CHAPTER 13

Synthetic Biology

SUMMARY

Biochemical pathways can be engineered to produce more desirable products, such as ethanol, dyes, or antibiotics, or degrade various substances, including pollutants and agricultural waste. In some cases, a patchwork of different pathways is generated that produces a novel biochemical pathway. The biochemical pathway for ethanol production could be altered for the production of ethanol from waste material as an additive to gasoline. Since ethanol is produced from the fermentation of sugars, pathways could be engineered to increase the range of fermentable sugars. Yeast and *Zymomonas*, a bacterium, are excellent organisms to carry out fermentation, since they typically produce only ethanol. *Escherichia coli* produces a mixture of fermentation end products. Ethanol would need to be purified in separate processes that could add time and cost to the project. *Zymomonas* grows only on glucose. One type of sugar, xylose, is readily available in plant cell walls of plant waste materials. The xylose fermentation pathway has been engineered into *Zymomonas* with great success. Recycling of plant-based products, such as paper, into ethanol could be accomplished using similar pathways.

Multiple biochemical pathways exist for the degradation of starch, cellulose, and aromatic compounds. Starch is a polymer of glucose that is formed in plants to store excess sugars. More efficient degradation of starch into individual sugar molecules could be accomplished by engineering the enzymes in the pathways to be more stable and effective. Cellulose is another polymer of glucose present in plant cell walls. Even though cellulose and starch are both polymers of glucose, the bonds formed between the glucose are different and thus require different enzymes to break them apart. Pathways for cellulose degradation already exist in microorganisms that degrade plant material, such as soil microbes and fungi. However, the pathways are slow. Since a large amount of paper waste, which is derived from plant material and contains cellulose, is making its way to landfills, engineering a more efficient pathway for the degradation of cellulose is becoming increasingly important. The use of "second-generation biofuels" would produce more fuels from material that is traditionally wasted, such as the cellulose, hemicellulose, and lignin that make up plant cell walls. The use of these waste products for fuel generation would keep sugarcane, corn, or vegetable oil in the food production line instead of being converted into first-generation biofuels."

Biodiesel is produced from plant oils and animal fats by treating them with methanol and sodium hydroxide. This process converts the oils and fats to fatty acid methyl esters plus glycerol and water; the latter two are waste products and removed prior to use. Currently, biodiesel is largely produced from feedstock plant oils. However, it is possible to generate second-generation biodiesel from nonedible plants or engineered algae or bacteria. The engineered microbes would produce large quantities of fatty acids that would then be converted to biodiesel.

In terms of agriculture, frost damage occurs to many crops each year. When water freezes, it expands and causes tissue damage. Ice crystals do not form without having some sort of nucleation seed. This seed is often provided by bacteria living on the surface of plants that have nucleation factors, which promote the ice formation. The only way to prevent ice from forming on the plant, and subsequent tissue damage, is to eliminate the nucleation factors. The nucleation factors, such as the InaZ protein on the surface of *E. coli,* can be genetically disrupted. The crops can then be sprayed with the mutant bacteria to help prevent ice formation and crop destruction. Ice nucleation proteins are present on outer membranes of bacteria and project one of their protein ends outward. These proteins could be used in surface display applications where other proteins are displayed on the outside of bacterial cells and used to detect and remove viruses or used in the display of enzymes on the external surface of the cells.

Fossil fuels, such as coal and oil, contain a large amount of sulfur, which is problematic when these fossil fuels are burned. Acid rain is a direct result of sulfur entering the atmosphere from the burning of these fuels. Organisms such as *Thiobacillus* and *Sulfolobus* convert inorganic sulfur to a soluble form that can be more easily removed. The organic sulfur found within the fossil fuels is present as a thiophene, which ironically is thought to be of biological origin from archaebacterial metabolism. The pathway to completely desulfurize thiophenes is called the 4S pathway. No known microbe has the capability to perform all of the steps, but several microbes possess partial pathways. Engineering these partial pathways into one microbe would generate the complete 4S pathway, which would provide a means to eliminate the sulfur from the fossil fuel prior to combustion.

One area of focus in research is to engineer the biosynthetic pathway for penicillin in the *Penicillium* mold to increase production of this antibiotic. Cephalosporin is a related antibiotic to penicillin that is made by a similar biosynthetic pathway in a related fungus. Additionally, antibiotics are modified to function as new antibiotics, which help prevent antibiotic resistance in microorganisms, particularly resistance from those microbes of medical importance. These derivatives are produced by modifying enzymes in biosynthetic pathways for the production of antibiotics with new properties. Most of the new derivatives are β-lactams, many of which are further classified as carbapenems. For antibiotic resistance, β-lactams are degraded by the enzyme β-lactamase. Most carbapenem antibiotics are resistant or even inhibitory to β-lactamases.

Some bacteria have the ability to manufacture plastics. These biosynthesized plastics are biodegradeable. In the future, plants could be engineered to contain the enzymes needed for the biosynthesis of these plastics, which would greatly decrease the cost and increase the yield. In addition, pathways can be engineered as sensors of environmental conditions, such as a pollutant, and respond by giving off some sort of signal (e.g., fluorescence).

Unnatural nucleic acids, such as XDNA and XNA, have been produced that still act as genetic material, yet have a range of mild to radical changes in the bases or backbone. These synthetic genetic materials are developed to be resistant to degradation by nucleases, and some are even able to transmit genetic information. XDNA is a synthetic nucleic acid containing the addition of benzene rings to the nitrogen-containing bases (A, T, G, C). XDNA strands are able to pair with normal bases *in vitro*, but the DNA double helix is wider than normal. Small segments of XDNA can even be replicated *in vivo*. XNA is another synthetic nucleic acid that has alternative ring structures in place of the natural pentose sugars. Examples include TNA, ANA, FANA, LNA, CeNA, and HNA. All of these examples differ by the specific ring structure introduced into the nucleic acid backbone. XNAs are complementary to RNA and DNA, resistant to nucleases, and able to transmit genetic information in models containing engineered DNA polymerases.

Bacteria that have been designed and engineered to perform a specific function are called designer bacteria. Designer bacteria include those microbes that are engineered to produce biofuels or dispose of waste through bioremediation. Bacteria can even be designed to mend cracks in concrete or kill pathogens resistant to antibiotics. In both cases, the bacteria communicate via quorum sensing.

In the end, pathway engineering and synthetic biology produce novel activities for both the biochemical pathways and the organisms to which the pathways belong. Pathway engineering and synthetic biology are used in a wide range of applications, from biofuels to antibiotics to designer bacteria with specific missions.

Case Study A Synthetic Substrate of DNA Polymerase Deviating from the Bases, Sugar, and Leaving Group of Canonical Deoxynucleoside Triphosphates

Mi-Yeon Jang et al. (2013). *Chemistry & Biology* 20, 416–423.

The use of alternative chemical building blocks in the formation of nucleic acid has advantages, including increased stability of the new structure relative to existing natural DNA, better understanding of the involvement of certain functional groups in synthesis of the nucleic acid, and development of an artificial genetic memory system for *in vivo* applications.

In this research article, the authors use 4′-inverse nucleotides (XNA) in which the 4′-CH_2OP is substituted with 4′OCH_2P. This process generates a nucleotide with two anomeric centers plus a phosphate group. The electrostatic potential of the resulting nucleotides is changed, and the molecules have increased stability and are resistant to degradation by enzymes. Additionally, previous research groups have identified substitutions for both the pyrophosphate group on natural nucleoside triphosphates and the nitrogenous bases. The pyrophosphate can be substituted with nonphosphate groups such as L-aspartic acid, L-phosphonoalanine, and iminodipropionic acid. The nitrogenous base can be successfully substituted with 5-chlorouracil. Information regarding substitutions to both the pyrophosphate and nitrogenous base is based on previous research articles.

