Template-Independent Enzymatic Oligonucleotide Synthesis (TiEOS): Its History, Prospects, and Challenges

Michael A. Jensen*†‡ and Ronald W. Davis†‡

† Stanford Genome Technology Center, Department of Biochemistry, Stanford University, Palo Alto, California 94304, United States
‡ Department of Genetics, Stanford University, Palo Alto, California 94304, United States

ABSTRACT: There is a growing demand for sustainable methods in research and development, where instead of hazardous chemicals, an aqueous medium is chosen to perform biological reactions. In this Perspective, we examine the history and current methodology of using enzymes to generate artificial single-stranded DNA. By using traditional solid-phase phosphoramidite chemistry as a metric, we also explore criteria for the method of template-independent enzymatic oligonucleotide synthesis (TiEOS). As its key component, we delve into the biology of one of the most enigmatic enzymes, terminal deoxynucleotidyl transferase (TdT). As TdT is found to exponentially increase antigen receptor diversity in the vertebrate immune system by adding nucleotides in a template-free manner, researchers have exploited this function as an alternative to the phosphoramidite synthesis method. Though TdT is currently the preferred enzyme for TiEOS, its random nucleotide incorporation presents a barrier in synthesis automation. Taking a closer look at the TiEOS cycle, particularly the coupling step, we find it is comprised of additions > n+1 and deletions. By tapping into the physical and biochemical properties of TdT, we strive to further elucidate its mercurial behavior and offer ways to better optimize TiEOS for production-grade oligonucleotide synthesis.

In the mid-20th century, there were several key breakthroughs in the fields of genetics and biochemistry that have led us to current day medicine. This was a cascade of events that started with X-ray-induced gene knockout studies in 1941, which made the connection that genes were directly involved in enzyme function.1 It soon followed that genes themselves were comprised of nucleic acids (DNA),2 and a double helix was an orderly structure of nucleic acids that stored genetic information3 and could be precisely replicated by a DNA polymerase.4

It was then understood the power to manipulate such biological systems was harnesses in the form of single-stranded (ss) DNA with a defined sequence.5 As such, a great deal of work in the 1950s went toward developing methods for chemically synthesizing polynucleotides from nucleic acid monomers.6−8 At about the same time, the potential to synthesize polynucleotides enzymatically was also being realized.9 However, nucleotide (nt) addition was limited to homopolymeric tract formation, and efforts were eclipsed by the chemical method, which allowed synthesis in a controlled, stepwise manner.10 Its development continued over the next 30 years,10,11 which culminated in the traditionally accepted solid-phase phosphoramidite method used on all major synthesis platforms to date (Figure 1).

One of the major drawbacks with the phosphoramidite method is the use of hazardous chemicals described in each step of Figure 1. These are listed by the U.S. Environmental Protection Agency for community awareness and emergency planning.13 Since the 1990s, there has been a growing trend toward “green” technology development in terms of sustainable chemistry.14 Possible advantages of using an enzyme for ssDNA synthesis also include (i) longer strand generation,15 (ii) a lower error rate,16 (iii) a faster cycle time, and (iv) a lower cost of production. These have very important implications in many fields of research, particularly in the areas of synthetic biology17 and DNA data storage.18,19 Another advantage of oligonucleotide production in an aqueous medium is portability. This comes into play especially in diagnostics and point-of-care devices,20,21 as well as for in-field applications from forensics to synthesis in space (stanfordssi.org).

Until recently, the use of enzymes to make ssDNA continued a very slow and unproductive journey alongside the chemical method. In this Perspective, we first evaluate the origins of enzymatic oligonucleotide synthesis; and with the knowledge gained from prior research, we explore the method defined as template-independent enzymatic oligonucleotide synthesis (TiEOS). We also observe parallels with the solid-phase phosphoramidite cycle as a metric for quality assessment to gain a better appreciation for the prospects and challenges inherent with using an enzyme to make ssDNA. From a biological point of view, we also pay particular attention to how additions > n+1 and deletions might be generated and controlled during the TiEOS cycle.

Received: September 20, 2017
Revised: February 27, 2018

DOI: 10.1021/acs.biochem.7b00937
Biochemistry XXXX, XXX, XXX–XXX
ENZYMES AND METHODOLOGIES USED FOR OLGONUCLEOTIDE SYNTHESIS: A BRIEF HISTORY

In 1955, Severo Ochoa discovered polynucleotide phosphorylase (PNPase) isolated from the microorganism Azotobacter vinelandii.9 Believed to be a component of RNA metabolism, PNPase was used in vitro to synthesize polyribonucleotides (Figure 2).

