

EFFICIENT PRODUCT CLEAVAGE IN TEMPLATE-FREE ENZYMATIC SYNTHESIS OF POLYNUCLEOTIDES

5 BACKGROUND

[0001] Interest in enzymatic approaches to polynucleotide synthesis has recently increased both because of increased demand for synthetic polynucleotides in many areas, such as synthetic biology, CRISPR-Cas9 applications, high-throughput sequencing, and the like, and because of the limitations of chemical approaches to polynucleotide synthesis, Jensen et al, 10 Biochemistry, 57: 1821-1832 (2018). Currently, most enzymatic approaches employ a template-free polymerase to repeatedly add 3'-O-blocked nucleoside triphosphates to a single stranded initiator or an elongated strand attached to a support followed by deblocking until a polynucleotide of the desired sequence is obtained. Among the challenges of devising a practical implementation of such enzymatic synthesis is to find a cost-effective and efficient 15 way to cleave a desired polynucleotide product from the initiator sequence and the support.

[0002] In view of the above, enzymatic synthesis of polynucleotides would be advanced if methods were available for high efficiency cleavage of polynucleotide products from their single stranded initiators.

20 SUMMARY OF THE INVENTION

[0003] The present invention is directed to methods and kits for template-free enzymatic synthesis of polynucleotides that include or enable a step of efficiently cleaving the polynucleotide products from its initiator using an endonuclease V activity.

[0004] In one aspect, methods of the invention include a method of synthesizing 25 polynucleotides of a predetermined sequence with the following steps: a) providing an initiator having a deoxyinosine penultimate to a 3'-terminal nucleotide having a free 3'-hydroxyl; b) repeating cycles of (i) contacting under elongation conditions the initiator or elongated fragments having free 3'-O-hydroxyls with a 3'-O-blocked nucleoside triphosphate and a template-independent DNA polymerase so that the initiator or elongated fragments are 30 elongated by incorporation of a 3'-O-blocked nucleoside triphosphate to form 3'-O-blocked elongated fragments, and (ii) deblocking the elongated fragments to form elongated fragments

having free 3'-hydroxyls, until the polynucleotide is formed; and c) treating the polynucleotide with an endonuclease V activity to cleave the polynucleotide from the initiator.

[0005] The present invention advantageously overcomes the above problems in the field of enzymatic polynucleotide synthesis by providing an initiator having a deoxyinosine at the penultimate position from its 3' end. This permits efficient cleavage of the single stranded initiator at its terminal nucleotide releasing a polynucleotide product with a 5'-monophosphate.

BRIEF DESCRIPTION OF THE DRAWINGS

[0006] Fig1. illustrates an experimental set up for demonstrating the cleavage efficiency of the present invention.

[0007] Fig. 2 shows data comparing efficiencies of USER/deoxyuridine cleavage and Endo V/inosine cleavage.

DETAILED DESCRIPTION OF THE INVENTION

[0008] The general principles of the invention are disclosed in more detail herein particularly by way of examples, such as those shown in the drawings and described in detail. It should be understood, however, that the intention is not to limit the invention to the particular embodiments described. The invention is amenable to various modifications and alternative forms, specifics of which are shown for several embodiments. The intention is to cover all modifications, equivalents, and alternatives falling within the principles and scope of the invention.

[0009] The practice of the present invention may employ, unless otherwise indicated, conventional techniques and descriptions of organic chemistry, molecular biology (including recombinant techniques), cell biology, and biochemistry, which are within the skill of the art. Such conventional techniques may include, but are not limited to, preparation and use of synthetic peptides, synthetic polynucleotides, monoclonal antibodies, nucleic acid cloning, amplification, sequencing and analysis, and related techniques. Protocols for such conventional techniques can be found in product literature from manufacturers and in standard laboratory manuals, such as Genome Analysis: A Laboratory Manual Series (Vols. I-IV); PCR Primer: A Laboratory Manual; and Molecular Cloning: A Laboratory Manual (all from Cold

Spring Harbor Laboratory Press); Lutz and Bornscheuer, Editors, Protein Engineering Handbook (Wiley-VCH, 2009); Hermanson, Bioconjugate Techniques, Second Edition (Academic Press, 2008); and like references.