The long-term goal of this study is to create a completely artificial genetic memory system that can be synthesized and maintained within cells, using the cell's own machinery, and to associate this system with a specific life-maintaining function.

How are the XNA monomers synthesized in this research article different from natural DNA monomers? What specific changes were introduced to create the XNA monomers?

All genetic monomers, whether natural or synthetic, contain a leaving group, backbone motif, and a nucleobase. In natural genetic monomers, the leaving group is a diphosphate, the pentose plus one phosphate constitutes the sugar-phosphate backbone, and the nucleobases are the traditional nitrogen-containing bases (A, T, C, G). In XNA monomers, all three components are unnatural. In this paper, the XNA monomers contain two anomeric centers at positions C1′ and C4′ on the leaving group. The backbone motif contains a 4′OCH_2P. The nitrogenous base is a chlorinated-derivative of uracil. (Refer to Figure 1 of the article.) Several analogs were constructed, which differed in various aspects of the three major parts of the monomer.

In the synthesized nucleosides, does having two anomeric centers cause any major conformational changes?

Yes. The effect of the C4′ anomeric center is stronger than that observed for the C1′ anomeric center, which causes some shifting in orientation of the nucleobase.

Can the XNA monomers be used *in vivo*?

A primer-template assay using various DNA polymerases was employed to determine if the XNA monomers could be used *in vivo*. These polymerases included HIV-1 reverse transcriptase (HIV-1 RT), Taq DNA polymerase, Vent(exo-) DNA polymerase, DNA pol III polymerase, and Therminator polymerase. The many XNA analogs produced by this research group had varying degrees of success depending on the polymerase used. The different polymerases had varying affinities for the XNA analogs. Several analogs were used by HIV-1 RT, which might have some significance for therapeutic targets. DNA pol III was able to incorporate only one analog out of all of the analogs produced. The most indiscriminate polymerases were the heat-stable Taq and Vent enzymes and, in particular, the Therminator polymerase. Only one analog was not recognized and incorporated by Therminator.

What is the significance of the polymerases used in the incorporation study?

HIV-1 RT is an error-prone polymerase involved in HIV virus replication. It also represents a potential target for HIV therapeutics. Taq DNA polymerase, Vent DNA polymerase, and Therminator polymerase are heat-stable polymerases. DNA polymerase III is not homologous to the other polymerases used. In addition, DNA pol III has proofreading abilities. By using a wide range of polymerases, the investigators were able to survey the effects of using XNA analogs to synthesize unnatural genetic information in multiple systems encompassing a wide variety of biotechnology applications.

Is it possible to form stable XNA:DNA duplexes? If so, what differences occur in XNA:DNA hybrid duplexes relative to natural double-stranded DNA?

Yes. XNA:DNA hybrids are produced and are stable structures with only slight differences. The bases in XNA:DNA hybrids are only slightly inclined, but not significantly. The hybrid helix does not become wider, nor does it become longer than natural double-stranded DNA. The major groove is wider and deeper, but the minor groove is slightly wider and less deep. At the atomic level, in the XNA strand, some of the oxygen atoms are more exposed to the solvent in the environment, which might result in hydrogen bond formation. The most significant difference is a puckering effect of the sugars in both strands of a hybrid duplex that is not observed in a natural DNA duplex.

The overall goal of this study is to produce synthetic genetic material that can be synthesized *in vivo* and used potentially for some life function of the organism. Does this study move toward attaining this goal?

Yes. This study moves the authors closer to their goal. The authors were able to construct various XNA analogs containing unnatural compounds in all three parts (leaving group, backbone, nucleobase) of the monomers. These XNA monomers were recognized by natural polymerases (as opposed to engineered polymerases) and incorporated into growing chains of synthetic genetic material.

The authors of this research article constructed completely novel chemical compounds in order to mimic natural nucleoside triphosphates found in genetic information. In their analogs, all three major parts of

Case Study A Synthetic Substrate of DNA Polymerase Deviating from the Bases, Sugar, and Leaving Group of Canonical Deoxynucleoside Triphosphates—cont'd

the genetic monomers were changed to unnatural compounds. Structurally, some slight differences occur when comparing the synthetic monomers to natural monomers. Additionally, DNA:DNA and XNA:DNA hybrids were modeled to determine any significant changes to the structure of the duplex. Only slight changes were observed. The authors also determined that most of their analogs were recognized by at least one natural polymerase. Of all the polymerases tested, the heat-stable Therminator polymerase could utilize the widest variety of synthetic analogs. Several other polymerases had varying degrees of success with the analogs. The fact that these synthetic monomers could be recognized and polymerized by natural enzymes advances the overall goal of the study to produce an artificial genetic memory system that can be synthesized *in vivo* and functions in some aspect of life.

Chemistry & Biology
Article

A Synthetic Substrate of DNA Polymerase Deviating from the Bases, Sugar, and Leaving Group of Canonical Deoxynucleoside Triphosphates

Mi-Yeon Jang,[1,3] Xiao-Ping Song,[1,3] Mathy Froeyen,[1] Philippe Marlière,[2] Eveline Lescrinier,[1] Jef Rozenski,[1] and Piet Herdewijn[1,2,*]
[1]Rega Institute for Medical Research, Katholieke Universiteit Leuven, Leuven, 3000, Belgium
[2]Institute of Systems and Synthetic Biology, University of Evry-Val-d'Essonne, Evry, F-91030, France
[3]These authors contributed equally to this work
*Correspondence: piet.herdewijn@rega.kuleuven.be
http://dx.doi.org/10.1016/j.chembiol.2013.02.010

SUMMARY

The selection of artificial nucleic acids to be used for synthetic biology purposes is based on their structural and biochemical orthogonality to the natural system. We describe the example of a nucleotide mimic that functions as a substrate for polymerases and in which the carbohydrate moiety as well as the base moiety and the leaving group are different from that of the natural building blocks. The nucleotides themselves have two anomeric centers, and different leaving group properties of substituents at both anomeric centers need to be exploited to perform selective glycosylation reactions for their synthesis. In addition, the reversibility of the polymerase reaction at the level of the template has been demonstrated when pyrophosphate functions as leaving group and not with the alternative leaving groups.

INTRODUCTION

To what extent could the structure of nucleic acids be distorted while maintaining their function of hereditary information carriers? Answering this question experimentally would be of importance for understanding the etiology of nucleic acids as pioneered by Eschenmoser (1999). It would also be of practical importance for deploying a xenobiology, i.e., an artificial biodiversity not interfering with natural species and ecosystems (Herdewijn and Marlière, 2009). In nature, only four different RNA (U, C, A, and G) and DNA (T, C, A, and G) monomers have been found. Nucleic acid polymerization in nature uses pyrophosphate as a leaving group, a phosphorylated hydroxymethyltetrahydrofuran as a backbone motif, and no element other than C, N, H, and O in nucleobases. Segregation between the natural and artificial information systems can be reached most efficiently when the composition of the building blocks (consisting of a sugar base and leaving group) used for the synthesis of the artificial information system are different from the composition of the natural building blocks used for DNA and RNA synthesis. Here, we present an example of the incorporation of nucleotide analogs in which these three entities are replaced by chemicals that have not been selected by nature for this purpose (Figure 1).

Synthetic biology follows either the engineering approach (to make an unnatural system using natural bricks) or the chemical approach (to make natural-like systems using unnatural bricks). We follow the second approach. In order not to interfere with the natural system, the chemistry that is used should be orthogonal to the chemistry that is used by nature. Our prime focus is on the development of orthogonal nucleic acids and to implement them in vivo. Here, we have evaluated 4′-inverse nucleic acids (XNA in which the 4′-CH_2OP group of the natural furanose sugar is replaced by a 4′-OCH_2P group, resulting in nucleotides with two anomeric centers and a phosphonate group) for their potential orthogonality. The conformational mobility of the sugar moiety may be tuned by the presence of both anomeric centers. This chemical modification should also lead to a change in the electrostatic potential at the surface of these nucleic acids. The replacement of a phosphate backbone by a phosphonate will increase the stability of such nucleic acids against enzymatic degradation. The long-term goal of this project is to synthesize such nucleic acids within the cell, using its own biochemistry, and to associate a function to the synthetic information system indispensable for the survival of the organism (Herdewijn and Marlière, 2009).