In 1959, F. J. Bollum described the first ssDNA polymerase, terminal deoxynucleotidyl transferase (TdT), capable of template-independent (primer-free) synthesis24,25 (Figure 3). Before characterization of this enzyme, we understood DNA polymerization to be constrained to primer-template duplex-driven synthesis.

In 1962, Bollum confirmed dNTPs were being added to the initiator at the 3′ oxygen, because blocking it with an acetyl group prevented further nucleotide addition. As a consequence, Bollum proposed TdT could be used for ssDNA polymerization with monomers blocked at the 3′ position for stepwise synthesis of oligonucleotides with a defined sequence.26 In polymerization to be constrained to primer—template duplex-driven synthesis.

In 1962, Bollum confirmed dNTPs were being added to the initiator at the 3′ oxygen, because blocking it with an acetyl group prevented further nucleotide addition. As a consequence, Bollum proposed TdT could be used for ssDNA polymerization with monomers blocked at the 3′ position for stepwise synthesis of oligonucleotides with a defined sequence.26 In

$$(pdN)_{n+3} + n ppdN \xrightarrow{TdT} (pdN)_{n+3} + n pp$$

Figure 3. TdT polymerization of ssDNA. Here, a 3′-hydroxylated initiator strand (also termed an acceptor) is required (in the presence of a divalent cation (e.g., Mg$^{2+}$)) for TdT to catalyze ssDNA using a deoxyribonucleoside triphosphate monomer (dTTP, also termed a donor). Here dNTP (ppdN) monomers are the substrate for addition of TdT to the 3′ end of an initiator strand (pdN)$_n$, to generate ssDNA (pdN)$_{n+3}$ with the release of pyrophosphate (pp). 3$^*$ indicates the length must be at least 3 nt. There are two primary isoforms of the TdT gene, TdTL (3′ exonuclease activity only, arrow right to left) and TdTS (3′ terminal transferase only, arrow left to right).
1965, Letsinger and Mahadevan reported the first chemical DNA synthesis (dCpT) on a polymer support.\textsuperscript{27} Now with the potential to anchor the starting material (initiator) to a solid substrate, single monomer additions could be repeated on a per cycle basis, and unincorporated substrate could be washed from the reaction mixture between steps without loss of initiator.

In 1971, Mackey and Gilham used PNPase to synthesize RNA of a defined sequence; via the introduction of nucleotides blocked at the 2′ (3′) end, it was theorized RNA monomers could also be added to an initiator in a controlled stepwise manner.\textsuperscript{28} Here, they proposed a blocking group should (i) be chemically stable, (ii) have conditions for removal that are compatible with the enzymatic reaction, (iii) be small enough for the enzyme to be fully incorporated, and (iv) completely block any further enzymatic nucleotide additions. To demonstrate this, they successfully coupled 5′-diphosphate-2′O-(α-methoxyethyl)-uridine to the 3′ end of an adenosine trinucleotide initiator in the presence of PNPase (Figure 4). A year later, Gilham and Smith would apply the same principle to generate ssDNA.\textsuperscript{29}

![Figure 4. PNPase synthesis of polyribonucleotides of defined sequence.](image)

**Figure 4.** PNPase synthesis of polyribonucleotides of defined sequence. (A) Ribonucleotide representation, either adenosine (A) or uridine (U), diphosphate (pp), with 2′ (3′)-O-(α-methoxyethyl) blocking group (ME). (B) Solution-phase enzymatic RNA synthesis with 2′-protected ribonucleotides. (I) Adenosine trinucleotide (initiator) is coupled to 5′ diphosphate uridine protected at the 2′ position with α-methoxyethyl in the presence of PNPase at 37°C for 7 h to generate 2′-protected tetranucleotide. (II) The α-methoxyethyl protecting group is removed with acid [pH 2 (3) for 15 min]. (III) The next nucleotide, 2′-protected 5′ ADP, is coupled to the tetranucleotide initiator (conditions are the same as those in step I). (IV) The structure of a blocked pentanucleotide is then confirmed by hydrolysis with pancreatic ribonuclease.

In 1978, England and Uhlenbeck introduced the first ligation method by which oligoribonucleotide synthesis was performed with T4 RNA ligase using 5′, 3′ ribonucleoside bisphosphates (prNp).\textsuperscript{30} In 1984, Schott and Schrade reported single-step oligonucleotide synthesis with TdT,\textsuperscript{31} during which they added unprotected dNTPs to initiator strands of variable length, and in 1999, the first report of solid-phase enzymatic DNA synthesis was made using T4 RNA ligase (Figure 5).\textsuperscript{32} This demonstrated the basic repeating synthesis cycle whereby protected monomers could be added in a stepwise manner and then deblocked at the 3′ end for further additions of defined sequence; as the initiator strand remains tethered without loss during wash steps, it is finally released by enzymatic cleavage postsynthesis.

