was the starting geometry for the oligo-DNA substrate. The oligo DNA was placed into a box of water (8.09 nm × 5.11 nm × 5.16 nm) containing 0.15 M NaCl ions.

Model of CePNKP in solution: The 3D homology model was the starting geometry for the CePNKP enzyme. The enzyme was placed into a box of water (10.25 nm × 6.78 nm × 6.69 nm) containing 0.15 M NaCl ions.

Molecular Dynamics simulations. Each solution was sequentially subjected to: (1st) energy minimization, (2nd) *NVT* equilibration, (3rd) *NPT* equilibration for energy relaxation, and (4th) *NPT* equilibration to produce the trajectory each molecule's motion in water at 1 atm and 300 K over time.

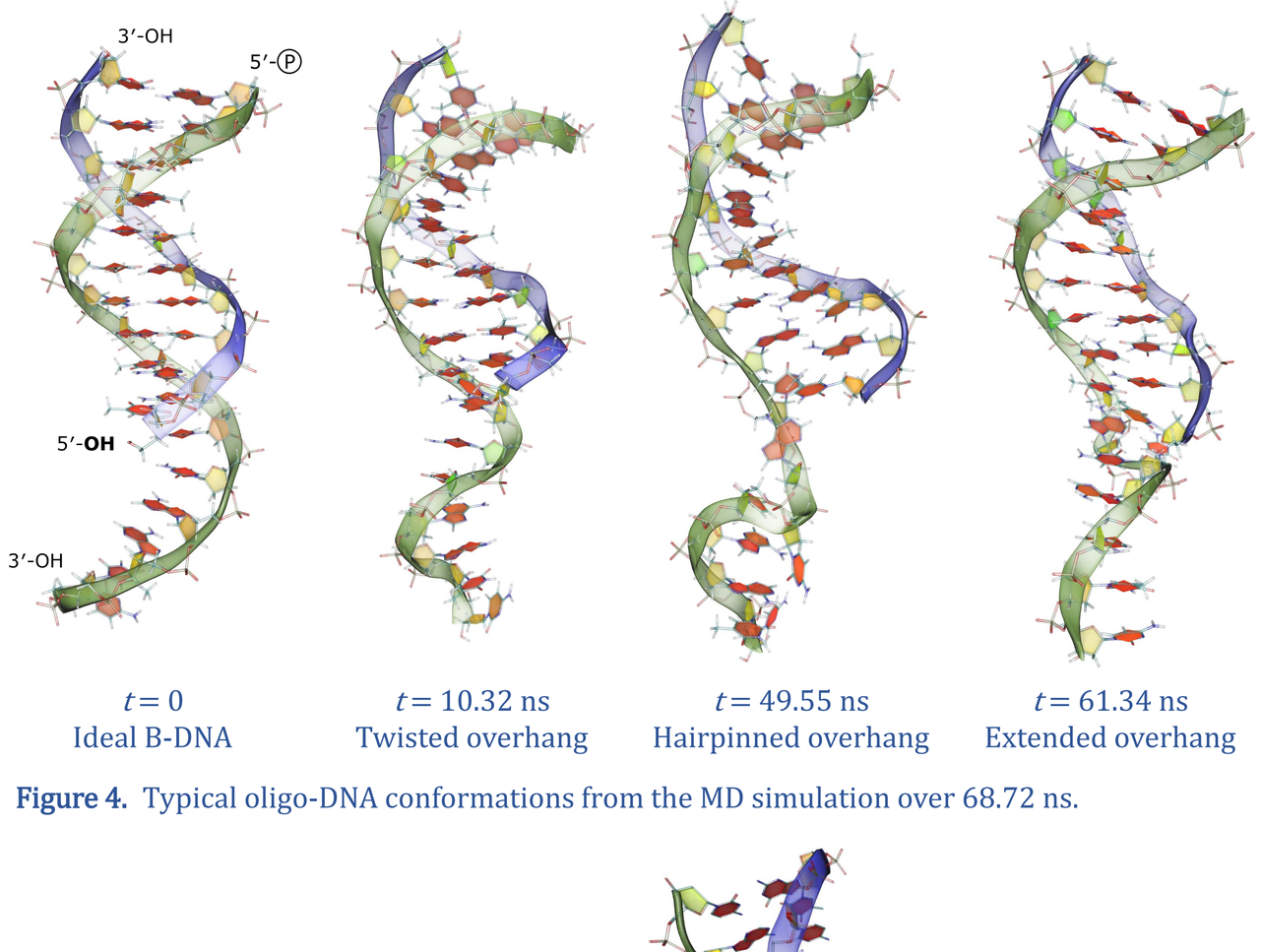

Figure 4. Typical oligo-DNA conformations from the MD simulation over 68.72 ns.