In recent studies, it was demonstrated that the pyrophosphate group of nucleoside triphosphates can be substituted by nonphosphate-based leaving groups such as L-aspartic acid (Adelfinskaya and Herdewijn, 2007), L-phosphonoalanine (Yang et al., 2011), and iminodipropionic acid (Song et al., 2011a) (Figure S1 available online). The selection of the 5-chlorouracil base is based on its successful introduction as a substitute for the thymine base in the genome of *Escherichia coli* (Marlière et al., 2011). The ''ribo'' congener of the phosphonate nucleoside was synthesized earlier as a potential antiviral agent (Kim et al., 1991) and as a mimic of a ribonucleoside monophosphate.

An additional advantage of using such drastically modified nucleotides for incorporation in studies using different polymerases is that the function of pyrophosphate in the enzymatic polymerization reaction can be better studied. As well, the nucleotides with a pyrophosphate leaving group, as with an alternative leaving group, have been synthesized and tested.

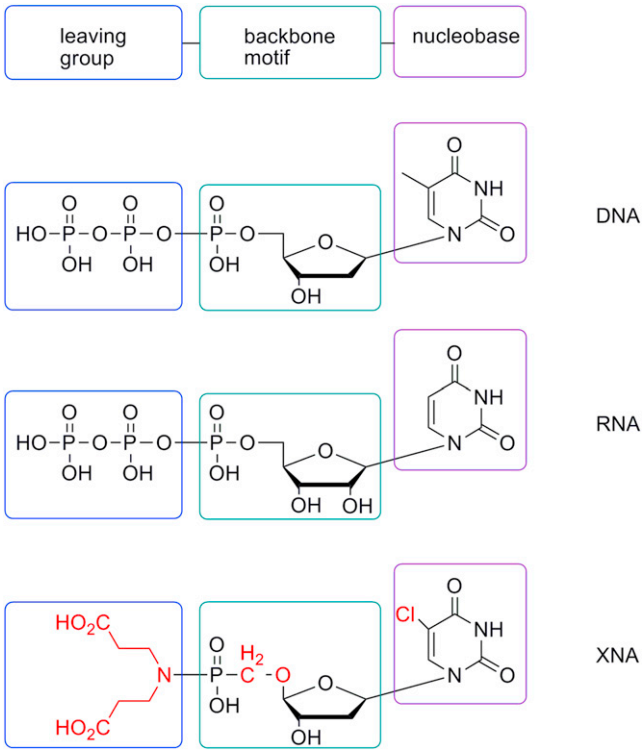

Figure 1. Structure of the Monomers for DNA, RNA, and XNA Synthesis

The monomer is composed of leaving group, backbone motif, and nucleobase. In XNA monomer, all three components are unnatural.

See also Figure S1.

Indeed, by using Therminator polymerase as catalyst, the pyrophosphate that is released with the incorporation of a nucleotide in the primer was found to induce pyrophosphorolysis in the template, followed by incorporation of nucleotide entities. This template degradation reaction does not occur when nucleotide building blocks and other leaving groups are used for XNA synthesis, which in some cases might be an advantage when implementing such xenonucleotide building blocks in vivo.

RESULTS

Synthesis of Modified Nucleotides

The nucleotides **10a–c** were synthesized from 1-*O*-methyl-2-deoxy-D-ribofuranose (Hoffer, 1960) (Figure 2). The primary hydroxyl group at C-5 of compound **1** was selectively protected with a *tert*-butyldimethylsilyl (TBDMS) group, and the secondary hydroxyl group was protected by benzoylation. When compound **2** was treated with tetrabutylammonium fluoride (TBAF) in tetrahydrofuran (THF) to deprotect the TBDMS group, the 3-*O*-benzoyl group migrated to the 5-OH position. This side reaction can be suppressed by the addition of a catalytic amount of acetic acid. The 5-hydroxymethyl group of **3** was oxidized to carboxylic acid by treatment with bis(acetoxy)iodobenzene in the presence of a catalytic amount of 2,2,6,6-tetramethylpiperidyl-1-oxyl (TEMPO) (Epp and Widlanski, 1999). Oxidative decarboxylation via a modified Hunsdiecker reaction transformed acid **4** into the 4β-acetate **5** as a single isomer (Dhavale et al., 1988). Reaction of acetate **5** with hydroxymethyl phosphonic acid diisopropyl ester in the presence of trimethylsilyl trifluoromethanesulfonate (TMSOTf) at −20°C led to the exclusive formation of the desired phosphonate **6**. Demethylation of anomeric methoxy group and acetylation of the hydroxyl group was carried out with $Ac_2O/AcOH/H_2SO_4$ in dichloromethane (DCM) to yield compound **7** (Földesi et al., 2001). The nucleobases 5-chlorouracil, thymine, and N^6-benzoyladenine were introduced after silylation with *N,O*-bis(trimethylsilyl) acetamide (BSA) and using TMSOTf as a Lewis catalyst at 0°C to yield a mixture of the α/β diastereoisomers. Separation of individual isomers was accomplished by column chromatography. The 3′-*O*-benzoyl group was removed with ammonia in methanol. Unfortunately, we were not able to remove the N^6-benzoyl group of compound **8c**. At room temperature, no reaction took place, whereas at an elevated temperature (60°C), the compound decomposed. As an alternative synthetic route, 6-chloropurine was introduced as a nucleobase, affording compound **8d**. The crude product was used in the next step, and compound **9d** was purified by high-performance liquid chromatography (HPLC). The overall yield of this reaction was very poor (~10%), as 4-isomers were formed (7α/7β/9α/9β = 1/1/1/1). Hydrolysis of the phosphonate esters **9a**, **9b**, and **9d** was carried out with iodotrimethylsilane (TMSI) in the presence of hexamethyldisilazane (HMDS) at 0°C within 1 hr. Prolonged reaction time and elevated temperature causes C1′-epimerization.

The diphosphono-phosphonates, **11a–c** were prepared from the phosphonate diacids **10a**, **10b**, and **10c**, respectively, following a literature procedure (Koh et al., 2005; Mackman et al., 2007). To obtain the phosphonoamidate analogs **16a-c**, **17a-c**, and **18**, the methyl esters of iminodipropionic acid (Nakatani et al., 2001), L-aspartic acid, and methyl 2-amino-3-(dimethoxyphosphoryl)propanoate (Schneider et al., 2006) were coupled to the phosphonate diacids **10a–c** following minor modification of the method described by Abraham et al. (1996). In this step, the dimerized compounds **15a–c** were obtained as side products. According to thin layer chromatography (TLC) analysis, first dicyclohexylcarbodiimide (DCC)-mediated P-N bond formation occurred. Prolonged reaction time caused the cleavage of the P-N bond of compounds **12a–c**, **13a–c**, and **14**, yielding compounds **15a–c**. For example, if the reaction mixture was heated for more than 5 hr, only dimerized compounds were isolated. Deprotection of **12a–c** and **13a–c** was easily accomplished using 0.4 M NaOH in methanol. However, deprotection of compound **14** was cumbersome. When the methyl ester was hydrolyzed in basic conditions and followed by deprotection of the phosphonate diester functions with TMSI, the compound was decomposed yielding **10b**. Fortunately, the reverse procedure (first deprotection of the phosphonate diester function and then hydrolysis of methyl ester) afforded a small amount of the desired compound **18**.