Despite the capability of PNPase to synthesize both ribo- and deoxyribonucleotides, limitations outweigh its benefits. (i) Purification is an extremely difficult, multistep process. (ii) PNPase has a preference for purine-rich initiators, which may limit sequence design. (iii) As the primer molecules become longer, accumulation of orthophosphate drives the reaction backward in phosphorolysis (strand degradation) (Figure 2).\textsuperscript{33}

Advantages of using T4 RNA ligase as a method of oligonucleotide production are also eclipsed by its limitations. (i) Initiator strands with uracil are less reactive. (ii) Ligation times are very long (≤144 h).\textsuperscript{32} (iii) The rate of T4 RNA ligase activity decreases after incubation at 37°C.\textsuperscript{30} (iv) If free unblocked nucleotides are present in the reaction mixture, monomers will react with themselves and/or add to the initiator to generate homopolymeric tracts;\textsuperscript{34} this can substantially reduce the target product yield if reactions favor synthesis of unblocked donor generation. (v) If the unblocked donor is ≥8 nt, intramolecular cyclization may occur,\textsuperscript{35} thus competing with acceptor (initiator) polymerization.

Though terminal transferase would prove to be a far better candidate for enzymatic oligonucleotide synthesis compared to its predecessors, it is still less than perfect. The remainder of this Perspective delves into its role in the synthesis cycle, with a particular focus on types and causes of failure strand generation.

### TERMINAL DEOXYNUCLEOTIDYL TRANSFERASE (TdT)

TdT is part of the X Family of low-fidelity DNA polymerases that includes, Pol β, λ, μ, σ1, and δ.\textsuperscript{36} Pol β is involved in base excision repair (BER) and nonhomologous end joining (NHEJ), with Pol λ, μ, and TdT being active in NHEJ and V(D)J recombination in eukaryotes.\textsuperscript{37} The primary function of TdT is to increase antigen receptor diversity through random nucleotide incorporation in the vertebrate adaptive immune system.
system. This creates a vast catalog of immunoglobulins (10^{14}) and T-cell receptors (10^{18}) that recognize and target most foreign pathogens encountered during an organism’s life span.\textsuperscript{38} In general, the terminal transferase gene (e.g., human) has two splice variants, TdTL (long amino acid sequence, associated with only 3' → 5' exonuclease activity) and TdTS (short amino acid sequence, acts solely as a terminal transferase).\textsuperscript{39-41} It is understood in vivo nucleotide additions and deletions during antigen receptor gene rearrangement are modulated by the co-expression of these two variants. For the remainder of this Perspective, we will be referring to the isoform, TdTS, which is typically isolated (e.g., from calf thymus)\textsuperscript{42} for use in a number of biological applications, including TUNEL, RACE, modified end labeling, and homopolymer tailing.

As TdT is the only enzyme fully dedicated to template-independent synthesis, both Pol λ and μ partially share these properties but are not suitable for synthesis automation. While processivity is most active with dsDNA binding and extension, Pol λ and μ can accommodate single-strand addition to some degree. For example, Pol λ requires a 3' overhang (>9 nt) to perform template-independent synthesis, and polymerization occurs with only certain sequences.\textsuperscript{45} Pol μ, which shares 41% amino acid identity with TdT, is a dual-mode DNA polymerase where it acts either by polymerization extension (template-dependent) or by randomly adding nucleotides across from an abasic site in a sequence-independent nucleotidyl transferase manner.\textsuperscript{46}

### TEMPLATE-INDEPENDENT ENZYMATIC Oligonucleotide SYNTHESIS (TiEOS)

To develop a cycle of repeating steps for enzymatic ssDNA synthesis in an automated capacity (Figure 6), we first need to consider criteria gleaned from past history exploring various enzymes and substrates used to generate polynucleotides. These necessary components include (i) a solid substrate for retaining the polynucleotide during synthesis, (ii) an initiator strand (>3 nt) for enzyme binding and polymerization, (iii) an enzyme that is template-independent, (iv) protected nucleotides for controlled, stepwise addition, and (v) a means of final product release from the support. In doing this, we define the method as template-independent enzymatic oligonucleotide synthesis (TiEOS) (Figure 6).

