Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____ _____

Xiao Zhang Date

Application of fluorescent nucleoside analogs in direct evolution of *Drosophila melanogaster* deoxyribonucleoside kinase

By

Xiao Zhang
Master of Science

Chemistry

_____

Stefan Lutz, Ph.D.
Advisor

_____

Vince Conticello, Ph.D.
Committee Member

_____

Dennis Liotta, Ph.D.
Committee Member

Accepted:

_____

Lisa A. Tedesco, Ph.D.
Dean of the James T. Laney School of Graduate Studies

_____

Date

Application of fluorescent nucleoside analogs in direct evolution of *Drosophila melanogaster* deoxyribonucleoside kinase

By

Xiao Zhang
B.A., Nanjing University, 2006

Advisor: Stefan Lutz, Ph.D.

An abstract of
a thesis submitted to the
Faculty of the James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Master of Science
in Chemistry
2009

Abstract

Application of fluorescent nucleoside analogs in direct evolution of *Drosophila melanogaster* deoxyribonucleoside kinase

By Xiao Zhang

Nucleoside analogs (NAs) are widely used as prodrugs, such as in highly active antiretroviral therapy (HAART) and chemotherapy. NAs require cellular enzymes to convert them into their active triphosphate form in order to be incorporated into DNA by low fidelity polymerase such as HIV reverse transcriptase or polymerase β in cancer cells during replication cycle, functioning as reverse transcriptase and polymerase inhibitor. Phosphorylation of NAs to monophosphates by deoxyribonucleoside kinases (dNKs) is usually the 'bottle neck' in prodrug activation. Engineering 2′-deoxyribonucleoside kinases with higher activity and orthogonal specificity for nucleoside analogs is under intensive study. However, the directed evolution of dNKs is hampered by a lack of efficient library screening techniques. A method combining fluorescent nucleoside analogs with fluorescent-active cell sorting (FACS) has been developed in our lab and evaluated by evolving *Drosophila melanogaster* deoxyribonucleoside kinase (*Dm*-dNK) using fluorescent 2′,3′-dideoxythymidine (fddT) as substrate. As part of my thesis, I have prepared a fluorescent 3′-fluoro-2′-deoxyuridine (fFT) and explored its use to evolve kinases that are specific to 3′-fluoro-2′-deoxyuridine (FT). My results show that the partial hydrogen bond interaction between the 3′ fluoro moiety and active site residues is beneficial for enrichment of desirable *Dm*-dNK mutants by FACS. Seperately, I have investigated alternative fluorescent probes that do not clash with residues in the active site, employing 1, 3-dipolar cycloaddition of alkenophiles with diaryl-tetrazoles.

Application of fluorescent nucleoside analogs in direct evolution of *Drosophila melanogaster* deoxyribonucleoside kinase

By

Xiao Zhang
B.A., Nanjing University, 2006

Advisor: Stefan Lutz, Ph.D.

A thesis submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Master of Science
in Chemistry
2009

Acknowledgment

First and foremost, I would like to thank Dr. Stefan Lutz for his guidance and support for my graduate study in his lab and in the Department of Chemistry. In his lab, I have learned not only the enthusiasm for science but also the meticulous and critical attitude towards work.

Secondly, I would like to thank Dr. Dennis Liotta and Dr. Vince Conticello. They teach excellent classes, while being really resourceful and patient for instructing me in my research work.

Also, my thank goes to former and present lab members, including Linfeng Liu, Ying Yu, Yichen Liu, Dr. Priti Soni, Dr. Manuela Trani, Dr. Joey Lichter, Dr. Pinar Lyidogan, Ashley Bagwell, Patrick Baldwin, and Hai Tran. They taught me plenty little tricks that are important for success in experiments.

Last but least, my special thanks goes to Dr. Yongfeng Li, a former student in Dr. Dennis Liotta's group, for a tremendous help in synthesis of nucleoside analogs.

Table of Contents

1 Introduction……..……1

1.1 DNA replication……1
1.2 Nucleoside analogs and nucleoside kinases……..2
1.3 Protein engineering of deoxyribonucleoside kinases……4
1.4 Library screening and fluorescent-active cell sorting……...6

2 Results and Discussion……...11

2.1 Synthesis of fluorescent 3-fluoro-2′-deoxyuridine (fFT)……11
2.2 Protein expression and purification……..12
2.3 Kinetic assay for 3-fluoro-2′-deoxyuridine (FT)……13
2.4 Screening 1st-round of mutagenesis library by FACS……14
2.5 Synthesis of 5-vinyl-2′-deoxyuridine and tetrazole……....15
2.6 Kinetic assay for 5-vinyl-2′-deoxyuridine……..16
2.7 Photoinducible 1, 3-dipolar cycloaddition……..17

3 Materials and Methods……...19
Experimental procedure……20

References……...29

List of Figures

Fig. 1 Incorporating nucleotide into elongating strand in DNA replication……...1
Fig. 2 Phosphorylation of nucleosides to nucleoside triphosphates……...2
Fig. 3 A few nucleoside analogs used in anti-viral and anti-cancer therapy……..3
Fig. 4 Structure of *Dm*-dNK……..…….5
Fig. 5 A few fluorescent nucleoside analogs……..7
Fig. 6 E. coli cells accumulated fT exhibit fluorescence……8
Fig. 7 Structure of ddT and FT……...8
Fig. 8 Steric clash between furano moiety and *Dm*-dNK side chains……..…10
Fig. 9 SDS-PAGE analysis of overexpression and purification of *Dm*-dNK……...12
Fig. 10 FACS result……...14
Fig. 11 Fluorometer-monitored 1, 3-dipolar cycloaddition……...18
Fig. 12 Structure of 5-allyl-2′-deoxyuridine……..19

Scheme 1 Synthesis of fFT……...11
Scheme 2 Reactions in the coupled enzyme kinetic assay……...13
Scheme 3 Synthesis of 5-vinyl-2′-deoxyuridine……...15
Scheme 4 Synthesis of tetrazole……...16
Scheme 5 Photoinducible 1, 3-dipolar cycloaddition……...17
Scheme 6 Mechanism of 1, 3-dipolar cycloaddition……18

List of Tables

Table 1 Catalytic efficiencies for deoxyribonucleoside kinases…….4
Table 2 Kinetic parameters of wt *Dm*-dNK towards FT……...……13
Table 3 Kinetic parameters of wt *Dm*-dNK towards 5-vinyl-2′-deoxyuridine…….17

1 Introduction

1.1 DNA replication

DNA is one of the essential biopolymers of life; it contains genetic information that is used to instruct and regulate the development and function of all organisms and some virus. According to the revised central dogma[1], DNA can be replicated directly by DNA polymerase[2, 3] using itself as template when transmit genetic information between parent and progeny. Some retrovirus use RNA as the storage of genetic information; the RNA genome is reverse-transcribed into DNA by reverse transcriptase. Subsequently, the DNA is integrated in host genome and replicated as part of the host's DNA[4]. During DNA replication, appropriate nucleotides donated by nucleoside triphosphates are added on to the free 3′-OH group of the elongating strand (Fig. 1). *In vivo*, nucleoside triphosphates can be synthesized from small molecule via the *de novo* biosynthetic pathway, or they come from phosphorylating nucleosides recycled from degradation in cells or food source (Fig. 2).

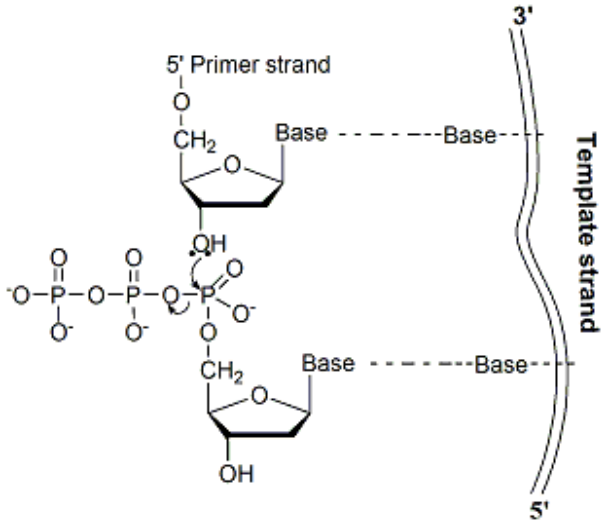


Fig. 1 Incorporating nucleotide into elongating strand in DNA replication. The 3′-hydroxy group on the elongating strand is attached to the α-phosphryl group of the oncoming nucleoside triphosphate.

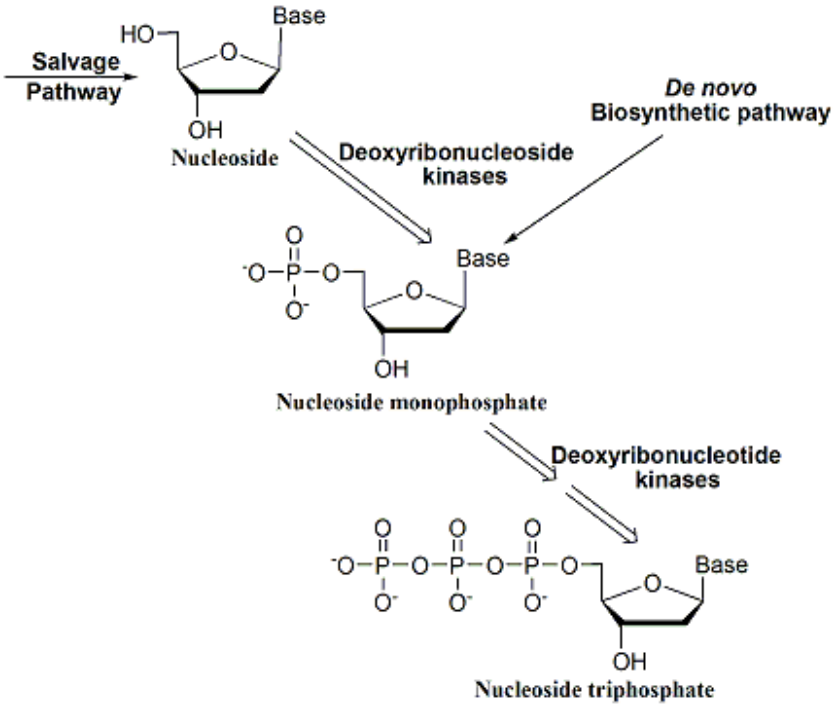


Fig. 2 Phosphorylation of nucleosides to nucleoside triphosphates by nucleoside and nucleotide kinases *in vivo*.

1.2 Nucleoside analogs and nucleoside kinases

According to World Health Organization (WHO), cancer is the leading cause of death around the world: in 2007 alone, it caused 7.9 million deaths[5]. Also HIV presents a tremendous problem to human health; there are 33 million people living with HIV worldwide, based on the summary of AIDS[6].

Cancer cells are characterized by their uncontrolled cell replication, due to alteration in genes that regulate cell growth and differentiation[7, 8]; while HIV virus infect normal T-cells and hijack the cells' machinery to replicate and assemble large amount of virus[9, 10]. Studies have shown that blocking DNA replication can stall cell proliferation and trigger

cell death[8, 11]. Therefore, nucleoside analogs (NA) are developed as DNA polymerase or HIV reverse transcriptase inhibitor to stall DNA replication/reverse transcription.