Nucleotide Incorporation Experiments

To explore the substrates ability of the synthesized nucleotide analogs (compound **11a–c**, **15a–c**, **16a–c**, **17a–c**, and **18**) for incorporation into a growing DNA chain, they were evaluated in a primer-template assay using several polymerases: HIV-1 reverse transcriptase (HIV-1 RT), Taq DNA polymerase, Vent

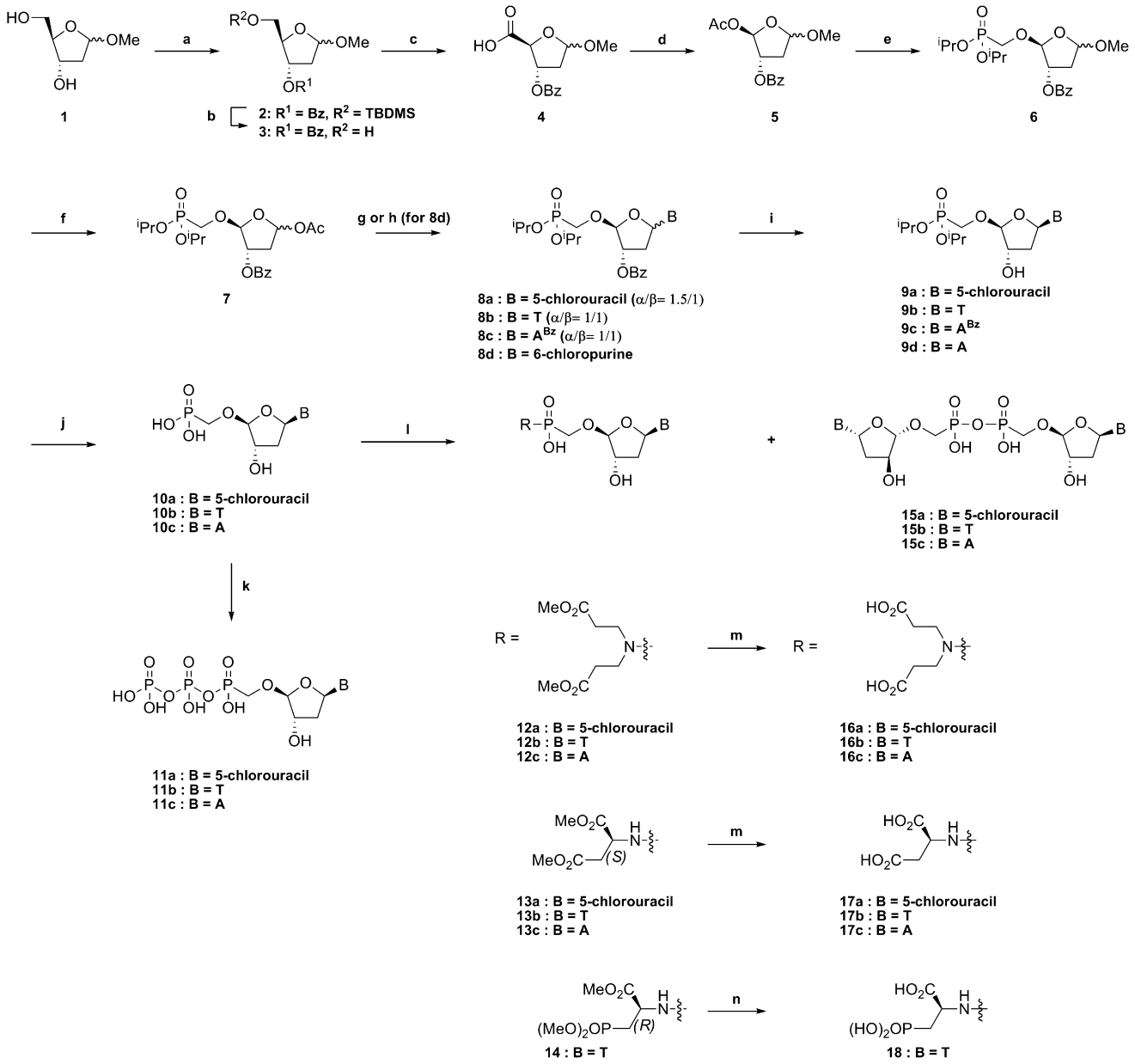

Figure 2. Synthesis of the Phosphonate Nucleoside Analogues

Reagents and conditions used for the synthetic scheme: (a) 1) TBDMSCl, pyridine, 0°C to room temperature (rt), 2) benzoyl chloride, DMAP; (b) TBAF, AcOH, THF, 0°C to rt; (c) $PhI(OAc)_2$, TEMPO, CH_3CN/H_2O (1/1, v/v), rt, overnight; (d) $Pb(OAc)_4$, pyridine, THF, rt, overnight; (e) TMSOTf, hydroxymethyl phosphonic acid diisopropyl ester, CH_2Cl_2, −20°C, overnight; (f) Ac_2O, AcOH, H_2SO_4, CH_2Cl_2, 0°C, 15 min; (g) *N*,*O*-bis(trimethylsilyl)acetamide, base, TMSOTf, 0°C, overnight; (h) *N*,*O*-bis(trimethylsilyl)acetamide, 6-chloropurine, TMSOTf, rt, 3 hr; (i) 7N NH_3 in MeOH, rt, overnight; (j) HMDS, TMSI, CH_3CN/CH_2Cl_2 (1/1, v/v), 0°C, 1 hr; (k) Bu_3N, *N*,*N*-carbonyldiimidazole, tris(tetrabutylammonium)hydrogen pyrophosphate, DMF, rt, overnight; (l) amine, DCC, water/t-BuOH(1/3, v/v), 80°C, 1 hr; (m) 0.4 M NaOH/MeOH (1/1, v/v), rt, 3 hr; (n) 1) HMDS, TMSI, CH_3CN/CH_2Cl_2 (1/1, v/v), 0°C, 1 hr; 2) 0.4 M NaOH/MeOH (1/1, v/v), rt, 1 hr.

(exo-) DNA polymerase, DNA pol III polymerase, and Therminator polymerase. Incorporation efficiency was analyzed by polyacrylamide-gel electrophoresis. The natural deoxyadenosine triphosphate (dATP) and deoxythymidine triphosphate (dTTP) were used as reference.

HIV-1 RT is an error-prone polymerase responsible for viral replication. HIV-1 RT shows flexibility and high tolerance toward modified nucleotides, and these properties make it a primary target in the treatment of HIV infection (Bebenek et al., 1993). Phosphoramidate analogs (Figure S1) are efficient substrates of HIV-1 RT. However, no incorporation occurred when, respectively, L-Asp-phosphonate analogs (**17a–c**), iminodipropionic acid-phosphonate [(IDP)-phosphonate] analogs (**16a–c**), and 3-phosphono-L-Ala-phosphonate analog (**18**) were used as substrates in the polymerization reaction. Only efficient incorporation was observed when using

Table 1. Single Nucleotide Incorporation of Modified Nucleotide Analogues by Various Polymerases

Cmpd (500 μM)	HIV-1 RT	Pol III	Taq	Vent	Therminator
11a	76	0	73	100	100
11b	100	0	75	98	100
11c	93	47	55	100	100
15a	0	0	0	0	88
15b	0	0	0	0	0
15c	89	0	46	100	84
16a	0	0	0	51	100
16b	0	0	0	0	86
16c	0	0	40	49	100
17a	0	0	10	40	99
17b	0	0	10	0	87
17c	0	0	54	97	95
18	0	0	20	10	100

Enzyme concentrations used for single nucleotide incorporation: [HIV-1 RT] = 0.025 U/μl; [Pol III] = 1.6 U/μl; [Taq] = 0.15 U/μl; [Vent] = 0.1 U/μl; [Therminator] = 0.05 U/μl, incubation time was 120 min, and the incorporation was expressed as percentage of total radio-emitting oligonucleotides. P1T1 and P1T2 are primer/template complexes used for incorporation reaction [Primer (P1) 5′-CAGGAAACAGCTATGAC-3′; Template (T1) 3′-GTCCTTTGTCGATACTG**TCCC**-5′; Template (T2) 3′-GTCCTTTGTCGATACTG**ACTGC**-5′], P1T1 was used for compounds **11c**, **15c**, **16c**, and **17c**, P1T2 was used for compounds **11a**, **11b**, **15a**, **15b**, **16a**, **16b**, **17a**, **17b**, and **18**. Cmpd, compound; HIV-1 RT, human immunodeficiency virus-1 reverse transcriptase; Pol III, polymerase III. See also Figures S2 and S3.

the pyrophosphate-phosphonate [(PPi)-phosphonate] analogs (**11a–c**) and the dimerized compound **15c** (Table 1). Among them, compound **11b** showed the best result, and efficient incorporation was observed when the substrate concentration was decreased to 10 μM (Figure S2A).