#### ADDITIONS > N+1

In living polymerization,\textsuperscript{48,49} where “chain ends remain active until killed”, TdT can generate a homopolymer tract of up to 8000 nt.\textsuperscript{51} It was determined that the length of the product generated from the 3' end of the initiator is proportional to the monomer to initiator ratio (M/I).\textsuperscript{50} As such, the extent of polymerization can be controlled by combining a higher M/I ratio and a longer incubation period so that the initiator is fully extended. For our purposes, we are primarily interested in a controlled, stepwise nucleotide addition in the TiEOS cycle. Once added to the initiator strand, the 3’ region of a single dNTP_{Pr} [Pr is a protecting group (see Figure 8)] is protected from further extension until it is selectively converted to the reactive hydroxyl group (dNTP_{Pr} → dNTP); as such, dNTP_{Pr} is also termed the reversible terminator (RT). Because RT stability, efficiency of TdT incorporation, and completeness of deprotection are all major factors affecting TiEOS product quality and final yield, we next focus on how additions > n+1 are generated in the TiEOS cycle.

As a metric of quality assessment, it is important to compare TiEOS with the chemical method in terms of stepwise synthesis. Using step 2 in Figure 1 as a reference, the incoming nucleoside phosphoramidite is coupled (in the presence of an activator) to the first base preattached to the solid support. Here, DMT at the 5’ position of each phosphoramidite monomer prevents homopolymeric tract formation. If, however, premature detritylation occurs due to residual acid leftover from step 1 (inadequate washing) or an acidic activator [5-ethylthio-1H-tetrazole pK_{a} (5.2) > 4, 5-dicyanoimidazole pK_{a} (4.28)\textsuperscript{51}], product > n+1 may be generated. Side reactions can also occur during synthesis at the internucleotide linkage and nucleobases themselves. To prevent this, the phosphor oxygens and nucleobase exocyclic amines are blocked by cyanoethyl, and isobutyl or benzoyl protecting groups, respectively (Figure 1A,B).

In TiEOS, strands > n+1 can be generated by unblocked nucleotides during the couple step. Figure 7 demonstrates the pattern observed for uncontrolled homopolymeric tract formation and may be representative of a TiEOS cycle reaction contaminated by unblocked nucleotides.\textsuperscript{34}

If the dNTP_{Pr} reservoir contains hydroxylated material, multiple dNTPs could be added consecutively onto the unblocked initiator (n). This can be a very complicated problem in terms of (i) percent homopolymeric tract formation and (ii) nucleotide additions and deletions during synthesis. However, this mechanism is rare and would only occur if the reaction is not properly controlled and monitored with an appropriate terminator (RT) system.
In 1994, Metzker and colleagues developed a set of reversible terminators as a gel-free alternative to the Sanger sequencing method, which includes 3′-O-methyl-dTTP (dATP) and 3′-O-(2-nitrobenzoyl)-dTTP (Figure 8B,C).52 Here they demonstrated these RTs could arrest template-dependent polymerase activity of Bst and AmpliTaq. Knapp et al. also reported fluorescently labeled RTs blocked with 2-cyanoethyl (Figure 8D) prevented further nucleotide addition by TdT.53 Ju et al. also discussed development of fluorescently labeled RT, allyl (Figure 8E), for use in sequencing by synthesis (SBS) reactions.54 Chen et al. also offered 3′-O-allyl, hydroxyamine, and azidomethyl (Figure 8E–G, respectively) as alternative RTs for SBS.55

With respect to Figure 8, there are two RT options that may be of particular interest for the TiEOS application. The first is photocleavable 2-nitrobenzyl (Figure 8C), which has been shown to cause TdT arrest after its incorporation onto blunt-ended, duplex DNA.56 This could be very beneficial in terms of high-throughput synthesis automation; however, at the time of this writing, only the dATP monomer with 2-nitrobenzyl is available for purchase (TriLink Biotechnologies). The second RT with strong potential in TiEOS is 3′ hydroxyamine (Figure 8F),57 which presents with mild reagent deblocking conditions using aqueous sodium nitrite and is currently available for all dNTPs (Firebird Biomolecular Sciences).

Moreover, the native form of TdT appears to allow some incorporation of these 3′-blocked monomers (Figure 8) as shown with 2-nitrobenzyl.56 Boulé et al. also observed TdT had an only weak preference for dNTP over 2′ OH NTP (RNA) polymer growth of the dA9 initiator.58 Winz et al. also used TdT to label DNA through copper-catalyzed azide–alkyne cycloaddition.59 Even in our own lab, we demonstrated that TdT incorporates 3′-tert-butoxy ethoxy-dTTP (dTTP-TBE)60 (Figure 9).

