

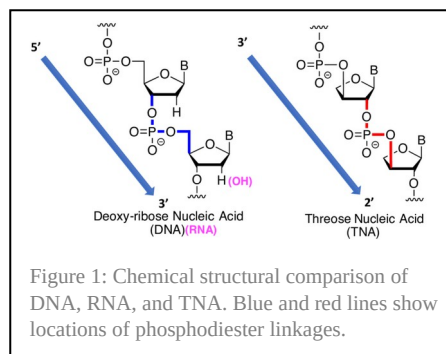
Towards Structural Elucidation of Closed Ternary Complexes of Kod-RSGA with Threose Nucleic Acid Triphosphates.

Background

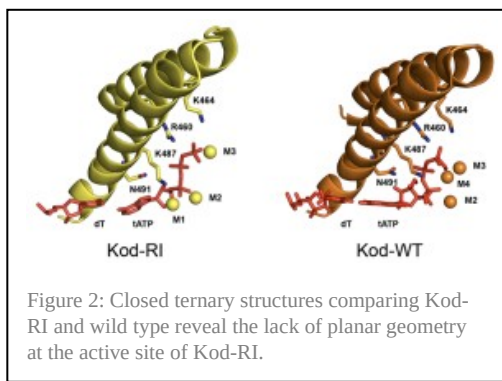
All living organisms must precisely replicate their genome in order to pass their genetic information on to their progeny and to ensure cell viability. Highly specialized enzymes called DNA polymerases catalyze DNA replication through template-guided deoxyribonucleotide additions to a primer strand. The first DNA polymerase was discovered in *Escherichia coli* by Arthur Kornberg in 1956¹, and since then many other polymerases composing eight different phylogenetic families have been found in nature². To date, the ability of polymerases to generate custom nucleic acid sequences represents enormous potential for the deployment of these nucleic acids in various applications such as diagnostics, therapeutics, and information storage³.

Despite the massive potential found in these natural nucleic acids, one drawback is their susceptibility to degradation by nucleases. Modern synthetic biology has addressed this issue with the creation of xeno nucleic acids (XNAs), nucleic acids with modified sugar backbones^{4,5}. α -L-threose nucleic acid (TNA) is an XNA which replaces the five-carbon ribose sugar in DNA/RNA with a four-carbon threose sugar (Figure 1).

In contrast to the 5'-3' phosphodiester linkages which characterize DNA and RNA, TNA extends in a 3'-2' direction⁶, thereby making TNA resistant to the nuclease degradation that afflicts natural nucleic acids. Despite these differences, TNA can form stable antiparallel duplexes with itself, DNA, and RNA, using traditional Watson-Crick base pairing, and can successfully replicate in a Darwinian evolution system⁷. These characteristics make TNA an attractive candidate for applications in therapeutics and diagnostics⁸.



Engineered polymerases are required to recognize XNAs as substrates. In 2016 the Chaput lab evolved Kod-RI, an engineered polymerase capable of TNA synthesis⁹. Kod-RI is a variant of Kod DNA polymerase from the archaeal species *Thermococcus kodakarensis*, with mutations present at A485R and E664I. While this engineered polymerase is capable of TNA synthesis, it proceeds at a rate of about 1 nucleotide addition per minute, which represents a >10,000-fold decrease from the wild-type (Kod-WT) protein, which proceeds at about 200 nucleotide additions per second⁹. Structural analysis of the ternary complex of Kod-RI bound to a primer/template duplex and an incoming TNA

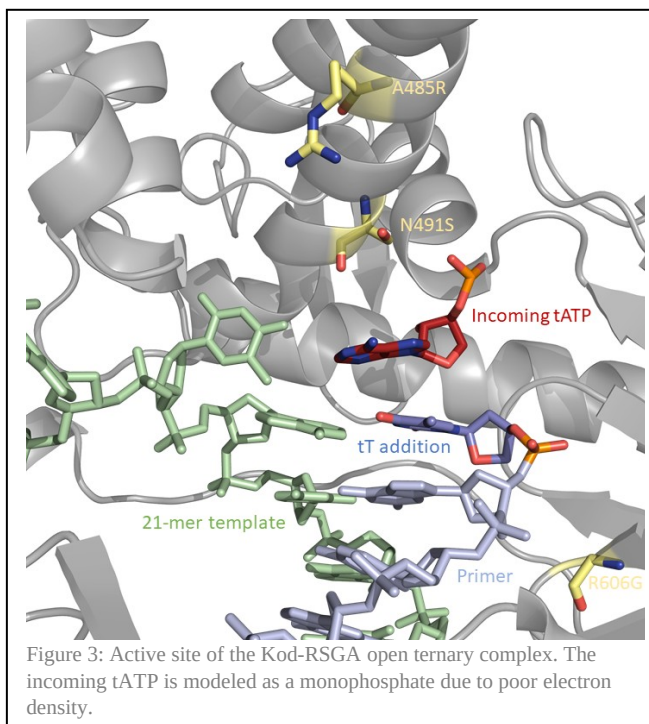


triphosphate revealed sub-optimal geometry in the active site compared to Kod-WT⁷, which may lead to its reduced catalytic rate (**Figure 2**).