[0010] The present invention is based in part on a recognition and appreciation of the efficiency of using an endonuclease V activity and a deoxyinosine penultimate to the 3' terminal nucleotide of an initiator to cleave a polynucleotide product from an initiator, as compared to other cleavable nucleotides, such as deoxyuridine. In one aspect, it is believed that synthesis initiation by a terminal deoxynucleotidyl transferase (TdT) on an initiator with a penultimate deoxyinosine is much more efficient than initiation on other cleavable nucleotide configurations.

[0011] Fig. 1 provides a diagram of a template-free enzymatic synthesis method employing initiators with a penultimate deoxyinosine. Shown in this depiction are initiators (100) attached by their 5' ends to solid support (105). Each initiator (102) has a 3'-penultimate deoxyinosine (104) next to 3'-terminal nucleotide (106) that has a free 3' hydroxyl. After a predetermined number of cycles of enzymatic incorporation and de-blocking, a polynucleotide product is formed that is attached to solid support (105) by initiators (102). The polynucleotide product is cleaved from initiators (102) and support (105) by treating the attached product with an endonuclease V activity which recognizes the presence of the deoxyinosine and cleaves the strand on the 3' side (112) of terminal nucleotide (106) of the initiators. In some embodiments, the endonuclease V activity is provided by using a prokaryotic endonuclease V. In still other embodiments, the endonuclease V is an E. coli endonuclease V. As used herein, the term "endonuclease V activity" means an enzyme activity that catalyzes the following cleavage reaction in a single stranded DNA: 5' ...NNINNNN ...-3' $\rightarrow$ 5'-... NNIN + 5'-PO₄-NNNN ... -3' where N is any nucleotide and I is deoxyinosine. Cleavage (114) of polynucleotides (116) by an endonuclease V activity leaves a 5'-monophosphate on the polynucleotides, which optionally may be removed by a step of treating them with a 5'-phosphatase.

[0012] Enzymes with endonuclease V activity are available from commercial enzyme suppliers, for example, New England Biolabs (Beverly, MA, USA), NzyTech (Lisbon, Portugal). Such enzymes may be used with the supplier's recommended cleavage buffers (e.g. 50mM K-Ac, 20mM Tris-Ac, 10mM Mg-Ac, 1mM DTT at pH 7.9). Typical cleavage

conditions are as follows: 70U of Endo V in 50µl of Nzytech buffer at 37°C for 500 pmol synthesis scale on resin. Typical cleavage times are from 5 to 60 minutes, or from 10 to 30 minutes. Optionally, endonuclease activity of the above enzymes may be heat inactivated by incubation at 65°C or higher for 20 minutes. Optionally, the Nzytech endonuclease V
5 comprises a His tag that allows convenient removal of the enzyme from reaction mixtures in preparation of final products.

Template-Free Enzymatic Synthesis

[0013] Template-free enzymatic synthesis of polynucleotides may be carried out by a
10 variety of known protocols using template-free polymerases, such as terminal deoxynucleotidyl transferase (TdT), including variants thereof engineered to accommodate more efficiently 3'-O-blocked deoxynucleoside triphosphates (3'-O-blocked dNTPs), e.g. Ybert et al, International patent publication WO/2015/159023; Ybert et al, International patent publication WO/2017/216472; Hyman, U.S. patent 5436143; Hiatt et al, U.S. patent 5763594;
15 Jensen et al, Biochemistry, 57: 1821-1832 (2018); Mathews et al, Organic & Biomolecular Chemistry, DOI: 0.1039/c6ob01371f (2016); Schmitz et al, Organic Lett., 1(11): 1729-1731 (1999).