Figure 9 shows possible distributive properties of TdT over time (5, 10, and 15 min).61 Percent conversion n → n+1 for each sample was calculated at 22% (I), 52% (II), and 55% (III) (based on the percent area for each target peak using chromatogram analysis reports). Reactions were stopped with ethylenediaminetetraacetic acid (EDTA), which is a cation chelating agent and, therefore, shuttles away metal ions necessary for polymerase activity. As indicated in sample II, the shoulder peak may represent a second dTTP addition, same as III with +2 and +3 additions. Because heat-labile dNTP monomers (e.g., dNTP-TBE) were specifically developed for use in hot-start polymerase chain reaction (PCR) protocols,62 TBE will be lost at high temperatures; also, dNTP-TBE to dNTP conversion is on a sliding scale, and even at 37 °C, TBE protection is compromised, especially over longer periods of incubation (Figure 9, II and III at 10 and 15 min). This is also a good example of possible unblocked dNTP contamination during the TiEOS coupling step (compare with the pattern in Figure 7).

As for any modified stock dNTP product, trace contamination may be present, which cannot be filtered out 100% by
Table 1. Examples of Random TdT Nucleotide Incorporation/Polymerization

<table>
<thead>
<tr>
<th>Method</th>
<th>Sequence</th>
<th>Efficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Solution-phase polymerization</td>
<td>G &gt; A &gt; T &gt; C</td>
<td>62%</td>
</tr>
<tr>
<td>Solid-phase polymerization</td>
<td>T &gt; C &gt; A &gt; G</td>
<td>67%</td>
</tr>
<tr>
<td>Influence of initiator 3’ nucleotide</td>
<td>A &gt; T &gt; C &gt; G</td>
<td>55%</td>
</tr>
<tr>
<td>Influence of nearest neighbor nucleotide</td>
<td>G &gt; A &gt; C &gt; T</td>
<td>54%</td>
</tr>
<tr>
<td>Mixed-base polymerization</td>
<td>AATCCGGCGCGGATCCGAGAGAGAGCTATACGAGCTTGGCGGCTGAGCCGGCTCGCGCCCCGGAGTCAC (3’ end)</td>
<td>80%</td>
</tr>
</tbody>
</table>

Analytical methods such as HPLC purification. For reversible terminators, in particular hydroxylated dNTPs must be removed prior to TiEOS application to prevent homopolymeric tract formation during synthesis. To achieve this, we suggest exhausting unblocked dNTPs through a simple polymerase extension reaction. For example, chemically synthesize two complementary primers (top and bottom, 20 and 35 nt, respectively), where the top primer is 5’-biotinylated and the bottom sequence contains a homopolymeric tract at the 5’ end (e.g., A15, G15, C15, or T15); with the top strand immobilized to streptavidin-coated superparamagnetic beads (SPMBs), hybridize the bottom strand to introduce a 3’ recess. To each set add the respective dNTP sample (dATP, dCTP, and dGTP) with a standard activator (20 nt) 40 pm, 20 units of TdT (bovine, NEB), 2 mM dNTP mix (equimolar A, G, C, and T), 1X CoCl2, 1X TdT buffer, water added to a final volume of 20 μL, 37 °C/60 min. Samples were sequenced using the Sanger method.

DELETIONS

While homopolymeric tract formation can be managed to a certain degree, a much more complicated issue arises with deletions in TiEOS. To put this in perspective, we must first understand conditions affecting coupling efficiency in the chemical method (Figure 1), and how product ≤ n−1 is minimized. The coupling reaction time for standard phosphoramidite monomers is typically 20–40 s and is largely dependent on the activator used and scale of synthesis. To avoid unretracted oligonucleotides (attached to the solid substrate), reagents are typically delivered in excess to ensure maximum coverage of the substrate-bound product. The general rule is 1 of 100 oligonucleotide molecules will fail to react during the coupling step. This is restated as a 99% coupling efficiency (CE), which is defined by the equation full-length product (FLP) = (CE)n, where n is the polymer length (n−1, if the first nucleoside is preattached to the support); the purity of FLP generated during chemical synthesis is directly proportional to the total number of monomers added.62 For example, synthesis of a 20 nt strand with a CE of 99% will generate 81.7% FLP, while synthesis of 40 nt will yield 66.9% FLP with the same CE. Depurination is also another factor that contributes to synthesis failures generated via the phosphoramidite method.63 This may be caused by residual acid leftover from the deblocking step, where adenosine is particularly susceptible to cleavage at the glycosidic bond, generating an apurinic site.