More recently, the Chaput lab engineered a new TNA polymerase, Kod-RSGA, capable of catalyzing about 10 nucleotide additions per minute. Kod-RSGA contains A485R, N491S, R606G, and T723A mutations, and represents a ten-fold increase over the previous generation of TNA polymerase¹⁰. We hypothesize that the Kod-RSGA active site will exhibit more favorable planar geometry than Kod-RI, promoting improved catalytic rates.

Project Update

Over the fall and winter quarters, we successfully determined a 3.05 Å Kod-RSGA ternary structure which reveals an incoming α -L threofuranosyl adenosine 3'-triphosphate (tATP) in the active site with the finger subdomain in an open conformation (**Figure 3**). This was primarily achieved using a 21-nucleotide long template that was previously reported to improve crystallization of Kod-WT ternary complexes¹¹, an approach we proposed in the last UROP funding period. While the open ternary structure can provide pertinent information regarding the mechanism of TNA synthesis, we will continue to actively pursue the structure of the closed ternary structure, which represents the catalytically competent state of polymerases.



Objectives

The goal of this project is to elucidate the closed ternary structure of Kod-RSGA in complex with a primer-template duplex and an incoming TNA triphosphate, allowing us to gain insights into the mechanism of TNA synthesis. To that end, we are proposing two new strategies. (1) We will soak existing crystals grown in the open ternary crystallization condition with excess tATP to induce the close conformation. (2) Upon inspection of existing Kod structures, there is a ~20 amino acid residue stretch at the C-terminal that is unstructured. We have performed site-directed mutagenesis to obtain two truncated Kod-RSGA constructs. These new truncated constructs, Kod-RSGA Δ 759 and Kod-RSGA Δ 749, may improve the ability for crystal formation and the chances for closed ternary complex crystals.

Specific Aims

- Kod-RSGA, Kod-RSGA Δ 759, and Kod-RSGA Δ 749 overexpression and purification
- Crystal screens and optimization trays for the ternary complexes of Kod-RSGA Δ 759 and Kod-RSGA Δ 749 bound to a template/primer duplex and an incoming tNTP.
- Soak existing open ternary crystals in excess tATP for harvest and structural study.
- Structural characterization of the ternary complexes of Kod-RSGA Δ 759 and Kod-RSGA Δ 749 by X-ray crystallography.

Methods

Kod Expression and Purification:

The expression and purification of Kod-RSGA will follow a previously published protocol⁶. In brief, the *kod-RSGA* gene in the vector pET21 (Novagen), pET21-Kod-RSGA, will be transformed into Acella® cells (Edge Biosystems), then grown at 37 °C in LB medium containing 100 $\mu\text{g mL}^{-1}$ ampicillin in an aerobic environment. When cell growth reaches an OD₆₀₀ of 0.8, 0.8 mM isopropyl β -D-thiogalactoside (IPTG) will be added to induce the expression of Kod-RSGA for 20 h at 18°C. Cells will be harvested by centrifugation for 20 min at 3315 \times g at 4 °C, then lysed by sonication in 40 mL of lysis buffer (10 mM Tris.Cl pH 7.5, 100 mM NaCl, 0.1 mM EDTA, 1 mM DTT, 10% glycerol, 5 mg egg hen lysozyme). The lysate will be centrifuged at 23,708 \times g for 30 min, and the clear supernatant will be heat-treated for 20 min at 70 °C and centrifuged again at 23,708 \times g for 30 min. The supernatant is then to be loaded onto 5 mL HiTrap Q HP and heparin HP columns (GE) assembled in series with the efflux of the Q column loaded in front of the heparin column. The Q column is to be washed with lysis buffer before removal, and the polymerase is to be eluted from the heparin column using a high salt buffer (10 mM Tris.Cl pH 7.5, 1 M NaCl, 0.1 mM EDTA, 1 mM DTT, 10% glycerol) in a linear gradient. The eluted fractions expected to contain the selected polymerase are to be visualized by SDS-PAGE, pooled, and concentrated using a 30 kDa cutoff Amicon centrifugal filter (Millipore). Further purification will be achieved by size exclusion chromatography (Superdex 200 HiLoad 16/600, GE) pre-equilibrated with Kod-RSGA buffer (50 mM Tris.Cl pH 8.5, 200 mM NaCl, 0.1 mM EDTA, 1 mM DTT). Purified polymerase will be concentrated to 10 mg mL⁻¹ for crystallization trials. We expect for Kod-RSGA Δ 759 and Kod-RSGA Δ 749 to express and purify the same as Kod-RSGA in this process.