Nucleoside analogs are nucleosides modified in the base or sugar moiety, such as 3′-azido-3′-deoxythymidine (AZT), 2′,3′-didehydro-2′,3′-dideoxythymidine (d4T), 2′,3′-dideoxyinosine (ddI), and 5-fluoro-1-(2R,5S)-[2-(hydroxymethyl)-1,3-oxathiolan-5-yl]cytosine (FTC) (Fig. 3). A number of NAs are used in anticancer and antiviral therapies or are currently in clinic trials[11-13]. NAs are not biologically active until they

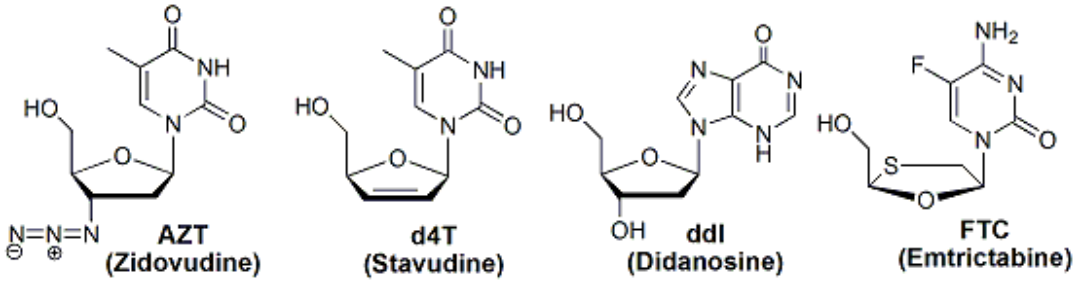


Fig. 3 A few nucleoside analogs used in anti-viral and anti-cancer therapy.

are phosphorylated to their active triphosphate form. Similar to natural nucleoside, NAs require the same set of nucleoside and nucleotide kinases to complete the phosphorylation process. The first step of activation is phosphorylation of NAs to nucleoside monophosphates by deoxyribonucleoside kinases, and this step is usually the rate-limiting step due to the high substrate specificity of human kinases toward natural nucleosides[14]. As a result, the kinetic property of nucleoside kinases is one of the factors that determine the sensitivity of cancer cells and virus to nucleoside analogs. Expression of exogenous or engineered nucleoside kinases in cancer cells enhances the cytotoxicity of nucleoside analogs [14-16]. However, the broad specificity of exogenous or engineered kinases may interfere with the tightly regulated pool of nucleosides and nucleotides *in*

vivo[17]. In addition, a large amount of nucleoside analogs failed preclinical tests, not because that they do not have anti-viral or anti-cancer ability but because natural kinases could not phosphorylate them into their active triphosphate anabolites[18, 19]. Engineering more efficient and specific kinases that prefer NAs to natural nucleosides is one promising approach for finding new and more potent nucleoside analogs.

1.3 Protein engineering of deoxyribonucleoside kinases

Several nucleoside kinases are under intense study. Engineered Herpes simplex virus type-1 thymidine kinase (HSV-1 TK) is being used in combine with the guanosine nucleoside analog ganciclovir (GCV) in suicide gene therapy[20, 21]. Evolving human thymidine kinase 2 (hTK2) by chimeragenesis yielded enzymes with different substrate specificity[22]. Among members of deoxyribonucleoside kinase family, *Drosophila melanogaster* deoxy-ribonucleoside kinase (*Dm*-dNK) is the only one with the unique ability to phosphorylate all four natural deoxyribonucleosides with high catalytic efficiency (Table 1.)[23, 24]. The broad substrate specificity and high catalytic rate of *Dm*-dNK renders it a promising candidate for protein engineering.

	dThd	dCyd	dAdo	dGuo
Dm-dNK	1.7×10^7	1.4×10^7	1.6×10^5	2.9×10^4
hTK1	8.0×10^6	n.d.	n.d.	n.d.
hTK2	1.9×10^4	1.1×10^4	n.d.	n.d.
hdCK	n.d.	7.3×10^4	2.6×10^3	2.7×10^3
hdGK	n.d.	n.d.	2.3×10^1	2.8×10^2
HSV-TK	3.8×10^5	n.d.	n.d.	n.d.

Table 1. Catalytic efficiencies $k_{cat}/K_m(M^{-1}s^{-1})$ for deoxyribonucleoside kinases.
Adapted from Ref. [22] n.d. = not detected.

The crystal structure of *Dm*-dNK was first reported by Johansson K. et al.[25] The

enzyme forms a homodimer, where each monomeric subunit contains a Rossman fold and assumes an α/β/α architecture with a multi-stranded β-sheet core. The dimer interface

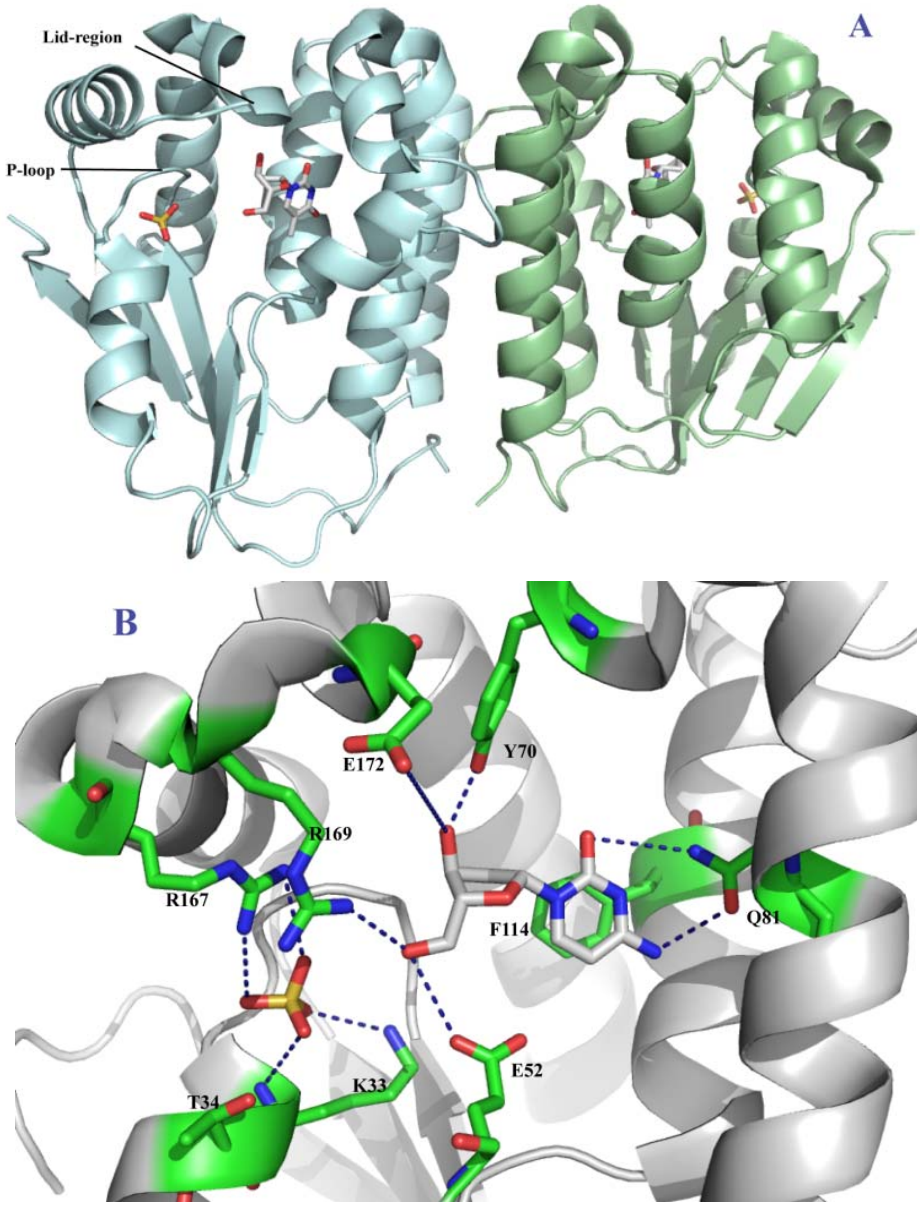


Fig. 4 Structure of *Dm*-dNK (1J90)[25]. A) The dimer with two-fold axis in the plane of the paper. B) Binding at active site. Substrate is shown as gray sticks while side chain residues are shown as green sticks.

consists of hydrophobic residues from two α-helices of each sub unit forming a four-helix bundle (Fig. 4A). There are two conserved sequence motif---the lid-region and the P-loop, which bind and position the phosphoryl donor and acceptor. In *Dm*-dNK, the P-loop binds the α and β phosphoryl group through the backbone of main chain and side chains of Lys33 and Thr34. In the phorsphoryl acceptor binding pocket,The Lid region contains conserved arginine Arg167 contributing to phosphoryl donor binding. The base portion of the nucleoside makes two hydrogen bonds with Gln81 and it stacks with Phe114 by π-interactions. Its deoxyribose moiety is anchored by two hydrogen bonds from Tyr70 and Glu172 to the 3′ hydroxyl group, as well as through a water molecule to Tyr179. The 5′ oxygen is hydrogen-bonded to Glu52 and Arg169 (Fig. 4B)[25]. Glu52 is suggested to be the catalytic residue that acting as a general base in the initial 5′-OH activation.

On the basis of wide-type *Dm*-dNK's ability to phosphorylate several NAs such as cytosine arabinoside (araC), Acyclovir (ACV), and 5-fluoro-2′-deoxyuridine (FdUrd) etc., direct evolution is applied to *Dm*-dNK, in order to obtain mutants with higher activity and specificity towards NAs[14, 26, 27].

1.4 Library screening and fluorescent-active cell sorting

Direct evolution is a potent method in protein engineering to evolve nucleoside kinases with desirable properties by mutagenesis techniques such as error-prone PCR and DNA shuffling[20-22]. The resulting combinatorial libraries are selected for thymidine kinase activity using genetic complementation in the auxotrophic *E.coli* KY895 host strain[28], and subsequently analyzed by screening for cytotoxicity to NAs. This method at best

selects for mutants with broader nucleoside kinase activity, which denying promising mutants with enhanced sensitivity and orthogonal NA kinase activity. Also, the TK activity of the mutants evolved by genetic complementation can potentially interrupt the balance of natural nucleotides and disrupt the tightly regulated 2′-deoxyribonucleoside metabolism in the cell.