Though the process by which failures are generated, contamination by product ≤ n−1 is still a problem. As such, these deletions may occur when the initiator is not fully consumed by the end of the cycle (n → n+1 conversion is <100%, and unreacted material is allowed to carry over into consecutive cycles).40–50 To emphasize a statement made earlier, the primary function of TdT is to increase antigen receptor diversity through random nucleotide incorporation. Because this adaptive immunity arose in the vertebrate system ~500 million years ago,64 it makes use of terminal transferase for nucleotide addition in a controlled, stepwise manner particularly challenging. While deletions may be due in part to an inadequate nucleotide:TdT:initiator ratio, we believe the main cause is random enzyme–nucleotide polymerization efficiency. This is where TdT appears to have a preference for adding one specific nucleobase over another, thus affecting the order and degree of nucleotide incorporation.40–50 Table 1 below provides examples of this mercurial behavior.

Physiological conditions in vivo, particularly nucleotide imbalance, may also greatly affect the TdT order of nucleotide incorporation.72,73 For example, in patients with adenosine deaminase deficiency where there is an accumulation of dATP,
N-region insertions during V(D)J recombination were 49% AT compared with 24% in normal B cells.73

In our own lab, we demonstrated TdT randomly incorporates dNTPs (G > C > A > T) onto an initiator (Table 1, mixed-base polymerization), even from a standard dNTP mix (each nucleobase represented in equimolar amounts). The initiator strand was used, TTATGCTCTGTGAAG (3′), also served as a forward primer to amplify FLP for sequencing; we added a poly A tail at the 3′ end for annealing T20 as a reverse primer. Also, the initiator nucleotide composition can affect TdT initiation efficiencies. For example, Tjong et al. found that dT10 was preferred over dC10, which resulted in an only 70% initiation efficiency.48

Short of manipulating the amino acid sequence of the terminal transerase itself, how might the information presented in Table 1 be beneficial in terms of generating ssDNA with higher yield and purity via synthesis automation? Assuming the species of TdT used in TiEOS (e.g., bovine) is consistent in its nucleobase preference, one could predict how efficient TdT might be at incorporating, for example, dATTPb on an initiator with T at the 3′ end. If this coupling event is expected to be only 95% successful (compared with dATTPb at 100%, same conditions), synthesis parameters could be optimized by either (i) delivering a larger volume of the dATTP stock solution or (ii) increasing the reaction time. Alternatively, each dNTPpb could be delivered in excess, and the reaction time could be adjusted to exceed the rate-limiting dATTPpb addition.

Despite any preference TdT may have for one particular nucleobase over another, a finely tuned ratio of enzyme to initiator is critical for addressing as many starting molecules as possible. Tang et al. found that by increasing this ratio from 0.1 (0.05 unit of TdT/μL) to 2 (1 unit/μL), the polydispersity index decreased from 1.31 to 1; here, the fraction of unextended initiator remained below 20% for each nucleotide added.48

Guanosine tracts of ≥4 nt may also be problematic in TiEOS. Because of the increased level of hydrogen bonding at the N7 ring position, G-quadruplexes are formed through intra-, bi-, or tetramolecular strand folding.44 The effects are exhibited in downstream applications such as PCR where polymerase arrest sites manifest as hairpin structures causing premature double-stranded product termination.75 It has been noted in several cases that guanosine homopolymer tracts either completely failed or showed TdT polymerization efficiencies significantly lower than that of dATP, dCTP, or dTTP.66,68 In addition to the effects of temperature and pH, the presence of metal cations can further stabilize G-quadruplexes generally in the following order: K+ > Ca2+ > Na+ > Mg2+ > Li+ and K+ > Rb+ > Cs+.70 Other reports suggest divalent cations may actually destabilize the G-quadruplex (e.g., Zn2+, Mg2+, Co2+, Mn2+, Ca2+, and Ni2+74). There are several options for minimizing secondary structure formation in G-rich sequences during TiEOS. For example, 7-deaza-dGTP disrupts hydrogen bonding between neighboring guanosines by displacing the ring nitrogen from position 7 to 8. Also, a library of reagent additives is available, which includes betaine, dimethyl sulfoxide, formamide, glycerol, NP-40, Tween 20, trehalose, and EcoSSB.77–81

Even with the phosphoramidite method, deletions may cause problems in application if not blocked from further reacting in downstream cycles. As previously mentioned, the 5′ oxygen of uncoupled molecules is acetylated during the capping step (Figure 1). These truncated species can then be removed enzymatically80 or by gel/column-based methods of filtration postsynthesis to purify the final product.83–85 In TiEOS, then, how might failures be minimized and/or removed? One
possibility is by introducing an enzyme with 3’ exonuclease activity such as Exo I, which catalyzes the complete degradation of 3’-hydroxylated ssDNA.