[0014] In some embodiments, an ordered sequence of nucleotides is coupled to an initiator nucleic acid using a TdT in the presence of 3'-O-reversibly blocked dNTPs in each synthesis
20 step. In some embodiments, the method of synthesizing an oligonucleotide comprises the steps of (a) providing an initiator having a free 3'-hydroxyl; (b) reacting under extension conditions the initiator or an extension intermediate having a free 3'-hydroxyl with a TdT in the presence of a 3'-O-blocked nucleoside triphosphate to produce a 3'-O-blocked extension intermediate; (c) deblocking the extension intermediate to produce an extension intermediate
25 with a free 3'-hydroxyl; and (d) repeating steps (b) and (c) until the polynucleotide is synthesized. (Sometime "an extension intermediate" is also referred to as an "elongation fragment.") In some embodiments, an initiator is provided as an oligonucleotide attached to a solid support, e.g. by its 5' end. The above method may also include washing steps after the reaction, or extension, step, as well as after the de-blocking step. For example, the step of
30 reacting may include a sub-step of removing unincorporated nucleoside triphosphates, e.g. by washing, after a predetermined incubation period, or reaction time. Such predetermined

incubation periods or reaction times may be a few seconds, e.g. 30 sec, to several minutes, e.g. 30 min.

[0015] The above method may also include capping step(s) as well as washing steps after the reacting, or extending, step, as well as after the deblocking step. As mentioned above, in some embodiments, capping steps may be included in which non-extended free 3'-hydroxyls are reacted with compounds that prevents any further extensions of the capped strand. In some embodiments, such compound may be a dideoxynucleoside triphosphate. In other embodiments, non-extended strands with free 3'-hydroxyls may be degraded by treating them with a 3'-exonuclease activity, e.g. Exo I. For example, see Hyman, U.S. patent 5436143. Likewise, in some embodiments, strands that fail to be deblocked may be treated to either remove the strand or render it inert to further extensions.

[0016] In some embodiments that comprise serial synthesis of oligonucleotides, capping steps may be undesirable as capping may prevent the production of equal molar amounts of a plurality of oligonucleotides. Without capping, sequences will have a uniform distribution of deletion errors, but each of a plurality of oligonucleotides will be present in equal molar amounts. This would not be the case where non-extended fragments are capped.

[0017] In some embodiments, reaction conditions for an extension or elongation step may comprising the following: 2.0 μ M purified TdT; 125-600 μ M 3'-O-blocked dNTP (e.g. 3'-O-NH₂-blocked dNTP); about 10 to about 500 mM potassium cacodylate buffer (pH between 6.5 and 7.5) and from about 0.01 to about 10 mM of a divalent cation (e.g. CoCl₂ or MnCl₂), where the elongation reaction may be carried out in a 50 μ L reaction volume, at a temperature within the range RT to 45°C, for 3 minutes. In embodiments, in which the 3'-O-blocked dNTPs are 3'-O-NH₂-blocked dNTPs, reaction conditions for a deblocking step may comprise the following: 700 mM NaNO₂; 1 M sodium acetate (adjusted with acetic acid to pH in the range of 4.8-6.5), where the deblocking reaction may be carried out in a 50 μ L volume, at a temperature within the range of RT to 45°C for 30 seconds to several minutes.

[0018] Depending on particular applications, the steps of deblocking and/or cleaving may include a variety of chemical or physical conditions, e.g. light, heat, pH, presence of specific reagents, such as enzymes, which are able to cleave a specified chemical bond. Guidance in selecting 3'-O-blocking groups and corresponding de-blocking conditions may be found in the following references, which are incorporated by reference: U.S. patent 5808045; U.S. patent

8808988; International patent publication WO91/06678; and references cited below. In some embodiments, the cleaving agent (also sometimes referred to as a de-blocking reagent or agent) is a chemical cleaving agent, such as, for example, dithiothreitol (DTT). In alternative embodiments, a cleaving agent may be an enzymatic cleaving agent, such as, for example, a phosphatase, which may cleave a 3'-phosphate blocking group. It will be understood by the person skilled in the art that the selection of deblocking agent depends on the type of 3'-nucleotide blocking group used, whether one or multiple blocking groups are being used, whether initiators are attached to living cells or organisms or to solid supports, and the like, that necessitate mild treatment. For example, a phosphine, such as tris(2-carboxyethyl)phosphine (TCEP) can be used to cleave a 3'-O-azidomethyl groups, palladium complexes can be used to cleave a 3'-O-allyl groups, or sodium nitrite can be used to cleave a 3'-O-amino group. In particular embodiments, the cleaving reaction involves TCEP, a palladium complex or sodium nitrite.