Designing a screening technique that yields orthogonal NA kinases, which have little or no activity toward natural deoxyribonucleosides, could be extremely beneficial. A method adopting fluorescent-active cell sorting was developed to achieve high-throughput screening for orthogonal NA kinases in our lab. By a few chemical modifications, a fluorescent moiety can be introduced in the nucleoside analogs [29-31] (Fig. 5). These fluorescent NAs (fNAs) can easily be transported in and out cells while their phosphorylated forms are charged and hence will be trapped inside the cells, enabling the cells to 'glow' when excited with light at certain wavelength (Fig. 6). After transformation, bacteria cells containing plasmids that encode for individual mutants of the constructed library are incubated with fNAs. Functional kinase mutants expressed in host cells will allow the cells to accumulate fNA phosphate inside. The higher the

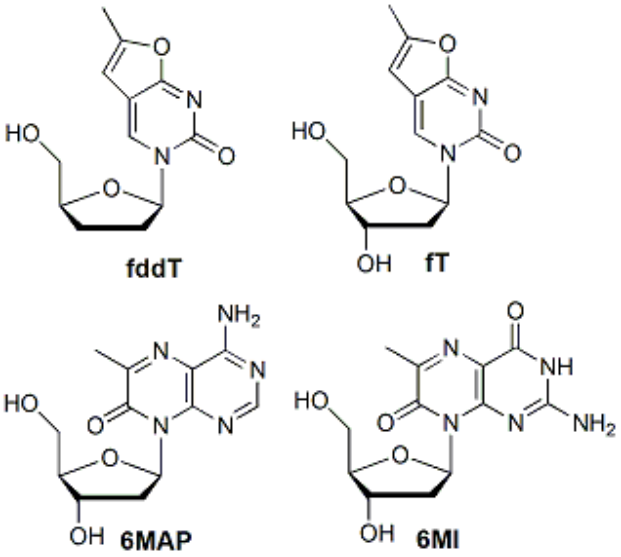


Fig. 5 A few fluorescent nucleoside analogs. Fluorescent moieties are the nucleo bases on the analogs.

enzyme activity is, the brighter the cells will glow. Based on the difference in fluorescent intensity between negative control and samples, cells which express mutants with high activity and specificity toward fNAs can be identified and selected from the pool and used for the next round of library construction. FACS not only offers a more efficient screening method, it also raises the opportunity for tailoring enzymes according to specific deoxyribonucleoside substrates. Nowadays, lots of NAs with potential high anti-viral or anti-cancer activity cannot be utilized because no kinase is able to convert them into their active triphosphate form. With the help of our fluorescent nucleoside analog phosphorylation screen (FLUPS), it is possible to evolve enzymes that can accommodate much bulkier modifications and give rise to more potent analogs in cancer or viral treatment.

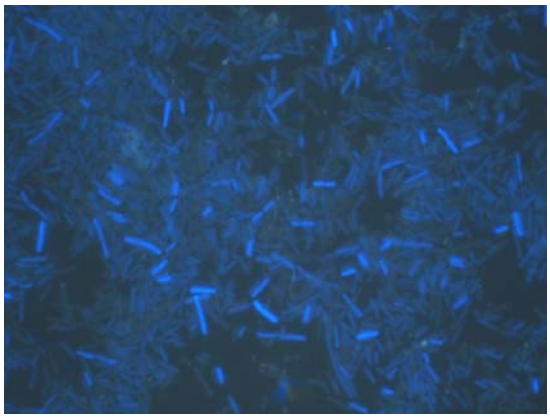


Fig. 6 E. *coli* cells accumulated fT exhibit fluorescence when excited under microscope. Fluorescent intensity varies due to different activity of mutant enzymes encoded in cells.

Previous work in our lab was done to evolve *Dm*-dNK with activity for 2′, 3′-dideoxythymidine (ddT) by FACS, using fluorescent ddT[32]. After four generation of mutagenesis, the mutants obtained exhibit >10,000-fold specificity change compared to wildtype. Structure studies of *Dm*-dNK showed that the 3′ hydroxyl group on the ribose moiety plays an important role in substrate binding (Fig. 4B). It anchors the nucleoside in the binding pocket by hydrogen-

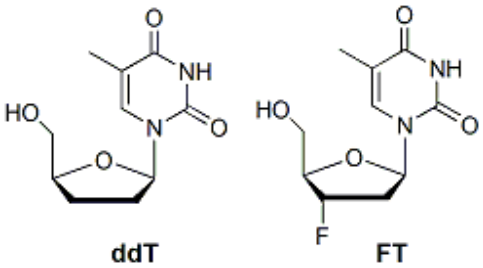


Fig. 7 Structure of ddT and FT

bonding to Glu172, Tyr70, and via a water molecule to Tyr179. 2′, 3′-dideoxythymidine (ddT) lacks any function group on its 3′ position, which may lower the binding affinity of the enzyme to a great extent. An additional fluorine atom on the 3′ position may compensate in part for the polar interaction between substrate and the enzyme. Thus, the *Dm*-dNK mutants after directed evolution should have higher activity and specificity for 3-fluoro-2′-deoxyuridine (FT) (Fig. 7). By constructing the fluorescent 3-fluoro-2′-deoxyuridine (fFT), we are able to evolve wildtype *Dm*-dNK via FACS.

However, the fluorescent furano moiety on the nucleic base is rigid and bulky which generated undesired mutation that is not beneficial to the actual nucleoside analog[32]. As shown in Fig. 8, the furano portion clashes with side chain residues Val84 and Met88 on *Dm*-dNK. In previous direct evolution of *Dm*-dNK with ddT, mutation T85M was found in active mutants, presumably responding to the steric clash and repositioning the helix with all three side chains. One solution is adding smaller and more flexible modification to nucleoside analogs. After the NAs being phosphorylated to their monophosphate form and trapped inside cells, we could then apply our agents to react with monophosphate nucleotides to tag them with a fluorophore. Song and Wang et al. discovered that tetrazole is capable of cycloaddition with alkene moieties of various electron-richness via photoclick chemistry to generate fluorescent pyrazolines[33]. Also, they demonstrated that the photoclick chemistry can be used to functionalizes alkene-containing protein in *E.* coli cells[34]. Developing photoclick chemistry between alkene-modified nucleotides and tetrazole is a promising method to reduce the risk of undesired mutations and obtain stronger fluorescent signal in cell sorting experiments.

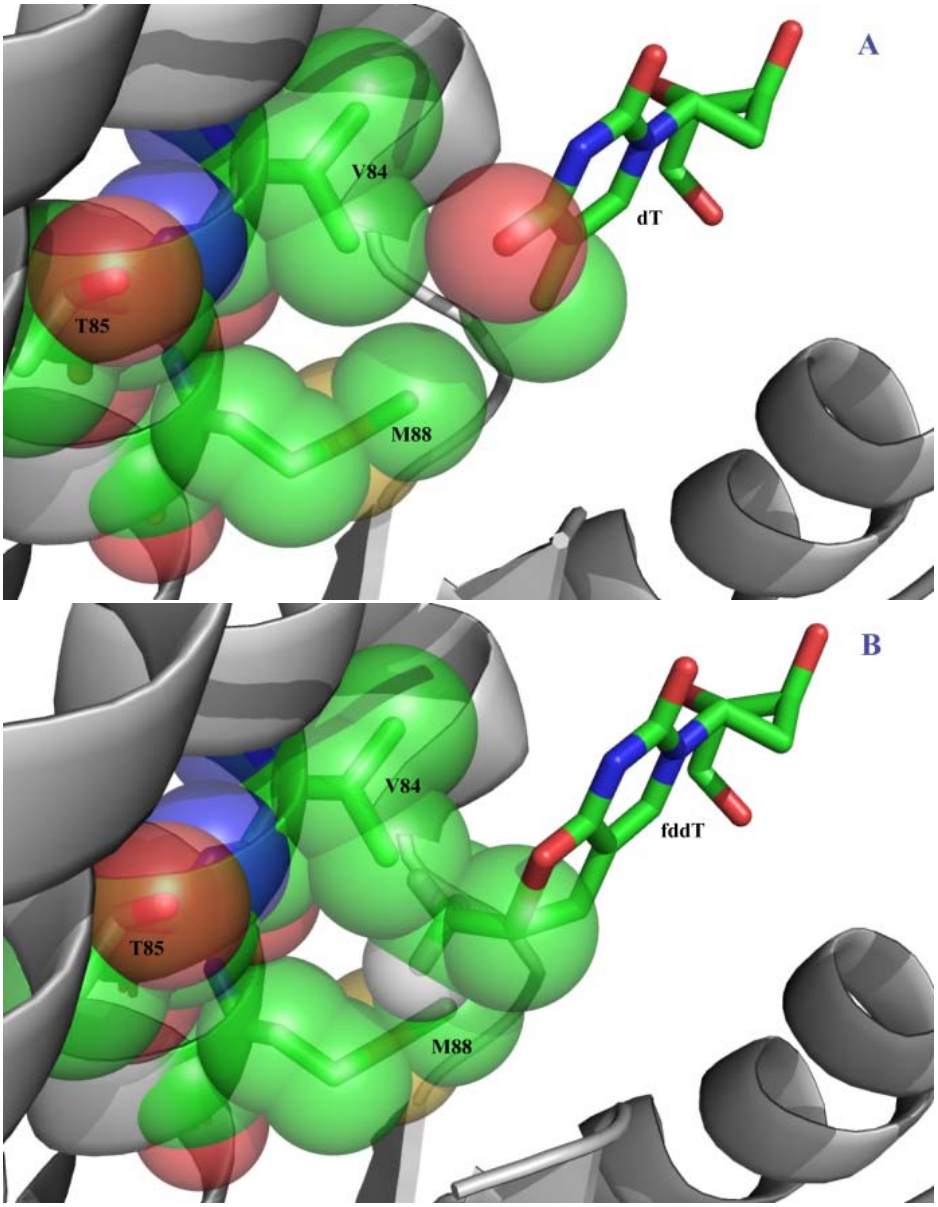
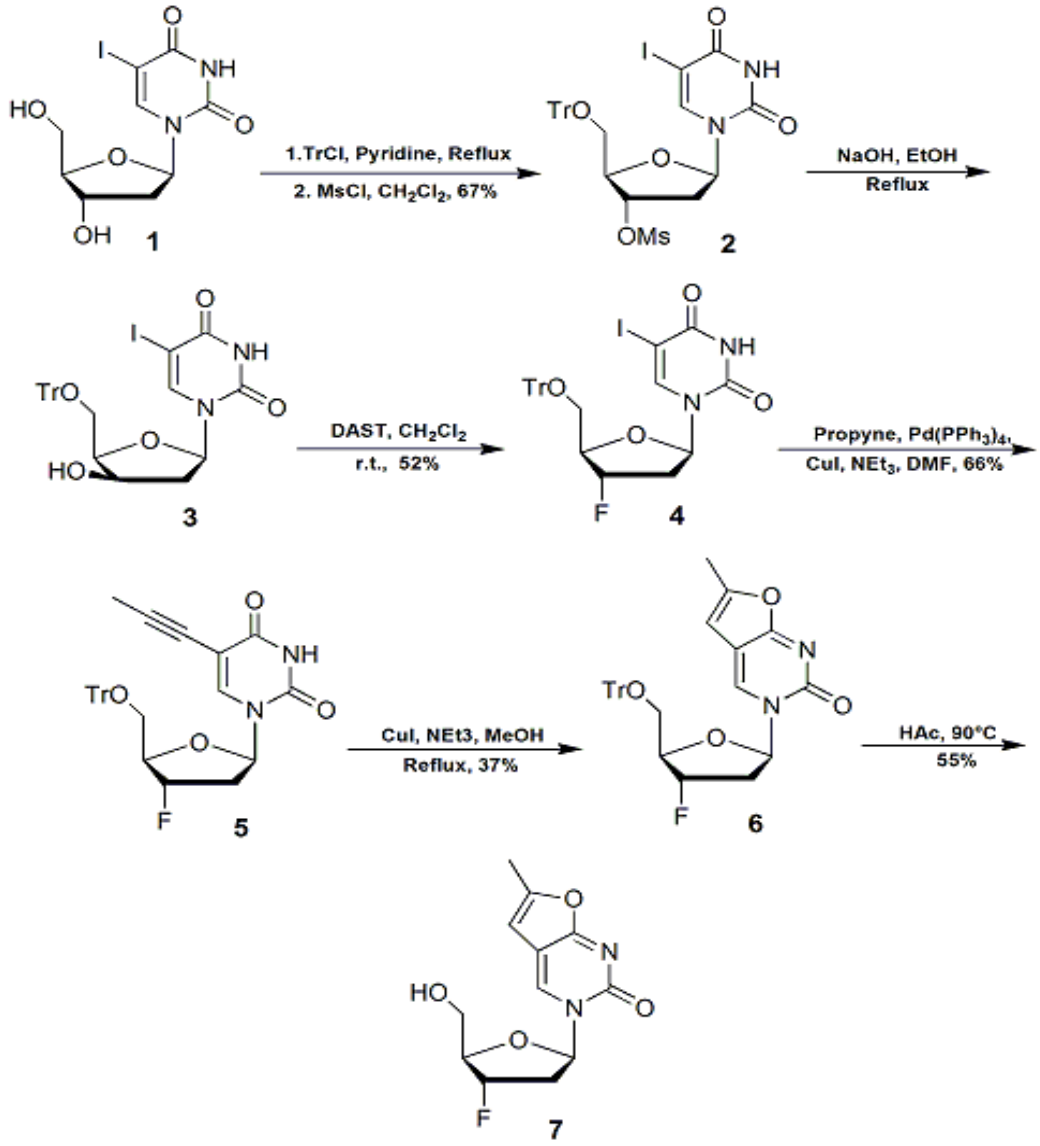


Fig. 8 Steric clash between furano moiety and *Dm*-dNK side chains. A) Crystal structure with dT as substrate. B) Modeled structure with fdT as substrate. The furano moiety was added to dT in crystal structure by PyMOL (DeLano Scientific, Palo Alto, CA, USA).