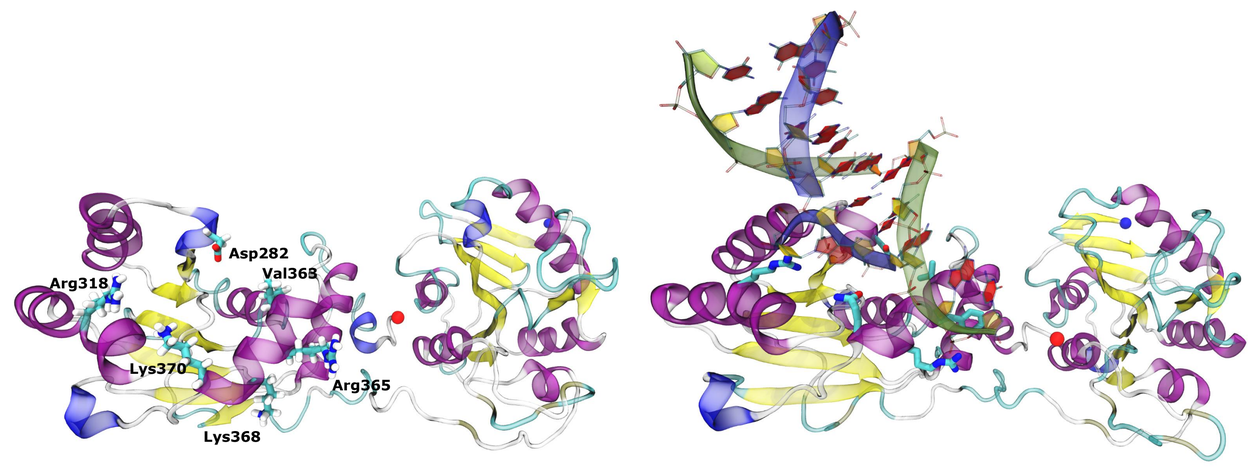

Figure 5. Ribbon representations of three-dimensional structures of PNKP. Side chains of catalytic and mutated (**Fig. 1**) residues depicted in stick representation.
Left: CePNKP structure at $t = 2.80$ ns from the MD simulation over 6.12 ns.
Right: mPNKP in complex with oligo DNA from crystallographic data (PDB: 3ZVN).

Conclusions

- The protein quantification and kinase assays will be repeated to tune up the experimental conditions that drive the phosphorylation reaction forward.
- 3D structural models for each CePNKP and oligo DNA were generated by MD simulations.
- Typical conformations and regions of flexibility in the binding sites of the structural models of CePNKP and oligo DNA were identified.
- The oligo DNA in complex with CePNKP will be generated by docking the structural models generated in this work. The complex will be studied by MD simulations.
- *In silico* mutagenesis will be performed in the structural model of the complex to gain insight into the binding of CePNKP with its DNA substrate.

References

(1) Weinfeld, M.; Mani, R. S.; Abdou, I.; Aceytuno, R. D.; Glover, J. N. M. *Trends Biochem. Sci.* **2011**, *36*, 262.
(2) Dumitrache, L. C.; McKinnon, P. J. *Mech. Ageing Dev.* **2017**, *161*, 121.
(3) Curtin, N. J. *Nat. Rev. Cancer* **2012**, *12*, 801.
(4) Bernstein, N. K.; Williams, R. S.; Rakovszky, M. L.; Cui, D.; Green, R.; Karimi-Busheri, F.; Mani, R. S.; Galicia, S.; Koch, C. A.; Cass, C. E.; Durocher, D.; Weinfeld, M.; Glover, J. N. M. *Mol. Cell* **2005**, *17*, 657.
(5) Bernstein, N. K.; Hammel, M.; Mani, R. S.; Weinfeld, M.; Pelikan, M.; Tainer, J. A.; Glover, J. N. M. *Nucleic Acids Res.* **2009**, *37*, 6161.
(6) MODELLER, Version 9.17; «https://salilab.org/modeller/»
(7) GROMACS, Version 2016.1; «http://www.gromacs.org/»