[0019] As noted above, in some embodiments it is desirable to employ two or more blocking groups that may be removed using orthogonal de-blocking conditions. The following exemplary pairs of blocking groups may be used in parallel synthesis embodiments, such as those described above. It is understood that other blocking group pairs, or groups containing more than two, may be available for use in these embodiments of the invention.

3'-O-NH ₂	3'-O-azidomethyl
3'-O-NH ₂	3'-O-allyl
3'-O-NH ₂	3'-O-phosphate
3'-O-azidomethyl	3'-O-allyl
3'-O-azidomethyl	3'-O-phosphate
3'-O-allyl	3'-O-phosphate

[0020] Synthesizing oligonucleotides on living cells requires mild deblocking, or deprotection, conditions, that is, conditions that do not disrupt cellular membranes, denature proteins, interfere with key cellular functions, or the like. In some embodiments, deprotection conditions are within a range of physiological conditions compatible with cell survival. In such embodiments, enzymatic deprotection is desirable because it may be carried out under

physiological conditions. In some embodiments specific enzymatically removable blocking groups are associated with specific enzymes for their removal. For example, ester- or acyl-based blocking groups may be removed with an esterase, such as acetylcysteine, or like enzyme, and a phosphate blocking group may be removed with a 3' phosphatase, such as T4 polynucleotide kinase. By way of example, 3'-O-phosphates may be removed by treatment with a solution of 100 mM Tris-HCl (pH 6.5) 10 mM MgCl₂, 5 mM 2-mercaptoethanol, and one Unit T4 polynucleotide kinase. The reaction proceeds for one minute at a temperature of 37°C.

[0021] A "3'-phosphate-blocked" or "3'-phosphate-protected" nucleotide refers to nucleotides in which the hydroxyl group at the 3'-position is blocked by the presence of a phosphate containing moiety. Examples of 3'-phosphate-blocked nucleotides in accordance with the invention are nucleotidyl-3'-phosphate monoester/nucleotidyl-2',3'-cyclic phosphate, nucleotidyl-2'-phosphate monoester and nucleotidyl-2' or 3'-alkylphosphate diester, and nucleotidyl-2' or 3'-pyrophosphate. Thiophosphate or other analogs of such compounds can also be used, provided that the substitution does not prevent dephosphorylation resulting in a free 3'-OH by a phosphatase.

[0022] Further examples of synthesis and enzymatic deprotection of 3'-O-ester-protected dNTPs or 3'-O-phosphate-protected dNTPs are described in the following references: Canard et al, Proc. Natl. Acad. Sci., 92:10859-10863 (1995); Canard et al, Gene, 148: 1-6 (1994); Cameron et al, Biochemistry, 16(23): 5120-5126 (1977); Rasolonjatovo et al, Nucleosides & Nucleotides, 18(4&5): 1021-1022 (1999); Ferrero et al, Monatshefte für Chemie, 131: 585-616 (2000); Taunton-Rigby et al, J. Org. Chem., 38(5): 977-985 (1973); Uemura et al, Tetrahedron Lett., 30(29): 3819-3820 (1989); Becker et al, J. Biol. Chem., 242(5): 936-950 (1967); Tsien, International patent publication WO1991/006678.

[0023] As used herein, an "initiator" (or equivalent terms, such as, "initiating fragment," "initiator nucleic acid," "initiator oligonucleotide," or the like) refers to a short oligonucleotide sequence with a free 3'-end, which can be further elongated by a template-free polymerase, such as TdT. In one embodiment, the initiating fragment is a DNA initiating fragment. In an alternative embodiment, the initiating fragment is an RNA initiating fragment. In one embodiment, the initiating fragment possesses between 3 and 100 nucleotides, in particular between 3 and 20 nucleotides. In one embodiment, the initiating fragment is single-stranded.