2 Results and Discussion

2.1 Synthesis of 3-fluoro-2′-deoxyuridine (fFT)



As shown in Scheme 1., Treatment of 5-iodo-2′-deoxyuridine with trityl chloride in pyridine followed by addition of methanesulphonyl chloride provide protected **2**[35].

Following the published procedure, the 3′ group was tautomerized to the xylo analog **3** by refluxing with sodium hydroxide in ethanol[36]. Treatment of **3** with DAST afforded fluorinated **4**, which was converted to **6** via Sonogashira reaction followed by cyclization in triethylamine/ methanol with copper(Ⅰ) iodide[30]. Compound **7** was obtained by deprotection in 80% aqueous acetic acid. The product was verified by nuclear magnet resonance (NMR). The overall yield of the synthesis was 10%.

2.2 Protein expression and purification

The gene for *Dm*-dNK was cloned into pMAL-c2X vector and expressed host *E.coli* strain TB1 (NEB, Ipswich, MA, USA). The pMAL vector introduces an MBP fusion protein which would bind to amylose resin. As a result, the fusion protein was purified by affinity chromatography on amylose resin column (Fig. 9). Then the protein was concentrated by Amicon-Ultra centrifugal filter unit (Amicon Bioseparations, Billercia, MA). The final concentration of Dm-dNK is 3.9×10^{-5} M.Sample aliquots were flash-frozen and stored at -80 ºC.

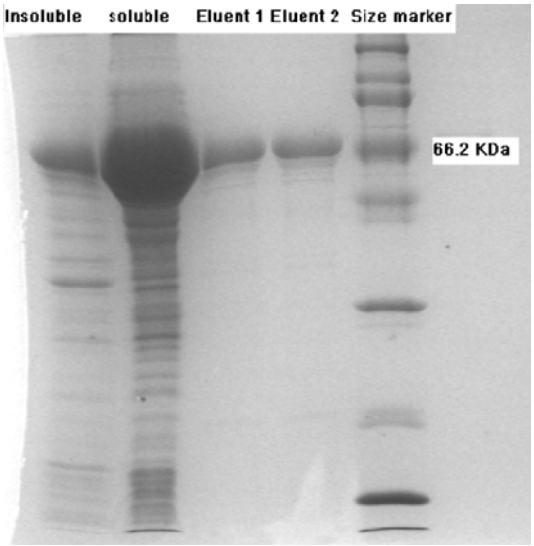
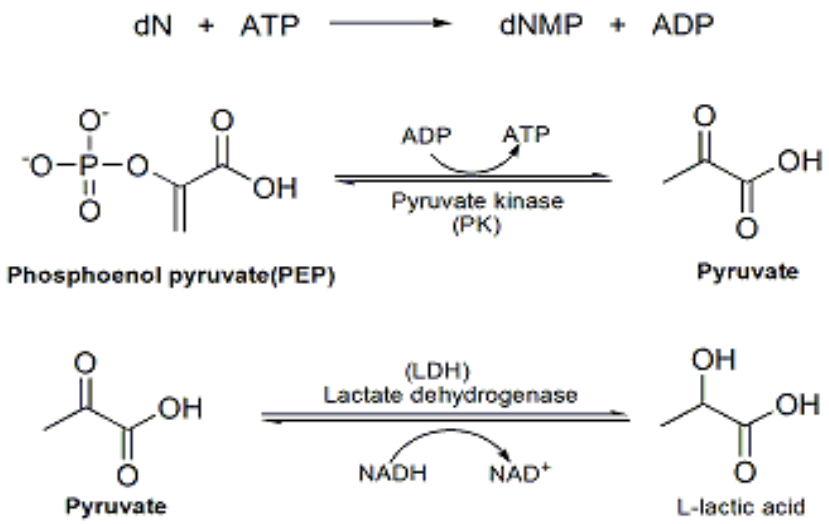


Fig. 9 SDS-PAGE analysis of overexpression and purification of wildtype Dm-dNK.

2.3 Kinetic assay for 3-fluoro-2′-deoxyuridine (FT)

In the coupled-enzyme spectrophotometric assay, transfer of the γ-phosphate group of ATP sets off a cascade reaction, in which NADH is oxidized to NAD^+[37]. The concentration change of NADH can be monitor by its decrease in absorbance at 340 nm. Since all reagents and products react in stoichiometric amounts, the measurement of [NADH] correspounds to the turnover of nucleosides (Scheme 2). Compared to thymidine, FT is a much worse substrate for wildtype *Dm*-dNK (Table 2).



Scheme 2. Reactions in the coupled enzyme kinetic assay.

Table 2. Kinetic parameters of wt *Dm*-dNK towards FT. Activity was determined at 37 °C, k_{cat} was calculated assuming the molecular weight of Dm-dNK is 69.3 KDa.

Substrate	V_{max}(μM•min^{-1})	Km (μM)	k_{cat} (s^{-1})	k_{cat} / Km(M•s^{-1})
Thymidine	1.82±0.21	2.27±0.86	7.79	3.43×10^6
FT	0.64±0.03	52.89±9.49	0.14	2.06×10^3

2.4 Screening 1st-round random mutagenesis library by FACS

Before actual library sorting, the cells enriched with fluorescent thymidine (fT) by wildtype *Dm*-dNK and human deoxycytidine kinase (hdCK) were sorted to ensure that the flow cytometer performs properly. Thymidine is not a substrate of hdCK (Table 1.), and as showed before, fT is phosphorylated by wildtype *Dm*-dNK. As a result, the intensity of the fluorescent signal in two types of cells should be apparently different.

For sorting the kinase library with fFT, *Dm*-dNK is used as negative control; since we are trying to identify mutants with increased activity for fFT compare to wildtype enzyme. Cells were incubated with 80 μM fFT for 2 h before sorting. As shown in Fig. 10, the two original cell cultures (*Dm*-dNK fFT & ep-Lib fFT) (ep-Lib stands for error prone PCR library, which is a random mutagenesis library) showed similar fluorescent intensity. The pool of cells after second round of sorting (ep-lib fFT sort 2) showed approximately 10 times higher fluorescent intensity over wildtype enzyme. The increase in fluorescent intensity after just one round of mutagenesis is much higher than the result obtained with substrate ddT. The result suggests the presence of additional functional group(s) that have the potential to help the enrichment of desirable mutants in cell sorting process. However, we

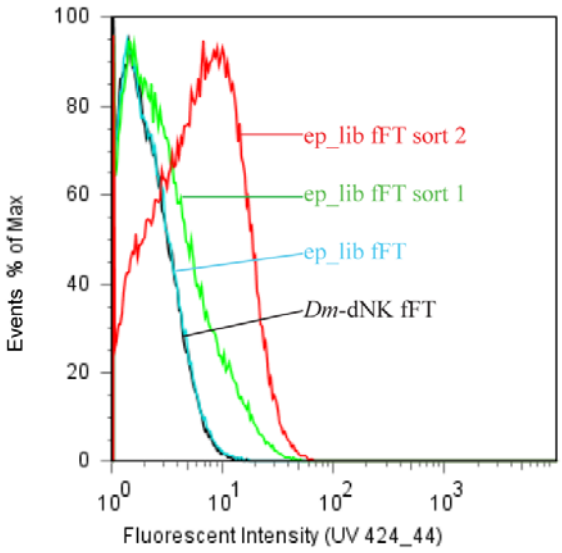
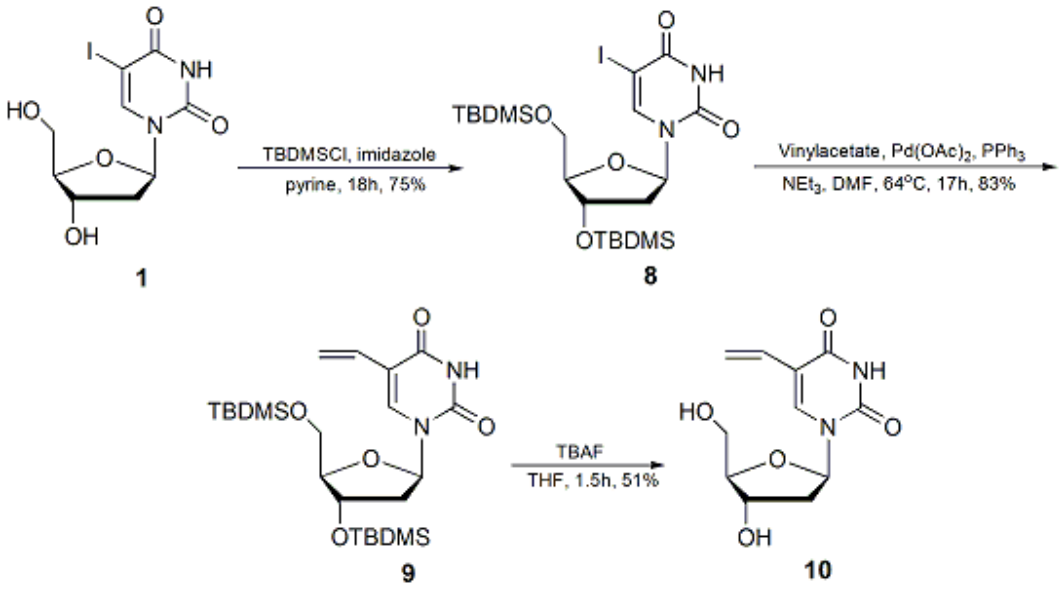


Fig. 10 FACS result. X-axis: normalized cell counts. Y-axis: fluorescence intensity.

did not carry on further experiments with FT due to its cytotoxicity *in vivo* via reductive pathways and fluorine elimination.