The synthesized compounds were also tested for incorporation in DNA by using several thermostable polymerases, and they demonstrated excellent recognition and incorporation of a number of synthesized compounds as substrate. Among them, Therminator polymerase showed the best results. All of the synthesized analogs, except compound **15b**, were recognized by Therminator polymerase and efficiently incorporated into a growing primer strand. At 100 μM, **11a–c** and **18** led to 100% primer extension within 10 min (Figures S2B–S2E). Even when the other compounds were less efficient, they were incorporated in over 84% yields at 500 μM in 120 min (Table 1). In some cases (Figure S2C), P+1 product increased by increasing the substrate concentration. This might explain why the modified nucleotide acts as an inhibitor at high concentrations during enzymatic polymerization reaction, whereas they are substrates at low concentration (Song et al., 2011a).

A nontemplated nucleotide incorporation at the 3′-end of a DNA fragment catalyzed by some polymerases has been described before (Hu, 1993). This may explain the large amount of P+2 products obtained by using Therminator polymerase (Figures S2B–S2E). Even at a very low enzyme concentration ([Therminator pol] = 0.00005 U/μl), P+2 is the main product that is formed (Figure S3) by terminal transferase activity.

In contrast to compounds **11b**, **11c**, and **18**, compound **11a** showed double spots at the P+2 level (Figure S2B). As repeated experiments showed the same patterns, further investigation of the identity of the P+2 compound was performed by mass spectrometry analysis. The mass difference between the two oligonucleotides is 20 Da and corresponds to the mass difference between **10a** and dTMP (Figure S4). This means that, in one case, the primer (P) is extended with two chlorouracil nucleotides (**10a**), and, in the other case, the primer (P) is extended with one chlorouracil nucleotide (**10a**) and dTMP. Careful analysis of the reaction mixture obtained after the incorporation study revealed that the 3′ end of template was degraded by two nucleotides (G and T) (Figure S4). dTMP, which is incorporated into the primer DNA strand, finds its origin in the degradation of the template by pyrophosphorolysis with the generation of deoxyguanosine triphosphate (dGTP) and dTTP (Figure S5). To further investigate the pyrophosphorolysis reaction of the template, we performed template degradation assays (Anderson et al., 1999). Significant degradation of the template was observed with compounds **11b** and **11c**, whereas in the presence of pyrophosphatase (PPase), the amount of degraded template is reduced (Figure S6). The degradation of the template is clearly induced by PPi. More degraded template is observed after 10 min rather than after 24 hr, because the degraded template is used as a primer for the incorporation of available nucleotides.

Compounds **11a–c** were also well incorporated using Vent(exo-) polymerase. However, the other compounds were less well incorporated using the same conditions (Table 1). In the case of Taq polymerase, the efficiency was significantly less appealing compared with Therminator or Vent polymerase (Table 1).

DNA Pol III polymerase is a member of the C family that shares no sequence homology with any of the other DNA polymerase families. In addition, it has proofreading capabilities that correct replication mistakes by means of exonuclease activity working in the 3′-5′ direction. Only one compound, **11c**, was recognized by the DNA Pol III polymerase α subunit and incorporated (Table 1).

The incorporation of more than one nucleotide was evaluated with Therminator polymerase. Full-length elongation was observed for compounds **11a-c**, **15c**, **16a**, and **17a**. The best results were obtained with the PPi-phosphonate **11b** (59.7% P+5 product) and **11c** (58% P+7 product) (Figure S7). The PPi-phosphonate compound **11a**, which contains the unnatural nucleobase, 5-chlorouracil, showed P+4 product (84%) as a major compound and a small amount of P+5 product (7%). Other compounds showed less efficient chain elongation properties, whereas stalling occurred after primer extensions with two or three nucleobases (Table S1).

The kinetic parameters for the incorporation of the fully modified nucleotide **16a** were determined on the basis of the single completed hit model (Creighton et al., 1995) (Table 2). Although the k_{cat} value of compound **16a** is only 9-fold lower than for the natural substrate dTTP, a large increase (190-fold) in the K_M value for the incorporation of **16a** was observed. As a result, the catalytic efficiency (k_{cat}/K_M) was decreased by a factor of 1,730.

Conformational Analysis of Monomeric Nucleotides

The synthesized phosphonate nucleosides differ from natural nucleosides by the presence of two anomeric centers i.e., C1′

Table 2. Kinetic Parameters for the Incorporation of dTTP and 16a by Therminator DNA Polymerase

Substrate	K_M (μM)	k_{cat} (min^{-1})	k_{cat}/K_M (min^{-1} μM^{-1})
dTTP	0.2768 ± 0.03915	219.78	794.00
16a	52.64 ± 4.86	24.14	0.4586

The incorporation experiments of dTTP and **16a** are carried out with primer-template complex P1T2 (P1: 5′-CAGGAAACAGCTATGAC-3′; T2: 3′-GTCCTTTGTCGATACTG**ACTGC**-5′) and Therminator DNA polymerase at 0.0005 U/μl (dTTP) or 0.00005 U/μl (**16a**).

and C4′. To analyze the impact of an additional anomeric center on the conformation of the nucleoside, the solution conformation of **10b** was determined. Because the magnitude of the anomeric effect is increased as the electron withdrawing character of the anomeric substituent increases (Paulsen et al., 1979), the anomeric effect of O4′-C4′-O4″ is stronger than that of O4′-C1′-N1. The anomeric effect at O4′-C4′-O4″ is driving O4″ into a pseudo-axial position (Mackman et al., 2007) (Figure 3A). Steric hindrance of the nucleobase in N-type conformation cooperates with the O4′-C4′-O4″ stereoelectronic effect to drive the nucleobase in a pseudo-equatorial orientation (Figure 3B). These predictions were confirmed by nuclear magnetic resonance (NMR) analysis in D_2O. From the observed $^3J_{HH}$, the predominant conformation of the sugar ring is described as P = 214° and φ = 60°. The 2′ *endo* conformation of the sugar ring is confirmed by nuclear Overhauser effect (NOE) interactions of thymine, H6 to H2′ of the sugar, whereas there was no observable interaction of the nucleobase with H3′ in the transverse rotating-frame Overhauser enhancement spectroscopy (T-ROESY) experiment (mixing time 200 ms; Figure S8).