Crystallization of ternary TNA polymerase complexes:

An initial binary complex will be made by incubating Kod-RSGA Δ 759 or Kod-RSGA Δ 749 at 37 °C for 30 minutes with 1.2 molar equivalence of the primer-template duplex (primer: 5' – GACCACGGCCAC – 3'; 21-mer template: 5' – TATGCAACTGT GGCCGTGGTC – 3') in the presence of a buffer supplemented with 20mM MgCl₂. Following this, a 4 M excess of 2'-deoxy- α -L-threofuranosyl thymine triphosphate will be added and incubated at 37 °C for 30 minutes. The ternary complex will be formed by addition of α -L-threofuranosyl adenine triphosphate (tATP). This final complex will be incubated at 37 °C for 30 minutes once again and will then be screened against approximately 800 conditions using a hanging-drop screening method via a Mosquito crystallization robot (TTP Techlab Ltd). The procedure will be repeated with modified tUTP in place of tATP, utilizing a different template

that will be compatible with the incoming tUTP. Positive crystal hits will be further optimized in crystal trays, and promising crystals will be harvested in liquid nitrogen and sent to the Advanced Light Source synchrotron facility (Berkeley Lab) to obtain diffraction patterns via X-ray crystallography.

Itemized Budget

Quantity	Item Description	Unit Price	Total Price
2 x 100 Plate Case	24 Well H.D. Plate	\$489.00	\$978.00
3 x 10oz case (~1200 slides)	24 Well Round Cover Slide	\$292.00	\$876.00
2 packs, 25 loops each	Mounted CryoLoop-10 Micron 0.05-0.1mm	\$79.00	\$158.00
2 packs, 25 loops each	Mounted CryoLoop-10 Micron 0.1-0.2mm	\$79.00	\$158.00
		TOTAL	\$2170.00

Student Responsibility

Our responsibilities will include expression and purification of Kod-RSGA, Kod-RSGA Δ 759, and Kod-RSGA Δ 749, setting up crystal trays with varied conditions, and optimizing crystal tray conditions that support promising crystal growth. As undergrads we will be working under the guidance of [Redacted] and [Redacted]. Working on this project with [Redacted] since the start of this academic year has provided a wonderful opportunity for our growth in science. The Chaput lab has been a wonderful learning experience, with a plethora of knowledge accessible to us via helpful conversations with a variety of friendly graduate students, as well as post-docs and experienced scientists.

References

- ¹Bessman MJ, Kornberg A, Lehman IR, Simms ES. Enzymic synthesis of deoxyribonucleic acid. *Biochim Biophys Acta*. 21(1):197-198 (1956) doi:10.1016/0006-3002(56)90127-5.
- ²Yang W, Gao Y. Translesion and repair DNA polymerases: diverse structure and mechanism. *Annu Rev Biochem*. 87:239-61(2018). <https://doi.org/10.1146/annurev-biochem-062917-012405>.
- ³Aschenbrenner J, Marx A. DNA polymerases and biotechnological applications. *Curr Opin Biotechnol*. 48:187-195 (2017). doi:10.1016/j.copbio.2017.04.005.
- ⁴Anosova I, Kowal EA, Dunn MR, Chaput JC, Van Horn WD, Egli M. The structural diversity of artificial genetic polymers. *Nucleic Acids Res*. 44, 1007–1021 (2016). doi:10.1093/nar/gkv1472.
- ⁵Chaput JC, Herdewijn P, Hollenstein M. Orthogonal Genetic Systems. *ChemBioChem* 21, 1408 (2020). <https://doi.org/10.1002/cbic.201900725>

⁶ Schöning K, Scholz P, Guntha S, Wu X, Krishnamurthy R, Eschenmoser A. Chemical etiology of nucleic acid structure: the alpha-threofuranosyl-(3' → 2') oligonucleotide system. *Science*. 290(5495):1347-1351. (2000) doi:10.1126/science.290.5495.1347

⁷ Yu H, Zhang S, Chaput JC. Darwinian evolution of an alternative genetic system provides support for TNA as an RNA progenitor. *Nature Chemistry*. 4, 183-187 (2012). <https://doi.org/10.1038/nchem.1241>

⁸ Chaput JC, Redesigning the genetic polymers of life. *Accounts of Chemical Research* 49:646-656 (2016). <https://doi.org/10.1021/acs.accounts.0c00886>

⁹ Dunn MR, Otto C, Fenton KE, Chaput JC. Improving Polymerase Activity with Unnatural Substrates by Sampling Mutations in Homologous Protein Architectures. *ACS Chemical Biology*. 11(5):1210-1219. (2016) <https://doi.org/10.1021/acscchembio.5b00949>

¹⁰ Nikoomanzar A, Vallejo D, Yik EJ, Chaput JC, Programmed Allelic Mutagenesis of a DNA Polymerase with Single Amino Acid Resolution. *ACS Synth. Biol.* 9, 7, 1873-1881 (2020). <https://doi.org/10.1021/acssynbio.0c00236>

¹¹ Kropp HM, Ludmann S, Diederichs K, Betz K, Marx A, Structural Basis for The Recognition of Deaminated Nucleobases by An Archaeal DNA Polymerase. *ChemBioChem* 22, 3360 (2021).