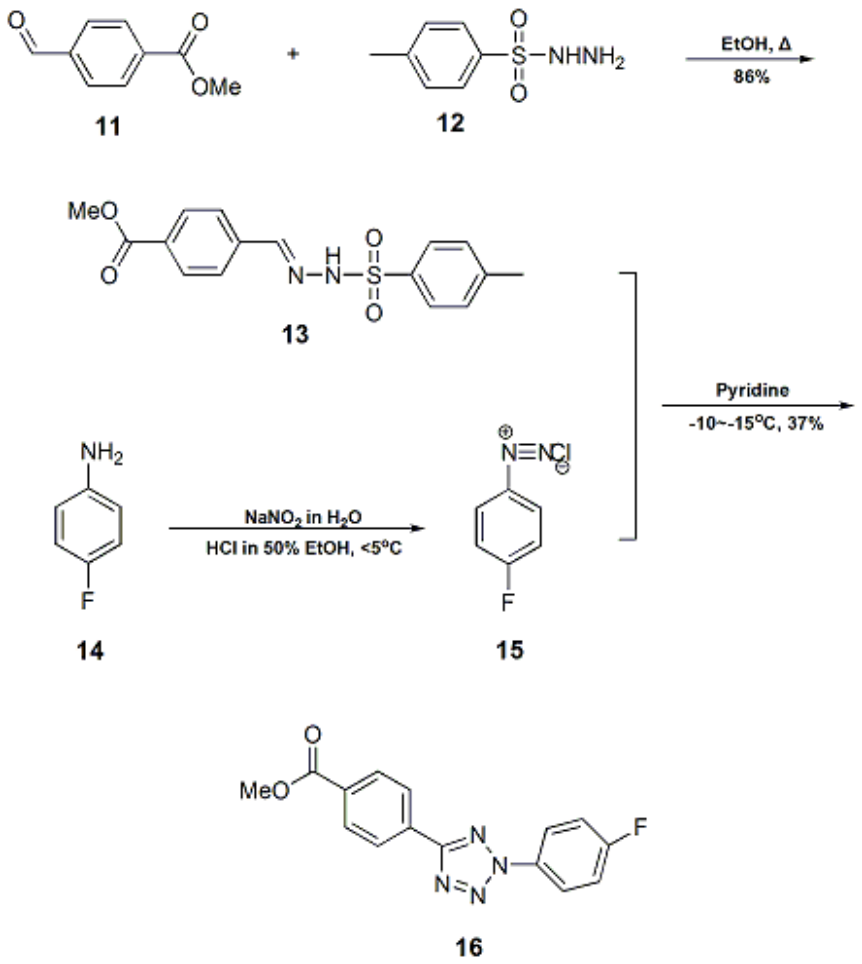
2.5 Synthesis of 5-vinyl-2′-deoxyuridine and tetrazole

5-vinyl-2′-deoxyuridine(**10**) was synthesized according to Scheme 3[38]. The 3′- and 5′-hydroxy groups of 5-iodo-2′-deoxyuridine (**1**) were protected by *tert*-butyldimethylsilyl chloride (TBDMSCl). Then the alkylation of compound **8** was catalyzed by palladium to afford compound **9**, which was deprotected to provide 5-vinyl-2′-deoxyuridine (**10**).



Scheme 3. Synthesis of 5-vinyl-2'-deoxyuridine.

Tetrazole was synthesized according to Scheme 4[39]. Methyl 4-formylbenzoate (**11**) was reacted with β-toluenesulfonyl hydrazide (**12**) in hot ethanol, forming hydrazone (**13**). Tetrazole (**16**) was obtained by reaction of hydrazone (**13**) and arenediazonium salt (**15**).



Scheme 4. Synthesis of tetrazole.

2.6 Kinetic assay for 5-vinyl-2′-deoxyuridine

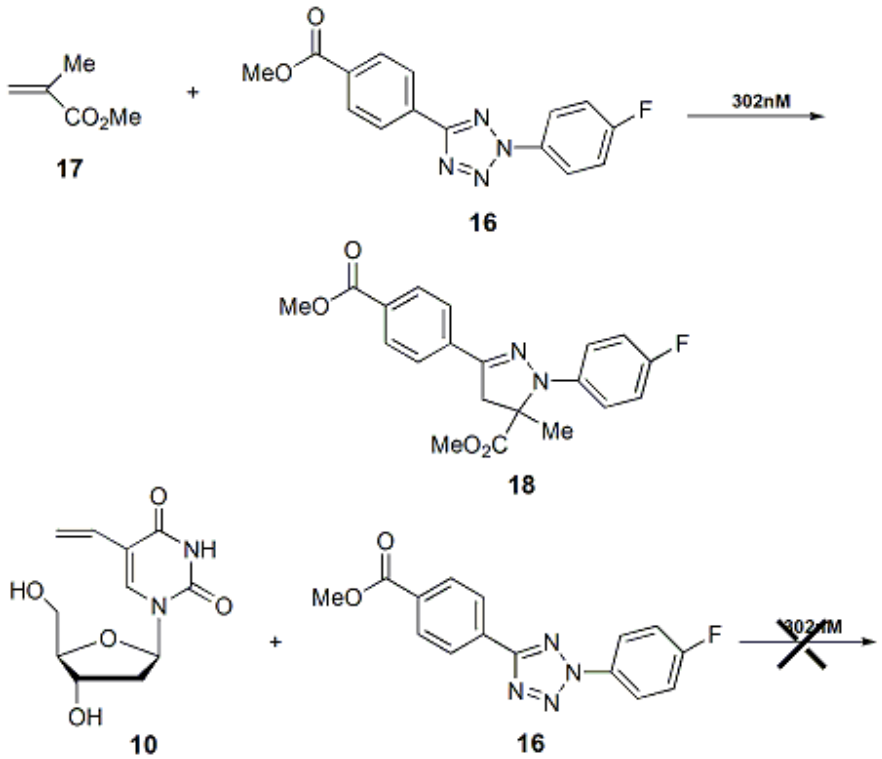
As described above, the coupled-enzyme spectrophotometric assay is used to measure kinetic character of wildtype *Dm*-dNK towards 5-vinyl-2′-deoxyuridine (Table 3). The flexibility and small steric hindrance of vinyl group on the 5-position does not significantly affect enzyme activity.

Table 3. Kinetic parameters of wt *Dm*-dNK towards 5-vinyl-2′-deoxyuridine. Activity was determined at 37°C, k_{cat} was calculated assuming the molecular weight of Dm-dNK is 69.3 KDa.

Substrate	V_{max}(μM•min^{-1})	Km (μM)	k_{cat} (s^{-1})	k_{cat} / Km(M•s^{-1})
Thymidine	1.82±0.21	2.27±0.86	7.79	3.43×10^{6}
5-Vinyl-2′-deoxyuridine	0.44±0.04	2.26±0.61	3.75	1.66×10^{6}

2.7 Photoinducible 1, 3-dipolar cycloaddition

Illumination of methyl methacrylate and tetrazole in water: acetonitrile (1:1, v/v) by 302nm UV light induced intense fluorescence in the reaction mixture, indicating pyrazoline adduct was produced by 1, 3-dipolar cycloaddition (Scheme 5). However, 5-



Scheme 5. Photoinducible 1, 3 -dipolar cycloaddition

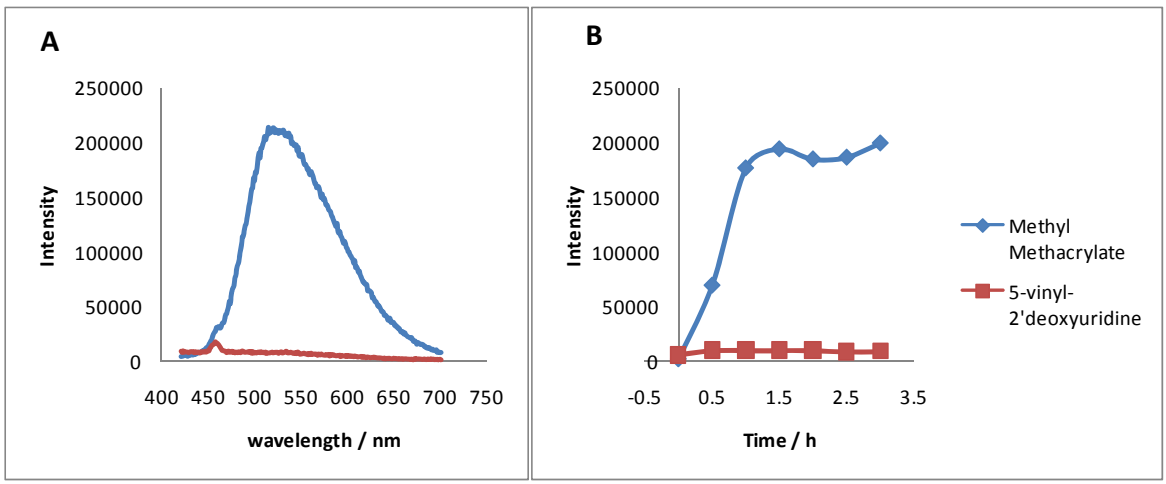
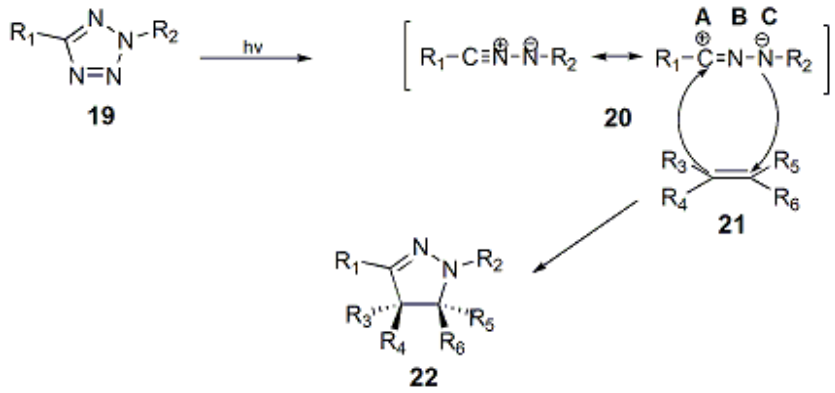


Fig. 11 Fluorometer-monitored 1,3-dipolar cycloaddition. A) The fluorescent spectra of two reactions at 3 h of excitation. B) The maxim fluorescent intensity of two reactions every half hour.

vinyl-2′-deoxyuridine could not react with tetrazole upon UV radiation (Fig. 11). During cycloaddition reaction (Scheme 6), nitrile imine (**20**) dipole is generated by photolysis of tetrazole precursor. In nitrile imine dipole, carbon A possesses an electron sextet with positive charge while nitrogen C is negative charged with unshared electron pair. When combined with a double bond system, the reaction leads to an uncharged 5-membered ring (**22**)[40].



Scheme 6. Mechanism of 1,3-dipolar cycloaddition

Electron-deficient alkenes facilitates electron transfer, as a result, they generally have higher yields[33]. The vinyl group on 5-vinyl-2′-deoxyuridine is conjugated with the uracil base, which hindered the reaction from proceeding. One possible solution is to add one extra carbon between alkene group and nucleic base, which is 5-allyl-2′-deoxyuridine (Fig. 12). The synthesis and characterization of 5-allyl-2′-deoxyuridine is in progress in our lab.