Molecular Modeling of a XNA:DNA Duplex

It is difficult to obtain information on the stability of duplexes between DNA and phosphonate oligonucleotides, because the synthesis of phosphonate oligonucleotides with mixed bases is a tedious process and the procedure does not fulfill the requirements for routine oligomer synthesis (as is the case for the synthesis of phosphodiester-linked oligomers) (Páv et al., 2011). Therefore, we have performed a molecular modeling experiment to investigate the potential structure of these heteroduplexes. The structures of the antiparallel dodecamer duplex of phosphonate nucleoside analogs XNA:DNA as well as DNA dodecamer duplex as reference were examined using Amber software (Case et al., 2005). The root-mean-square (rms) deviation from a standard B-DNA double helix is shown in Figure S9. The hybrid XNA:DNA double helix ends in a stable conformation slightly different from a B-DNA helix (Figure 4).

The curves 5.3 analysis parameters (Schlick, 2010) are shown in Table S2 and highlight the major differences of the XNA:DNA duplex from a DNA:DNA duplex. The inclination of the bases seems not to be significantly different in the new helix. The twist goes slightly down by a few degrees. The helix does not, however, become wider (helix diameter) or longer (rise parameter).

The major groove gets a little bit wider and deeper, and the minor groove becomes a little wider but less deep. The most striking change, however, is the puckering parameter (calculated

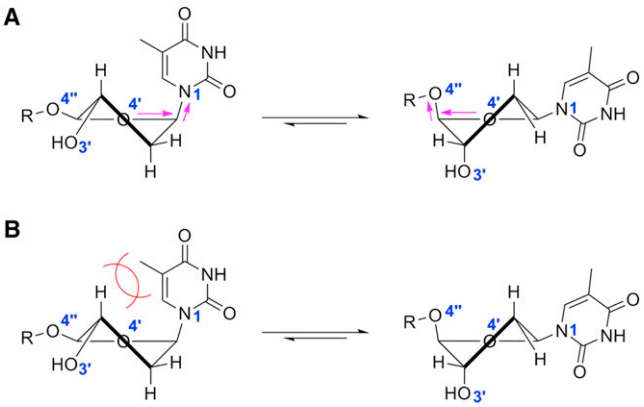

Figure 3. Stereoelectronic Effects that Drive the Conformation of the Sugar toward 3′*exo*/2′*endo*
(A) Anomeric effect of O4′-C4′-O4″.
(B) Steric hindrance of the nucleobase.

using PTRAJ in Amber10 software). The sugar puckering of nucleotides in a DNA duplex fluctuates between C1′*exo* and C2′-*endo* (Privé et al., 1991). In the XNA:DNA duplex, a different pucker is observed in the XNA strand than in the DNA strand (Table S3). In the DNA strand, puckering is quasi constant along the chain (Southeast to South). The XNA:DNA duplex induces a change in the pucker of the XNA units along the duplex: a change from p = 132° (Southeast region) to 50° (Northeast region) and back to 117° (East region) along the duplex.

The electrostatic potential around both duplexes was calculated using APBS (Baker et al., 2001) and then projected on the molecular surface (Pettersen et al., 2004) (Figure 5). In the normal dsDNA, a more negative potential is seen clearly in the minor groove. The wider major groove has a less negative potential. In the XNA:DNA duplex, this difference between major and minor grooves gets less well-defined, probably because the minor groove widens.

The radial distribution function of the solvent molecules around O5′ (normal DNA) and O4′ (XNA chain) is shown in Figure S10. In the XNA chain, the O4′ atom is more exposed to the solvent-making formation of hydrogen bonds with solvent water molecules. In the normal dsDNA and in the DNA strand in the hybrid XNA:DNA duplex, this O5′ atom is close to the phosphate, making it less accessible to the solvent.

DISCUSSION

In the present study, we investigated the substrate properties of tetrahydrofuranyloxy nucleotides with different leaving groups and different base moieties. The phosphonate group has the advantage that, in vivo, no dephosporylation by phosphatases is possible and no phosphorylation by kinases is needed (as would be the case when using nucleoside analogs instead of nucleotide analogs). Once the oligonucleotide is synthesized, its stability against enzymatic degradation by phosphodiester cleavage would be increased when compared with natural DNA. The synthetic challenge to prepare the monomers lies in the selective glycosylation reaction of a compound with two anomeric centers. This goal has been achieved based on

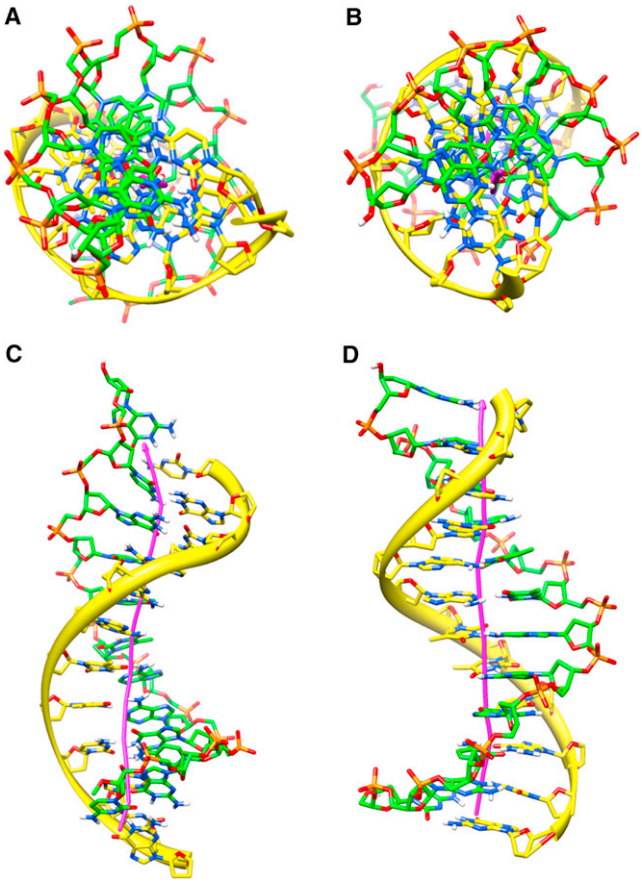

Figure 4. Structures of the XNA:DNA and dsDNA 12-mer Duplexes after 20 ns of Simulation Time

XNA:DNA duplex is very similar to the dsDNA duplex, except for the wider major groove and a different sugar puckering (see text). Top (A) and side (C) view of XNA:DNA. Top (B) and side (D) view of dsDNA. Figures are generated using Chimera.

See also Tables S2 and S3.

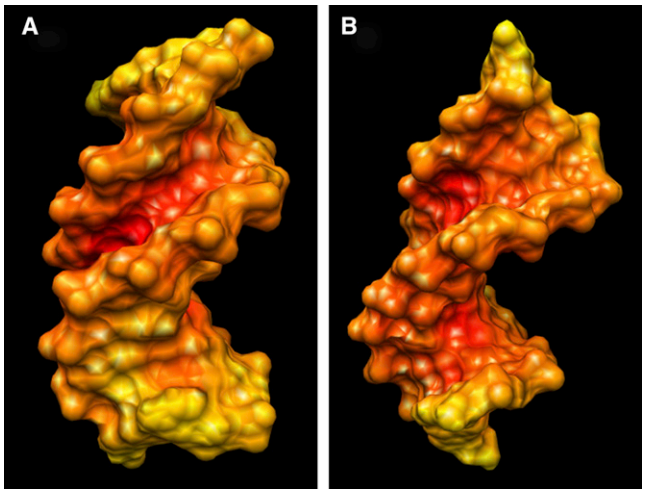

Figure 5. Electrostatic Potential Projected on the Connolly Surface of the XNA:DNA and dsDNA Duplexes

(A) Potential around the dsDNA duplex.

(B) Potential of the XNA:DNA duplex.