Figure 11. Proposed kinetic pathway of TdT distributive (template-independent) enzymatic synthesis. See Figure 10 for nomenclature. Here “n” is any strand ≥3 nt in length. Also, M1 and M2 can be either divalent cation, Mg2+ or Co2+, whereas Zn2+ may act as an allosteric cofactor. This pathway represents transferase catalytic activity for the TdTS variant of the TdT gene. The dashed arrow following step 12 indicates TdT either dissociates from the initiator and binds to another strand or remains bound to the original initiator during polymerization.

- **Figure 11**

**TdT KINETICS AND FACTORS THAT CONTRIBUTE TO RANDOM NUCLEOBASE INCORPORATION**

It should be clear by this point the effects of random nt incorporation by terminal transferase can be quite extensive (Table 1). This is most apparent when factored into oligonucleotide synthesis, where quality and product yield may be negatively impacted. A closer look at the physical properties and behaviors of TdT through its kinetic pathway could shed more light on this unpredictable behavior. Because of its kinship with terminal transferase, we compare central elements of the Pol β pathway (Figure 10) with that of TdT (Figure 11).

Unlike Pol β, TdT assumes a permanently closed ternary conformational complex; here, Arg258 in binding motif A of Pol β, which is essential for maintaining an open configuration, is absent in TdT. Instead, TdT contains loop 1, a Lariat-like conformation, which clamps down on the initiator via hydrogen bonding and precludes dsDNA from the active site. Upon entering the binding pocket, the triphosphate of the incoming dNTP is held in place by water molecules, where the sugar and nucleobase are loosely seated (steps 1–3). At this first possible checkpoint shown in Figure 10 (step 3a), Pol β determines whether dNTP pairing is a match with the template strand before metal ions commit the nucleotide to its geometry. Mismatched nucleotides that might otherwise be rejected at this stage in a template-dependent configuration are more blindly accepted in the kinetic pathway of terminal transferase. At the second fidelity checkpoint (Figure 10, step 7a), a major conformational shift occurs after both metal ions bind in the active site and properly align the nucleotide through an induced-fit mechanism; the enzyme goes from an open to closed ternary configuration.

If an improperly seated nucleotide has escaped the first fidelity checkpoint (Figure 10, step 3a), Pol β closure will sandwich the mismatch into the active site, increasing the distance from 3.4 Å (between the α-phosphate of the incoming nucleotide and the 3’ oxygen of the primer strand) to 3.9 Å. This in turn destabilizes the closed enzyme conformation of the active site. Under ideal conditions, the mismatched nucleotide is then expelled from the complex before the chemical step, and the polymerase returns to an open configuration. For TdT, there is no template for comparison; therefore, no fidelity checkpoints exist to dictate nucleotide specificity. As mentioned earlier, TdTL and TdTS act in concert in vivo; therefore, nucleotide additions and deletions are moderated. For TiEOS, we consider only the isoform with transferase activity (GenBank entry AAA36726.1) (Figure 11).

As determined by crystal structures of terminal transferase, a lack of specific contact between the incoming nucleobase and surrounding residues may explain indiscriminant dNTP allowance. This argument is strengthened by the fact TdT permits considerable flexibility toward various nucleobase modifications. For example, (i) Jarchow-Choy et al. demonstrated that their extended nucleobase dNTP analogues (xDNA) were incorporated by TdT with kinetic efficiencies comparable to those of the natural dNTP controls; similar, Berdis and McCutcheon incorporated the dNTP analogues 5-(nitro, phenyl, naphthyl, and cyclohexyl)-indolyl-2’-deoxyriboside triphosphate; (iii) Sørensen et al. used TdT to incorporate dNTP-coated proteins and other macromolecules (e.g., polyethylene glycol, dendrimer, and streptavidin); and (iv) as mentioned earlier, Winz showed use of azide-functionalized pyrimidine adducts with TdT for click chemistry applications in DNA labeling.

Unlike Pol β, which preferentially binds Mg2+, TdT can accept multiple nucleotide binding and catalytic divalent cations (Figure 11, steps 4–7); these include Mg2+, Co2+, and Zn2+ (allosteric cofactor). The specificity of nucleobase incorporation by TdT appears to be largely ion-dependent as demonstrated by Fowler and Suo: G > A > C > T and T > C > G > A, respectively [Mg2+ (purines) and Co2+ (pyrimidines)]. It has been shown, too, that micromolar quantities of Zn2+ increase the efficiency of nucleotide incorporation for both Mg2+ and Co2+ (generally at millimolar concentrations per reaction). Chang and Bollum identify Zn2+ as a nonessential allosteric cofactor for terminal transferase, which loosely interacts with the initiator and TdT binding site to induce conformational changes that increase the rate of catalysis. Whether Mg2+ or Co2+ competes for the active site when present in the same reaction, tailoring the TiEOS coupling reaction buffers separately for purines (Mg2+).