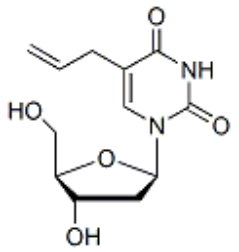


Fig. 12 Structure of 5-allyl-2′-deoxyuridine

3 Materials and Methods

Solvents and chemicals were commercially available from Aldrich-Sigma (St. Louis, MO, USA) and are of the highest available quality unless otherwise stated. 6-Methyl-3-(β-D-2-deoxyribofuranosyl)furano-[2,3-*d*]pyrimidin-2-one (fT) was purchased from Berry & Associates (Dexter, MI, USA). *Pfu* Turbo DNA polymerase was purchased from Stratagene (La Jolla, CA, USA). Pyruvate kinase and lactate dehydrogenase were from Roche Biochemicals (Indianapolis, IN, USA). Restriction enzymes were purchased from New England Biolabs (Ipswich, MA, USA). Plasmid DNA was isolated by using the QiaPrep Spin MiniPrep kit and PCR products were purified with the QiaQuick PCR purification kit (Qiagen, Valencia, CA, USA).

Instrumentation Proton NMR spectra were recorded on Varian INOVA at 400 MHz and Mercury at 300 MHz. Carbon-13 NMR spectra were recorded on Varian INOVA at 400 MHz. Chemical shifts were reported in ppm using deuterated solvents as

internal standards. Muliplicity was reported as follows: s = singlet, d = doublet, t = triplet, q = quartet, m= multiplet. All mass spectra were obtained from the Emory University Mass Spectrometry Center (Dr. F.Strobel) on JEOL JMS-SX102/SX102A/E. UV spectra were measured on a Varian Cary 100 Bio Spectrophotometer (Varian, Walnut Creek, CA). Flash chromatography was performed using Purasil 60Å silica gel (230-400 mesh). Fluorescence spectra were measured on FluoroMax-3 spectrofluorometer (HORIBA Jobin Yvon, Edison, NJ, USA).

Experimental procedure

5-iodo-5'-O-(triphenylmethyl)-3'-methanesulfonate-2'-deoxyuridine (2) A mixture of 5-iodo-2′-deoxyuridine (**1**) (1.0 g, 2.82 mmol) and trityl choloride (0.95 g, 3.39 mmol) in anhydrous pyridine (20 mL) was refluxed for 2.5 h. After 2.5 h methanesulphonyl chloride (0.68 mL, 8.5 mmol) was added to the ice-cooled solution. The reaction was carried out at room temperature overnight. Then the reaction was quenched with 1 mL water, and the solution was poured into 400 mL ice water. The mixture was extracted with 30 mL×4 dichloromethane, combine the organic layer. Then the organic layer was washed by saturated sodium chloride (120 mL), dried over magnesium sulfate, and evaporated *in vacuo*. The residue was purified by flash chromatography (gradient eluent: hexane/ethyl acetate (3:1 ~ 1:2). Yielded compound **2** 1.27 g (67%). ^{1}HNMR (400 MHz, $CDCl_3$) δ 2.45 (m, 1H, H2′), 2.77 (m, 1H, H2′), 3.03 (s, 1H, CH_3SO_2), 3.48 (m, 2H, H5′), 4.35 (m,1H, H4′), 5.34 (m, 1H, H3′), 6.33 (dd, 1H, J=5.2 and 8.8 Hz, H1′), 7.28-7.46 (m, 15H, trityl), 8.19 (s, 1H, 6H).

1-[2′-deoxy-5′-O-(triphenylmethyl)-D-threo-pentofuranosyl]-5-iodo-2,4(1H,3H)-pyrimidinedione (3) Compound **3** (1.25 g, 1.85 mmol) was dissolved in 125 mL ethanol. One equivalent sodium hydroxide (1.85 mmol, 0.07 g NaOH in 8.34 mL water) was added to the solution and refluxed for 1.5 h. Then another one equivalent of sodium hydroxide was added to the reaction, refluxed for 1 h. After cooling to room temperature, 2 M HCl was used to neutralize the reaction mixture and the solution was evaporated *in vacuo*. The residue was purified by flash chromatography (gradient eluent: hexane/ethyl acetate (1:1 ~ 1:2). Yielded compound **3** 0.838 g (74%). ^{1}HNMR (400 MHz, MeSO-d_6) δ (2′H under MeSO), 3.17 (m, 1H, H5′), 3.30 (m, 1H, H5′), 4.11 (m,1H, H4′), 4.17 (m, 1H, H3′), 5.29 (d, 1H, J=3.2 Hz, OH), 6.10 (d, 1H, J=8.0 Hz, H1′), 7.25-7.43 (m, 15H, trityl), 8.18 (s, 1H, 6H).

1-(2′,3′-dideoxy-3′-floro-5′-O-(triphenylmethyl)-β-L-ribofuranosyl)-5-iodopyrimidine-2,4(1H,3H)-dione (4) Compound **3** (0.95 g, 1.59 mmol) was dissolved in dry dichloromethane (55 mL), and DAST (0.38 mL, 2.9 mmol) was added. The solution was stirred at room temperature for 1h, and then diluted with dichloromethane (15 mL). Use saturated sodium bicarbonate solution (100 mL) and water (50 mL) to wash; the organic phase was separated, dried with sodium sulfate, and evaporated *in vacuo*. The residue was purified by flash chromatography (initial elute hexane/ethyl acetate (2:1), followed by hexane/ethyl acetate (1.5:1)). Compound **4** was isolated as a foam (0.49 g, 52%). ^{1}HNMR (400 MHz, $CDCl_3$) δ 2.32 (m, 1H, H2′), 2.77 (m, 1H, H2′), 3.45 (m, 2H, H5′), 4.36 (d, 1H, $J_{4'\text{-}H}$=28 Hz, H4′), 5.28 (dd, 1H, $J_{3'\text{-}F}$=53.6, H3′), 6.38 (dd, 1H, J=4.8 and 9.4 Hz, H1′), 7.28-7.41 (m, trityl), 8.19 (s, 1H, 6H).

^{13}CNMR (400 MHz, $CDCl_3$) δ 39.7 (d, C2′, $J_{C2'-F}$=21.36 Hz), 63.5 (d, C5′, $J_{C5'-F}$=10.68 Hz), 69.1 (C5), 84.7 (d, C4′, $J_{C4'-F}$=25.18 Hz), 85.5 (C1′), 88.1 (quaternary trityl), 94.5 (d, C3′, $J_{C3'-F}$=177.77 Hz), 127.7-128.7 (aromatic trityl), 143.2 (C6), 150.0 (C2), 160.0 (C4). ^{19}FNMR confirmed the existence of fluorine atom (173.6). M-H^+, $C_{28}H_{23}O_4N_2F_1{}^{127}I_1$, 597.07648; found: 597.07056.

1-(2′,3′-dideoxy-3′-fluoro-5′-O-(triphenylmethyl)-β-L-ribofuranosyl)-5-(1-propynyl) pyrimidine-2,4(1H,3H)-dione (5) Compound **4** (0.49 g, 0.83 mmol) was dissolved in anhydrous DMF (10 mL) in a three neck flask. The solution was deoxygenated under argon for 1.5h. Tetrakis(triphenylphophine) palladium ($Pd(PPh_3)_4$, 0.10 g. 0.083 mmol) and copper(I) iodide (CuI, 0.03 g, 0.16 mmol) were added. Two equivalent anhydrous triethylamine was injected and the flask was filled with propyne (~1.5 mL, 25.5 mL). The reaction mixture was protected from light and stirred 22h. The mixture was quenched with water (3 mL) and 50 mL water was added afterwards. Ethyl acetate (20 mL×4) was used to perform extraction. The organic layer was washed with water (50 mL), dried (Na_2SO_4), and evaporated *in vacuo*. The residue was purified with flash chromatography (initial elute: hexane/ethyl acetate (3:1), followed by hexane/ethyl acetate (2:1 and 1.5:1). Afforded compound **5** 283 mg (66%). ^{1}HNMR (400 MHz, $CDCl_3$) δ 1.68 (s, 3H, CH_3), 2.33 (m, 1H, H2'), 2.78 (m, 1H, H2′), 3.43 (d, 2H, H5′), 4.38 (d, 1H, $J_{4'-H}$=27.6 Hz, H4′), 5.31 (dd, 1H, $J_{3'-F}$=53.6 Hz, H3′), 6.39 (dd, 1H, J=5.2 and 9.2 Hz, H1′), 7.28-7.46 (m, trityl), 8.03 (s, 1H, H6). ^{13}CNMR (400 MHz, $CDCl_3$) δ 4.7 (CH_3), 39.7 (d, C2′, $J_{C2'-F}$=21.36 Hz), 63.7 (d, C5′, $J_{C5'-F}$=10.68 Hz), 69.9 (propynyl C1″), 84.7 (d, C4′, $J_{C4'-F}$=25.94 Hz), 85.5 (C1′), 88.2 (quaternary trityl), 91.2 (propynyl C2″), 94.6 (d,

C3′, $J_{C3'-F}$=177.77 Hz), 101.6 (C5), 127.7-128.7 (aromatic trityl), 143.6 (C6), 149.2 (C2), 161.7 (C4). M+H^+, $C_{31}H2_8O_4N_2F_1$, 511.19549; found: 511.20226.

3-(2′,3′-dideoxy-3′-fluoro-5′-O-(triphenylmethyl)-β-L-ribofuranosyl)-6-methylfuro [2,3-d]pyrimidine-2(3H)-one (6) Compound **5** (0.27 g, 0.53 mmol) was dissolved in anhydrous methanol (70 mL) and triethylamine (30 mL) (7:3). Copper(I) iodide (CuI, 0.02 g, 0.11 mmol) was added to the solution and the solution was refluxed for 4h. The solvent was removed *in vacuo*, and the residue was purified by flash chromatography (initial elute: hexane/ethyl acetate (2:1), gradually enlarge solvent polarity to hexane/ethyl acetate (1:1.5). Yielded compound **6** 0.10 g (37%). ^{1}HNMR (400 MHz, $CDCl_3$) δ 2.30 (s, 3H, CH_3), 2.23 (m, 1H, H2′), 3.16 (m, 1H, H2′), 4.49 (d, 1H, $J_{4'-H}$=24.0 Hz, H4′), 5.29 (dd, 1H, $J_{3'-F}$=53.6 Hz, H3′), 5.80 (d, 1H, J=1.2 Hz, H5), 6.47 (dd, 1H, J=5.6 and 9.6 Hz, H1′), 7.26-7.37 (aromatic trityl), 8.44(s, 1H, H4). ^{13}CNMR (400 MHz, $CDCl_3$) δ 14.3 (CH_3), 40.8 (d, C2′, $J_{C2'-F}$=21.36 Hz), 63.4 (d, C5′, $J_{C5'-F}$=9.92 Hz), 85.1 (d, C4′, $J_{C4'-F}$=25.94 Hz), 88.0 (quaternary trityl), 88.4 (C1′), 94.1 (d, C3′, $J_{C3'-F}$=179.40 Hz), 99.5 (C5), 108.0 (C4a), 127.6-128.6 (aromatic trityl), 143.3 (C4), 154.9 (C6), 156.0(C2), 172.2 (C7a). M+Na^+, 533.18471; found: 533.18481.