The electrostatic potential in the minor groove of XNA:DNA is more negative than in the major groove. In XNA:DNA the difference is less pronounced. Colors indicate potential values in dimensionless units of k_bT/e_c, k_b is the Boltzmann constant, T is temperature in K, and e_c is the charge of an electron. A red color means −16, yellow corresponds with −5.

different leaving group properties (in the presence of Lewis acids) of the anomeric substituents and the presence of an OBz group at the 3′-position of the sugar moiety to tune reactivity. Three different bases have been evaluated: one natural purine base (A), one natural pyrimidine base (T), and one synthetic base (5-chlorouracil) (Marlière et al., 2011). The leaving groups that have been investigated are the natural pyrophosphate (**11a–c**), the modified nucleotide itself (**15a–c**) (Song et al., 2011b), iminodipropionate (**16a–c**), L-aspartic acid (**17a–c**), and L-phosphonoalanine (**18**). It is clear that pyrophosphate remains the best leaving group, in particular when considering HIV-RT as a catalyst for the polymerization reaction. However, several thermostable polymerases, and in particular Therminator polymerase, also accept the amino acids as a leaving group for the enzymatic synthesis of the artificial nucleic acids with a phosphonate backbone. The efficient incorporation and chain elongation of **16a** (and to a somewhat lesser extent of **17a**) is striking. Kinetic analysis for the incorporation of compound **16a**, in comparison with dTTP, revealed that the decreased catalytic efficiency is mainly due to a higher K_M value (190-fold), whereas k_{cat} is only 9-fold lower. Although **16a** was incorporated rather efficiently into a growing DNA strand, the affinity of the enzyme for **16a** was substantially reduced compared with the natural substrate.

These are examples of fully modified nucleotides (with altered base, sugar, and leaving group) that are accepted as substrates in a DNA template-based incorporation assay. This is an important observation with respect to the selection of orthogonal chemistry as a potential artificial information system in vivo.

We have also studied the structure of an XNA:DNA duplex by molecular modeling. A 2′-*endo* conformation of the sugar is observed in the monomeric form. Structural differences between B-form dsDNA and the XNA:DNA duplex are minor variations in twist, roll, width, and depth of the grooves and a change in the electrostatic potential of the grooves, further pointing to the potential orthogonality of both structures. The highest variability is observed in the sugar puckering that is Southeast to South in the DNA strand of the XNA:DNA duplex and varies between Northeast and Southeast in the XNA strand. The presence of two anomeric centers may contribute to the conformational flexibility of the tetrahydrofuranyloxy nucleotides.

When Therminator polymerase was used for the incorporation study of **11a**, the gel shift experiments showed double spots at the P+2 level. The compounds that have been formed in the reaction mixture were analyzed by mass spectrometry, which led us to demonstrate the reversibility of the polymerization reaction. In addition to polymerization, DNA polymerases may catalyze several other reactions such as pyrophosphorolysis, which lead to the degradation of DNA and the formation of deoxynucleoside triphosphates (dNTPs). During the polymerization reaction, deoxynucleoside monophosphates are

incorporated in the growing nucleic acid chain with release of pyrophosphate. As the reactions are carried out in the absence of pyrophosphatases, PPi generated in the reaction mixture gives rise to pyrophosphorolysis. The PPi attacks the 3′ terminal of the template, releasing dGTP and dTTP, and the template is degraded by two nucleotides (Figure S5). The degraded template becomes primer and vice versa, and compound **10a** is incorporated into the degraded template (which became a sticky end) and followed by another **10a** or dTMP. It should be mentioned that incorporation and pyrophosphorolysis do not take place in the absence of the primer, which is in agreement with previous experiments showing that the presence of a template or a primer alone is not sufficient to sustain pyrophosphorolysis (Deutscher and Kornberg, 1969). Obviously, this phenomenon was not observed when using the substrates with alternative leaving groups. All these data reconfirm the importance of pyrophosphorolysis in the irreversibility of the polymerase reaction. The sequence-selective chemical cleavage approach (Maxam and Gilbert, 1977) might find its enzymatic counterpart in a sequence-selective enzymatic cleavage approach (sequencing by degradation with a base-specific exonuclease activity based on using different PPi mimics). In this context, the observation that a change of the leaving group of dNTPs leads to a base selective incorporation of nucleotides might be an interesting observation (Yang et al., 2011), although, in the present case, we did not observe the reversed reaction, i.e., a template degradation reaction using the pyrophosphate mimics and the investigated polymerases.

SIGNIFICANCE

By inverting the ordering of the carbon and oxygen atom of the 4′-CH_2OP group in natural nucleotides to a 4′-OCH_2P group, a phosphonate nucleoside is obtained with two anomeric centers. This sugar moiety can be considered as derived from tartaric dialdehyde, an unstable compound. A synthetic scheme was developed to prepare these nucleotide analogs, making use of differences in reactivity of the anomeric leaving groups. Based on the molecular modeling experiments of a 4′-inverse DNA:DNA duplex, the sugar moiety of these phosphono-nucleosides may take conformations situated between the extreme North and South forms, whereas the electrostatic potential of the grooves is altered when compared with dsDNA. The diphosphophosphonate nucleosides with an adenine, thymine, and 5-chlorouracil base are excellent substrates for several polymerases. When changing the leaving group to iminodipropionate, aspartate and phosphonoalanine, thermostable polymerases [Therminator, Vent(exo⁻), and Taq polymerase] still accepts the modified nucleotides as substrates. Therefore, the first instance of a nucleoside triphosphate mimic in which the base, backbone motif, and pyrophosphate leaving group found in the natural compounds have been replaced by unnatural chemical entities and that could still function as substrates for polymerases has been elaborated. This is an important step in view of developing artificial information systems for in vivo applications. This result stands as a significant milestone in the development of an artificial genetic memory in living cells.

EXPERIMENTAL PROCEDURES

DNA Polymerase Reactions

Oligonucleotides P1 and T1–T4 were purchased from Sigma Genosys. The concentrations were determined with a Varian Cary-300-Bio UV Spectrophotometer. The primer was 5′-^{33}P-labeled with 5′-[γ-^{33}P]-ATP (Perkin Elmer) using T4 polynucleotide kinase (New England BioLabs) according to the procedure provided by the supplier. The labeled oligonucleotide was further purified using Illustra Microspin G-25 columns (GE Healthcare).

End-labeled primer was annealed to its template by combining primer and template in a molar ratio of 1:2 and heating the mixture to 75°C for 10 min and followed by slow cooling to room temperature over a period of 1.5 hr. For the incorporation of compounds **11a–c**, **15a–c**, **16a–c**, **17a–c**, and **18**, a series of 20 μl batch reactions were performed with the enzyme [HIV-1 RT, pol III, Taq, Vent(exo⁻), and Therminator DNA polymerase]. The final mixture contained 125 nM primer template complex, reaction buffer (50 mM Tris HCl, 50 mM KCl, 10 mM $MgCl_2$, 0.5 mM spermidine, 10 mM dithiothreitol [DTT]; pH 8.3 for HIV-1 RT; 20 mM Tris HCl, 10 mM KCl, 2 mM $MgSO_4$, 0.1% Triton X-100, pH 8.8 for thermostable polymerase; 20 mM Tris HCl, 10 mM $MgCl_2$, 10 mM DTT, 20 μg/ml bovine serum albumin, 4% glycerol, pH 7.5 for polymerase III), appropriate concentration of enzyme, and different concentrations of modified nucleotide building blocks (1 mM, 500 μM, 200 μM, and 100 μM). In the control reaction, 10 μM of dATP or dTTP were used as reference. The mixture was incubated at 37°C or 75°C, respectively, and aliquots were quenched after 10, 20, 30, 60, and 120 min.

SUPPLEMENTAL INFORMATION

Supplemental Information includes ten figures, four tables, and Supplemental Experimental Procedures and can be found with this article online at http://dx.doi.org/10.1016/j.chembiol.2013.02.010.

ACKNOWLEDGMENTS

We thank KU Leuven (GOA, IDO), EC (Metacode), and FWO for financial support and the Hercules Foundation of the Flemish Government (grant 20100225–7) for mass spectrometry. The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013)/ERC grant agreement ERC-2012-ADG_20120216/ 320683.