**Figure 10.** Proposed kinetic pathway of TdT template-dependent enzymatic synthesis. See Figure 11 for nomenclature. Here “n” is any strand ≥3 nt in length. Also, M1 and M2 can be either divalent cation, Mg2+ or Co2+, whereas Zn2+ may act as an allosteric cofactor. This pathway represents transferase catalytic activity for the TdTS variant of the TdT gene. The dashed arrow following step 12 indicates TdT either dissociates from the initiator and binds to another strand or remains bound to the original initiator during polymerization.
Terminal transferase due to its distributive property.91 The speciﬁc TdT:dNTPPr:initiator ratio, (ii) tailoring dNTPPr buffers98 for the specific nucleotide being incorporated (Mg2+ for purines and Co2+ for pyrimidines) with Zn2+ as an enhancer, (iii) filtering the dNTPPr stock mix from unblocked monomers, (iv) digesting unreacted initiators with an exonuclease that targets 3′-hydroxylated ssDNA, (v) introducing additives to prevent secondary structure formation for G-rich polymers, and (vi) adjusting coupling times based on TdT-speciﬁc nucleotide incorporation. If TdT nucleotide incorporation cannot be controlled by the synthesis environment alone, there are options to modify the enzyme’s amino acid structure.99,100 To touch on this, the ﬁeld of protein engineering maintains a rapidly growing toolbox of methodologies for altering protein structure and function (e.g., rational design and directed evolution).101 Helpful tools, including reconstructed evolutionary adaptive path (REAP) analysis, combine a protein’s evolutionary and functional history to best predict which amino noncanonical amino acids (Ncas) can be accepted into the binding pocket e102–104 (see Table 2).105 Attention may also be focused on altering the nucleobase binding region to more tightly regulate dNTP incorporation. For example, as shown in the Protein Data Bank (PDB entry 4I27, ternary complex of mouse TdT with ssDNA and incoming nucleotide), there is a vast open space opposite the incoming nucleobase. If loop 1, which is in the proximity, can be extended to moderate nucleobase binding, this may allow for a more uniform incorporation of all four standard dNTPs (A, G, C, and T).

Finally, we consider the prospective cost of synthesizing oligonucleotides with an enzyme compared with that of the phosphoramidite method. While there are several factors that contribute to the total cost such as the scale of synthesis, strand length, synthesis platform, and throughput (e.g., total number of samples generated either by column, titer plate, or array), we base our estimates on synthesis of 1000 1000-nucleotide samples at 1 fm each (see Table 2).

Table 2 shows TiEOS can be signiﬁcantly more cost-eﬃcient than the phosphoramidite method if TdT is recycled ($136 and $2700, compared with $136000 if a fresh stock of TdT is introduced every cycle). Recycling TdT can be done by covalently attaching the enzyme to a solid substrate;107,108 as such, TdT has been shown to yield homopolymers up to 8000 nt during living polymerization,43 which is a testament to the enzyme’s longevity in a single reaction.

Because enzymatic oligonucleotide synthesis has not yet been reduced to standard practice, many unknowns still exist. For example, how much will TdT’s random nucleobase incorporation actually contribute to deletions via the unreacted initiator? Also, can dNTP analogues with 3′ blocking moieties be accepted into the binding pocket efﬁciently enough to match TdT nucleotide incorporation rates during living polymerization? Whether these concerns may be resolved or controlled by optimizing the buffers and/or reaction conditions alone is uncertain, but the option of protein engineering TdT for maximum performance holds even greater potential. Our objective for this Perspective was to shed light on such key issues, using the phosphoramidite cycle as a metric for improving stepwise, enzymatic oligonucleotide synthesis. Therefore, a sustainable method for producing artiﬁcial ssDNA promises to transform many ﬁelds of research and development, particularly in the area of synthetic biology. This in turn will greatly beneﬁt both the environment and healthcare system.
Perspective

Biochemistry

**AUTHOR INFORMATION**

Corresponding Author

*Department of Biochemistry, Stanford Genome Technology Center, 3165 Porter Dr., Palo Alto, CA 94304. Telephone: 650 721 5477. Fax: 650 721 5651. E-mail: m.a.jensen@stanford.edu.

**ORCID**

Michael A. Jensen: 0000-0001-9230-1027

**Funding**

This work was funded by National Institutes of Health Grant 2P01HG000205-24 to R.W.D.

**Notes**

The authors declare no competing financial interest.

**ACKNOWLEDGMENTS**

The authors thank Dr. Mohsen Nemat-Gorgani for his feedback regarding the manuscript.

**REFERENCES**


Perspective