3-(2′,3′-dideoxy-3′-fluoro-5′-hydroxy-β-L-ribofuranosyl)-6-methylfuro[2,3-d] pyrimidin-2(3H) -one (7) Compound **6** (0.10 g, 0.20 mmol) was dissolved in acetic acid (HAc/H_2O (8:2)). The solution was heated to 90ºC for 30min. The solvent was dried *in vacuo*, and the residue was purified by recrystalization. Afforded compound **7** 0.03 g (55%). ^{1}HNMR (400 MHz, MeSO-d_6) δ 2.31 (s, 3H, CH_3), 2.21 (m, 1H, H2′), 2.73 (m,

1H, H2′), 3.66 (m, 2H, H5′), 4.34 (d, 1H, $J_{4'\text{-}F}$=26. 8 Hz, H4′), 5.35 (dm, 1H, $J_{3'\text{-}F}$=56.4 Hz, H3′), 6.22 (m, 1H, H1′), 6.44 (s, 1H, H5), 8.58 (s, 1H, H4). ^{13}CNMR (400 MHz, $CDCl_3$/CD_3OD) δ 14.0 (CH_3), 41.4 (d, C2′, $J_{C2'\text{-}F}$=20.60 Hz), 62.6 (d, C5′, $J_{C5'\text{-}F}$=11.54 Hz), 88.3 (d, C4′, $J_{C4'\text{-}F}$=23.65 Hz), 90.1 (C1′), 95.9 (d, C3′, $J_{C3'\text{-}F}$=176.24 Hz), 101.3 (C5), 110.0 (C4a), 156.8 (C6), 157.7 (C2), 173.4 (C7a). $M+H^+$, $C_{12}H_{14}O_4N_2F_1$, 269.09321; found: 269.09271.

3', 5'-bis-O-tert- butyldimethylsilyl-5-iodo-2'-deoxyuridine (8) Imidazole (0.58 g, 8.58 mmol) and t-butyldimethylsilyl chloride (1.29 g, 8.54 mmol) were added to the solution of 5-iodo-2'-deoxyuridine (1.02 g, 2.90 mmol) dissolved in dry pyridine (15 mL). The solution was stirred at room temperature under argon atmosphere for 18h. After evaporation *in vacuo*, the residue was extracted with ethyl acetate (40 mL × 3) and water (40 mL). The organic layer was collected, dried over sodium sulfate, filtered, and dried *in vacuo*. The crude product was purified by silica gel column chromatography (hexane/EtOAc, 4:1). Yielded compound **8** 1.23 g (75%). ^{1}H-NMR (400 MHz, $CDCl_3$) δ 8.10 (s, 1H, H-6), 6.28 (dd, J=4.8 Hz, J=8.4 Hz, 1H, H1′), 4.42-4.39 (m, 1H, H3′), 4.00-3.99 (m, 1H, H4′), 3.90 (dd, J=1.2 Hz, J=11.4 Hz, 1H, H5′), 3.77 (dd, J=2.4 Hz, J=11.4 Hz, 1H, H5′), 2.34-2.28 (m, 1H, H2′), 2.04-1.97 (m, 1H, H2′), 0.96 (s, 9H, t-Bu), 0.90 (s, 9H, t-Bu), 0.17-0.09 (m, 12H, CH3-Si×4).

3', 5'-bis-O-tert-butyldimethylsilyl-5-vinyl-2'-deoxyuridine (9) Palladium (II) acetate (18 mg, 0.078 mmol), triphenylphosphine (36 mg, 0.135 mmol), and anhydrous triethylamine (0.92 mL, 6.6 mmol) were combined in anhydrous dimethylformamide

(4.21 mL) and stirred at 64 °C until intense red color developed. Compound **8** (0.51 g, 8.58 mmol) and vinyl acetate (4.21 mL, 45.7 mmol) dissolved in anhydrous dimethylformamide (2.7 mL) were then added, and stirred at 64°C under argon atmosphere for 17 h. The reaction mixture was filtered to remove the resulting precipitate, and the filtrate was evaporated to dryness *in vacuo.* The crude product was purified by silica gel column chromatography (hexane/EtOAc, 4:1). Afforded compound **9** 0.34 g (83%). ^{1}H-NMR (400 MHz, $CDCl_3$) δ 7.68 (s, 1H, H-6), 6.40-6.33 (dd, J=11.6 Hz, J=17.8 Hz, 1H, C<u>H</u>=CH_2), 6.33-6.30 (m, 1H, H1′), 6.02 (dd, J=1.6 Hz, J=17.8 Hz, 1H, vinyl-*cis*), 5.25 (dd, J=1.6 Hz, J=11.4 Hz, 1H, vinyl-*trans*), 4.43-4.40 (m, 1H, H3′), 3.99-3.97 (m, 1H, H4′), 3.88 (dd, 1H, J=2.8 Hz, J=11.4 Hz, H5′), 3.78 (dd, J=2.4 Hz, J=11.2 Hz, 1H, H5′), 2.35-2.29 (m, 1H, H2′), 2.18-1.99 (m, 1H, H2′), 0.92-0.86 (m, 18H, t-Bu), 0.12-0.05 (m, 12H, CH3-Si×4).

5-vinyl-2'-deoxyuridine (10) To a tetrahydrofuran solution (2.06 mL of 1M THF solution, 2.06 mmol) of compound **9** (0.33 g, 0.687 mmol) was added tetrabutylammonium fluoride and the solution was stirred at room temperature for 1.5h. The solvent was removed under reduced pressure and the residue was purified by silica gel column chromatography ($CHCl_3$/MeOH, 8:1). Afforded 5-vinyl-2'-deoxyuridine (**10**) 0.09 g (51%). ^{1}H-NMR (400 MHz, DMSO-d_6) δ 11.41 (s, 1H, NH) 8.12 (s, 1H, H-6), 6.37 (dd, J=11.6 Hz, J=17.6 Hz, 1H, C<u>H</u>=CH_2), 6.16 (t, J=5.6 Hz, 1H, H1′), 5.92 (dd, J=2 Hz, J=17.6 Hz, 1H, vinyl-*cis*), 5.25 (d, J=4.4 Hz, 1H, vinyl-*trans*), 5.14-5.10 (m, 2H, OH-3′, OH-5′), 4.27-4.25 (m, 1H, H3′), 3.80-3.78 (m, 1H, H4′), 3.63-3.58 (m, 2H, H5′), 2.16-2.08 (m, 2H, H2′). ^{13}CNMR (400 MHz, DMSO-d_6) δ 162.08 (C-4), 149.57 (C-2), 137.88

(C-6), 128.78 (<u>C</u>H=CH_2), 113.92 (CH=<u>C</u>H_2), 110.80 (C-5), 87.45 (C-1′), 84.39 (C4′), 70.03 (C3′), 60.97 (C5′) ~40 (C2′, covered in DMSO peak).). FTMS (ESI) calculated for $C_{11}H_{15}O_5N_2$ 255.09755 [$M+H^+$], found 255.09766.

Methyl 4-((2-((4-methylphenyl) sulfonyl) hydrazinylidene) methyl)benzoate (13) Toluenesulfonyl hydrazide (**12**) (0.56 g, 3.037 mmol) was dissolved in 20mL ethanol, then a hot solution of methyl 4-formylbenzoate (**11**) (0.50 g, 3.037 mmol) in 1 mL ethanol was added. The solution was stirred at 70 °C for 30 min. After cooling down, 25 mL water was added to the mixture to facilitate precipitation. After filtering the mixture, drying the precipitate to dryness, afforded compound **13** 0.90 g (86%). ^{1}H-NMR (400 MHz, $CDCl_3$) δ 8.03 (d, J=8.8 Hz, 2H, H2, H6), 7.89 (d, J=8.4 Hz, 2H, H3, H5), 7.77 (s, 1H, -C<u>H</u>=N-), 7.65 (d, J=8.4 Hz, 2H, H2′, H6′), 7.34 (d, J=8 Hz, 2H, H5′, H3′), 3.93 (s, 3H, -OC<u>H</u>$_3$), 2.42(s, 3H, -C<u>H</u>$_3$).

Methyl 4-(2-(4-fluorophenyl)-2H-tetrazol-5-yl)benzoate (16) An ice-cooled solution of sodium nitrite (42 mg, 0.602 mmol) in 0.24 mL water was added to a solution of 4-fluoroaniline (57 mL, 0.601 mmol) in 0.96 mL EtOH/H_2O (1:1) containing 0.16 mL concentrated HCl to obtain arenediazonium salt (**15**). Then add dropwise arenediazonium salt solution to a solution of compound **13** (200 mg, 0.601 mmol) in pyridine (10 mL) at -10 to -15 °C. Solvent was evaporated *in vacuo*, and the residue was purified by silica gel chromatography (Hexane/EtOAc, 6:1). Afforded compound **16** 67 mg (37%). ^{1}H-NMR (400 MHz, $CDCl_3$) δ 8.33 (d, J=8.4 Hz, 2H, H2, H6), 8.26-8.18 (m, 4H, H3, H5, H2′,

H6′), 7.60-7.55 (m, 2H, H5′, H3′), 3.91 (s, 3H, -OC$\underline{H}_3$). FTMS (ESI) calculated for $C_{15}H_{12}O_2N_4F_1$ 299.09388 [M+H^+], found 299.09399.

Photoinducible 1, 3-dipolar cycloaddition Dissolve 2 μL methyl methacrylate (**17**) or 5-vinyl-2'-deoxyuridine (**10**) (40 mM in DMSO) and 8 μL tetrazole (**16**) (10 mM in DMSO) in 2 mL H_2O: acetonitrile (1:1, v/v). The reaction mixture was irradiated with a Pen-Ray mercury lamp 90-0012-01 (11SC-1) (Ultra-Violet Products, CA, USA) for 3 h in a 1cm quartz cuvette at room temperature (the light was filtered with N-WG-295 long pass filter (Edmund Optics, NJ, USA)). The mixture was excited at 395nm and scanned in the region of 420 to 700 nm thought a 2 nm slit[34] every half hour in our fluorometer (FluoroMax-3®, Jobin Yvon INC. NJ).

Random mutagenesis library construction A random mutagenesis library of wildtype *Dm*-dNK was created with the GeneMorph II kit (Stratagene, La Jolla, CA, USA). The mutation frequency averaged 2 to 3 mutation per gene (determined by DNA sequencing). The PCR products were cloned into pBAD-HisA (Invitrogen, Carlsbad, CA, USA) via *Nco*I and *Hind*III restriction sites and electroporated into E. *coli* TOP10 (Invitrogen, Carlsbad, CA, USA).

Library screening by fluorescent-activated cell sorting Library screening was performed by inoculating 2 mL LB media supplemented with ampicillin (50 μg/mL) at 37 °C until OD (600nm) reached 0.4~0.5. The protein expression was induced with arabinose (0.2%) and incubated for 4 h before the cell culture was mixed with 80 μM fFT

and incubated for another 2 h at 37 ºC. Then, the cells were centrifuged and the pellet was washed three times with the PBS buffer (pH 7.4) before being suspended in the PBS buffer to ~1×10^8 cells/mL. Cell sorting was performed on a FACSVantage flow cytometer (Becton Dickinson, Franklin Lakes, NJ, USA). The event detection was set to forward and side scattering. Sorting was performed at <2000 events/s with excitation by a UV laser (351 - 364 nm) and emission detection through a band pass filter (424 ± 20 nm). After three rounds of sorting, cells were collected in SOC medium and inoculated for 2h, then plated on LB-agar for harvesting cells and DNA sequencing.