Received: September 10, 2012
Revised: January 15, 2013
Accepted: February 6, 2013
Published: March 21, 2013

REFERENCES

Abraham, T.W., Kalman, T.I., McIntee, E.J., and Wagner, C.R. (1996). Synthesis and biological activity of aromatic amino acid phosphoramidates of 5-fluoro-2′-deoxyuridine and 1-beta-arabinofuranosylcytosine: evidence of phosphoramidase activity. J. Med. Chem. *39*, 4569–4575.

Adelfinskaya, O., and Herdewijn, P. (2007). Amino acid phosphoramidate nucleotides as alternative substrates for HIV-1 reverse transcriptase. Angew. Chem. Int. Ed. Engl. *46*, 4356–4358.

Anderson, R.S., Bollum, F.J., and Beattie, K.L. (1999). Pyrophosphorolytic dismutation of oligodeoxy-nucleotides by terminal deoxynucleotidyltransferase. Nucleic Acids Res. *27*, 3190–3196.

Baker, N.A., Sept, D., Joseph, S., Holst, M.J., and McCammon, J.A. (2001). Electrostatics of nanosystems: application to microtubules and the ribosome. Proc. Natl. Acad. Sci. USA *98*, 10037–10041.

Bebenek, K., Abbotts, J., Wilson, S.H., and Kunkel, T.A. (1993). Error-prone polymerization by HIV-1 reverse transcriptase. Contribution of template-primer misalignment, miscoding, and termination probability to mutational hot spots. J. Biol. Chem. *268*, 10324–10334.

Case, D.A., Cheatham, T.E., 3rd, Darden, T., Gohlke, H., Luo, R., Merz, K.M., Jr., Onufriev, A., Simmerling, C., Wang, B., and Woods, R.J. (2005). The Amber biomolecular simulation programs. J. Comput. Chem. *26*, 1668–1688.

Creighton, S., Bloom, L.B., and Goodman, M.F. (1995). Gel fidelity assay measuring nucleotide misinsertion, exonucleolytic proofreading, and lesion bypass efficiencies. Methods Enzymol. *262*, 232–256.

Deutscher, M.P., and Kornberg, A. (1969). Enzymatic synthesis of deoxyribonucleic acid. 28. The pyrophosphate exchange and pyrophosphorolysis reactions of deoxyribonucleic acid polymerase. J. Biol. Chem. *244*, 3019–3028.

Dhavale, D.D., Tagliavini, E., Trombini, C., and Umani-Ronchi, A. (1988). Novel synthetic equivalents of differentially protected tartaric aldehydes. A simple route to useful c-4 chiral synthons. Tetrahedron Lett. *29*, 6163–6165.

Epp, J.B., and Widlanski, T.S. (1999). Facile Preparation of Nucleoside-5′-carboxylic Acids. J. Org. Chem. *64*, 293–295.

Eschenmoser, A. (1999). Chemical etiology of nucleic acid structure. Science *284*, 2118–2124.

Földesi, A., Trifonova, A., Dinya, Z., and Chattopadhyaya, J. (2001). Total synthesis of 2′,3′,4′,5′ ’-(2)h(5)-ribonucleosides: the key building blocks for NMR structure elucidation of large RNA. J. Org. Chem. *66*, 6560–6570.

Herdewijn, P., and Marlière, P. (2009). Toward safe genetically modified organisms through the chemical diversification of nucleic acids. Chem. Biodivers. *6*, 791–808.

Hoffer, M. (1960). α-Thymidin. Chem. Ber. *93*, 2777–2781.

Hu, G. (1993). DNA polymerase-catalyzed addition of nontemplated extra nucleotides to the 3′ end of a DNA fragment. DNA Cell Biol. *12*, 763–770.

Kim, C.U., Luth, B.Y., and Martin, J.C. (1991). Regiospecific and highly stereoselective electrophilic addition to furanoid glycals: synthesis of phosphonate nucleotide analogues with potent activity against HIV. J. Org. Chem. *56*, 2642–2647.

Koh, Y.H., Shim, J.H., Wu, J.Z., Zhong, W., Hong, Z., and Girardet, J.L. (2005). Design, synthesis, and antiviral activity of adenosine 5′-phosphonate analogues as chain terminators against hepatitis C virus. J. Med. Chem. *48*, 2867–2875.

Mackman, R.L., Zhang, L., Prasad, V., Boojamra, C.G., Douglas, J., Grant, D., Hui, H., Kim, C.U., Laflamme, G., Parrish, J., et al. (2007). Synthesis, anti-HIV activity, and resistance profile of thymidine phosphonomethoxy nucleosides and their bis-isopropyloxymethylcarbonyl (bisPOC) prodrugs. Bioorg. Med. Chem. *15*, 5519–5528.

Marlière, P., Patrouix, J., Döring, V., Herdewijn, P., Tricot, S., Cruveiller, S., Bouzon, M., and Mutzel, R. (2011). Chemical evolution of a bacterium’s genome. Angew. Chem. Int. Ed. Engl. *50*, 7109–7114.

Maxam, A.M., and Gilbert, W. (1977). A new method for sequencing DNA. Proc. Natl. Acad. Sci. USA *74*, 560–564.

Nakatani, K., Sando, S., Kumasawa, H., Kikuchi, J., and Saito, I. (2001). Recognition of guanine-guanine mismatches by the dimeric form of 2-amino-1,8-naphthyridine. J. Am. Chem. Soc. *123*, 12650–12657.

Paulsen, H., Luger, P., and Heiker, F.R. (1979). Aspects of conformational analysis of pentopyranosyl acetates, benzoates, and halides. In Anomeric Effect: Origin and Consequences, (ACS Symposium Series 87) (Washington: American Chemical Society), pp. 63–79.

Páv, O., Košiová, I., Barvík, I., Pohl, R., Buděšínský, M., and Rosenberg, I. (2011). Synthesis of oligoribonucleotides with phosphonate-modified linkages. Org. Biomol. Chem. *9*, 6120–6126.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004). UCSF Chimera—a visualization system for exploratory research and analysis. J. Comput. Chem. *25*, 1605–1612.

Privé, G.G., Yanagi, K., and Dickerson, R.E. (1991). Structure of the B-DNA decamer C-C-A-A-C-G-T-T-G-G and comparison with isomorphous decamers C-C-A-A-G-A-T-T-G-G and C-C-A-G-G-C-C-T-G-G. J. Mol. Biol. *217*, 177–199.

Schlick, T. (2010). Nucleic acids structure minitutorial. In Molecular Modeling and Simulation: An Interdisciplinary Guide, Second Edition (New York: Springer), pp. 129–162.

Schneider, A., Rodrigues, O.E.D., Paixão, M.W., Appelt, H.R., Braga, A.L., and Wessjohann, L.A. (2006). Stereoselective synthesis of Boc-protected l-seleno- and tellurolanthionine, l-seleno- and tellurocystine and derivatives. Tetrahedron Lett. *47*, 1019–1021.

Song, X.P., Bouillon, C., Lescrinier, E., and Herdewijn, P. (2011a). Iminodipropionic acid as the leaving group for DNA polymerization by HIV-1 reverse transcriptase. ChemBioChem *12*, 1868–1880.

Song, X.-P., Maiti, M., and Herdewijn, P. (2011b). Enzymatic synthesis of DNA employing pyrophosphate-linked dinucleotide substrates. J. Syst. Chem. *2*, 3.

Yang, S., Froeyen, M., Lescrinier, E., Marlière, P., and Herdewijn, P. (2011). 3-Phosphono-L-alanine as pyrophosphate mimic for DNA synthesis using HIV-1 reverse transcriptase. Org. Biomol. Chem. *9*, 111–119.