Protein expression and purification The *Dm*-dNK gene was cloned into pMAL-c2X vector for expression as fusion proteins with maltose binding protein (MBP) according to the manufacturer's protocol. Following induction of protein expression with IPTG at 37 °C for 4 h, cells were harvested by centrifugation (4000 *g*, 20 min, 4 ºC). Cell pellets were resuspended in buffer A (20 mM Tris-HCl pH 7.4, 300 mM NaCl, 1mM EDTA) and lyzed by sonication on ice. Supernatant was collected after centrifugation (9,000 *g*, 30 min, 4 ºC) and loaded onto amylose resin (New England Biolabs). After two washes with five column volumes of buffer A, the target protein was eluted with three column volumes of buffer A supplemented with 10 mM maltose. Product fractions were concentrated in an Amicon-Ultra centrifugal filter unit (Amicon Bioseparations, Billercia, MA) and protein concentration quantified by measuring absorbance measurements at 280 nm (MBP-*Dm*dNK, ε = 106,230 $M^{-1}cm^{-1}$, calculated according to Pace *et al.*[41]). Protein purity was verified by SDS-PAGE. Sample aliquots were flash-frozen and stored at -80 ºC.

Kinetic analysis The pyruvate kinase/lactate dehydrogenase coupled-enzyme spectrophotometric assay[22] was employed for kinetic characterization at 37 °C, in a 500 ll reaction mixture containing 50 mM Tris– HCl (pH 7.5), 100 mM KCl, 2.5 mM MgCl2, 0.18 mM NADH, 0.21 mM phosphoenolpyruvate, 1 mM ATP, 1 mM 1,4-dithio-DL-threitol, 30 U/ml pyruvate kinase, and 33 U/ml lactate dehydrogenase. Substrate concentration varied from 1 μM to 500 μM. All experiments were performed in triplicate. Apparent kcat (calculated assuming one active site per enzyme monomer) and KM values were determined by fitting data to the Michaelis–Menten equation, using non-linear regression analysis in Origin (OriginLab, Northhampton, MA).

Reference:

1. Crick, F., *Central Dogma of Molecular Biology.* Nature, 1970. **227**: p. 561-563.
2. Marians, J.K., *Prokaryotic DNA replication.* Annu. Rev. Biochem., 1992. **61**: p. 673-719.
3. So, A.G., Downey, K.M., *Eukaryotic DNA replication.* Critical Reviews in Biochemistry and Molecular Biology, 1992. **27**: p. 129-155.
4. Mougel, M., Houzet, L., and Darlix, J.L., *When is it time for reverse transcription to start and go?* Retrovirology, 2009. **6**(24): p. 1-9.
5. *Cancer.* URL: http://www.who.int/mediacentre/factsheets/fs297/en/index.html, Feb, 2009.
6. *HIV/AIDS data and statistics.* URL: http://www.who.int/hiv/data/en/index.html, 2009.

7. Kaufmann, W. K., Paules, R.S., *DNA damage and cell cycle checkpoints.* FASEB, 1996. **10**(238-247).

8. Meikrantz, W., Schlegel, R., *Apoptosis and the Cell Cycle.* Journal of Cellular Biochemistry, 1995. **58**: p. 160-174.

9. Chan, D.C., Kim, P.S., *HIV Entry and Its Inhibition.* Cell, 1998. **93**: p. 681-684.

10. Robey, E., Axel, R., *CD4: Collaborator in Immune Recognition and HIV Infection.* Cell, 1990. **60**: p. 697-700.

11. Agrawal, L., Lu, X., Jin, Q., and Alkhatib, G., *Anti-HIV Therapy: Current and Future Directions.* Current Phamaceutical Design, 2006. **12**: p. 2031-2055.

12. Clercq, E.D., *New Development in Anti-HIV Chemotherapy.* Current Medicinal Chemistry, 2001. **8**: p. 1543-1572.

13. Shore, S., Raraty, M.G.T., Ghaneh, P., and Neoptolemos, J.P., *Chemotherapy for pancreatic cancer.* Aliment Pharmacol Ther., 2003. **18**: p. 1049-1069.

14. Solaroli, S., Johansson, M., Balzarini, J., Karlsson, A., *Enhanced toxicity of purine nucleoside analogs in cells expressing Drosophila melanogaster nucleoside kinase mutants.* Gene Therapy, 2007. **14**: p. 86-92.

15. Zheng, X., Johansson, M., and Karlsson, A., *Retroviral Transduction of Cancer Cell Lines with the Gene Encoding Drosophila melanogaster Multisubstrate Deoxyribonucleoside Kinase.* J Biol Chem, 2000. **275**: p. 39125-39129.

16. Zheng X., Johansson, M., and Karlsson, A., *Nucleoside Analog Cytotoxicity and Bystander Cell Killing of Cancer Cells Expressing Drosophila melanogaster Deoxyribonucleoside Kinase in the Nucleus or Cytosol.* Biochemical and Biophysical Research Communications, 2001. **289**: p. 229-233.

17. Mathews, K.C., and Song, S., *Maintaining precursor pools for mitochondrial DNA replication.* FASEB, 2007. **21**: p. 2294-2303.

18. Shi, J., McAtee, J. J., Schlueter Wirtz, S., Tharnish, P., Juodawlkis, A., Liotta, D.C., and Schinazi, R.F., *Synthesis and biological evaluation of 20,30-didehydro-20,30-dideoxy-5-fluorocytidine (D4FC) analogues: discovery of carbocyclic nucleoside triphosphates with potent inhibitory activity against HIV-1 reverse transcriptase.* J. Med. Chem.,, 1999. **42**: p. 859-867.

19. Gentry, G.A., *Viral thymidine kinases and their relatives.* Pharmacol. Ther., 1992. **54**: p. 319-355.

20. Black, M.E., Newcomb, T.G., Wilson, H.M., Loeb, L.A., *Creation of drug-specific herpes simplex virus type 1 thymidine kinase mutants for gene therapy.* Proc. Natl Acad. Sci., 1996. **93**: p. 3525-3529.

21. Black, M.E., Kokoris, M.S., and Sabo, P., *Herpes Simplex Virus-1 Thymidine Kinase Mutants Created by Semi-Random Sequence Mutagenesis Improve Prodrug-mediated Tumor Cell Killing.* Cancer Research, 2001. **61**: p. 3022-3026.

22. Gerth, M., Lutz, S., *Non-homologous recombination of deoxyribonucleoside kinases from human and Drosophila melanoganster yields huamn-like enzymes with novel activities.* J. Mol. Biol., 2007. **370**: p. 742-751.

23. Munch-Petersen, B., Piskur, J., and Sondergaard, L., *Four deoxynucleoside kinase activities from Drosophila melanogaster are contained within a single monomeric enzyme, a new multifunctional deoxynucleoside kinase.* J Biol Chem, 1998. **273**(7): p. 3926-3931.

24. Munch-Petersen, B., Knecht, W., Lenz, C., Songdergaard, L., Piskur, J., *Functional expression of a multisubstrate deoxyribonucleoside kinase from Drosophila melanogaster and its C-terminal deletion mutants.* J Biol Chem, 2000. **275**: p. 6673-6679.

25. Johansson, K., Ramaswamy, S., Ljungcrantz, C., Knecht, W., Piskur, J. et al., *Structural basis for substrate specificities of cellular deoxyribonucleoside kinases.* Nature Struct. Biol., 2001. **8**: p. 616-620.

26. Hebrard, C., Dumontet, C., and Jordheim, L. P., , *Development of gene therapy in association with clinically used cytotoxic deoxynucleoside analogues.* Cancer Gene Theraphy, 2009. **16**: p. 541-550.

27. Knecht, W., Rozpedowska, E., Le Breton, C., and Piskur, J., et al. , *Drosophila deoxyribonucleoside kinase mutants with enhanced ability to phosphorylate purine analogs.* Gene Therapy, 2007. **14**: p. 1278-1286.

28. Black, M.E., Loeb, L.A., *Identification of important residues within the putative nucleoside binding site of HSV-1 thymidine kinase by random sequence selection: analysis of selected mutants in vitro.* Biochemistry, 1993. **32**: p. 11618-11626.

29. Barrio, J.R., Secrist, J.A., Leonard, N.J., *Fluorescent adenosine and cytidine derivatives.* Bioochem Biophys Res Commun, 1972. **46**: p. 597-604.

30. McGuigan, C., Carangio, A., Snoeck, R., et al., *Synthesis and antiviral evaluation of some 3'-fluoro bicyclic nucleoside analogues.* Nucleosides, Nucleotides & Nuclic Acids, 2004. **23**: p. 1-5.

31. Hawkins, M.E., *Fluorescent Pteridine Nucleoside Analogs.* Cell Biochemistry and Biophysics, 2001. **34**: p. 257-281.

32. Liu L., Li, Y., Liotta, D., and Lutz, S., *Directed evolution of an orthogonal nucleoside analog kinase via fluorescence-activated cell sorting.* Nucleic Acids Research, 2009. **37**(13): p. 4472-4481.

33. Wang, Y., Rivera Vera, C.I., , and Lin, Q., *Convenient Synthesis of Highly Functionalized Pyrazolines via Mild, Photoactivated 1,3-Dipolar Cycloaddition.* Organic Letters, 2007. **9**(21): p. 4155-4158.

34. Song ,W., Wang, Y., Jin, Q., and Lin, Q., *Selective Functionalization of a Genetically Encoded Alkene-Containing Protein via 'Photoclick Chemistry' in Bacterial Cells.* J. Am. Chem. Soc., 2008. **130**: p. 9654-9655.

35. Janta-Lipinski, M., Costisella, B., Ochs, H., et al., *Newly synthesized L-enatiomers of 3'-fluoro-modified β-2'-deoxyribonucleoside 5'-triphosphate inhibit Hepatitis B DNA polymerases but not the five cellular DNA polymerases α,β,γ,δ andεnor HIV-1 reverse transcriptase.* J.Med. Chem, 1998. **41**: p. 2040-2046.

36. Martin, J.A., Bushnell, D.J., Duncan, I.B., et al, *Synthesis and antiviral activity of monofluoro and difluoro analogues of pyrimidine deoxyribonucleosides against human innmunodeficiency virus (HIV-1).* J. Med. Chem, 1990. **33**: p. 2137-2145.

37. Schelling, P., Flockers, G., Scapozza, L.,, *A spectrophotometric assay for quantitative determination of kcat of herpes simplex virus type 1 thymidine kinase substrates.* Anal Biochem, 2001. **295**: p. 82-87.

38. Fujimoto, K., Matsuda, S., Takahashi, N., and Saito, I.,, *Template-directed photoreversible ligation of deoxyoligonucleotides via 5-vinyldeoxyuridine.* J. Am. Chem. Soc., 2000. **122**: p. 5646-5647.

39. Ito, S., Tanka, Y., Kakeh, A., and Kondo, K.,, *A facile synthesis of 2,5-distubstituted tetrazole by the reaction of phenylsulfonylhydrazones with arenediazonium salts.* Bull. Chem. Soc. Jpn. , 1976. **49**: p. 1920-1923.

40. Huisgen, R., *1,3-Dipolar Cycloadditions: Past and Future.* Angew. Chem., Int. Ed., 1963. **2**: p. 565-598.

41. Pace, C.N., Vajdos, F., Fee, L., Grimsley, G., and Gray, T., *How to measure and predict the molar absorption coefficient of a protein.* Protein Sci., 1995. **4**: p. 2411-2423.

42. Zheng, X., Johansson, M., and Karlsson, A., *Nucleoside Analog Cytotoxicity and Bystander Cell Killing of Cancer Cells Expressing Drosophila melanogaster Deoxyribonucleoside Kinase in the Nucleus or Cytosol.* Biochemical and Biophysical Research Communications, 2001. **289**: p. 229-233.