In an alternative embodiment, the initiating fragment is double-stranded. In a particular embodiment, an initiator oligonucleotide synthesized with a 5'-primary amine may be covalently linked to magnetic beads using the manufacturer's protocol. Likewise, an initiator oligonucleotide synthesized with a 3'-primary amine may be covalently linked to magnetic beads using the manufacturer's protocol. A variety of other attachment chemistries amenable for use with embodiments of the invention are well-known in the art, e.g. Integrated DNA Technologies brochure, "Strategies for Attaching Oligonucleotides to Solid Supports," v.6 (2014); Hermanson, Bioconjugate Techniques, Second Edition (Academic Press, 2008); and like references.

[0024] Many of the 3'-O-blocked dNTPs employed in the invention may be purchased from commercial vendors or synthesized using published techniques, e.g. U.S. patent 7057026; International patent publications WO2004/005667, WO91/06678; Canard et al, Gene (cited above); Metzker et al, Nucleic Acids Research, 22: 4259-4267 (1994); Meng et al, J. Org. Chem., 14: 3248-3252 (3006); U.S. patent publication 2005/037991. In some embodiments, the modified nucleotides comprise a modified nucleotide or nucleoside molecule comprising a purine or pyrimidine base and a ribose or deoxyribose sugar moiety having a removable 3'-OH blocking group covalently attached thereto, such that the 3' carbon atom has attached a group of the structure:



wherein $-Z$ is any of $-C(R')_2-O-R''$, $-C(R')_2-N(R'')_2$, $-C(R')_2-N(H)R''$, $-C(R')_2-S-R''$ and $-C(R')_2-F$, wherein each R'' is or is part of a removable protecting group; each R' is independently a hydrogen atom, an alkyl, substituted alkyl, arylalkyl, alkenyl, alkynyl, aryl, heteroaryl, heterocyclic, acyl, cyano, alkoxy, aryloxy, heteroaryloxy or amido group, or a detectable label attached through a linking group; with the proviso that in some embodiments such substituents have up to 10 carbon atoms and/or up to 5 oxygen or nitrogen heteroatoms; or $(R')_2$ represents a group of formula $=C(R''')_2$ wherein each R''' may be the same or different and is selected from the group comprising hydrogen and halogen atoms and alkyl groups, with the proviso that in some embodiments the alkyl of each R''' has from 1 to 3 carbon atoms; and wherein the molecule may be reacted to yield an intermediate in which each R'' is exchanged for H or, where Z is $-(R')_2-F$, the F is exchanged for OH, SH or NH_2 , preferably OH, which intermediate dissociates under aqueous conditions to afford a molecule with a free 3'-OH; with

the proviso that where Z is $-C(R')_2-S-R''$, both R' groups are not H. In certain embodiments, R' of the modified nucleotide or nucleoside is an alkyl or substituted alkyl, with the proviso that such alkyl or substituted alkyl has from 1 to 10 carbon atoms and from 0 to 4 oxygen or nitrogen heteroatoms. In certain embodiments, -Z of the modified nucleotide or nucleoside is
5 of formula $-C(R')_2-N_3$. In certain embodiments, Z is an azidomethyl group.

[0025] In some embodiments, Z is a cleavable organic moiety with or without heteroatoms having a molecular weight of 200 or less. In other embodiments, Z is a cleavable organic moiety with or without heteroatoms having a molecular weight of 100 or less. In other
10 embodiments, Z is a cleavable organic moiety with or without heteroatoms having a molecular weight of 50 or less. In some embodiments, Z is an enzymatically cleavable organic moiety with or without heteroatoms having a molecular weight of 200 or less. In other embodiments, Z is an enzymatically cleavable organic moiety with or without heteroatoms having a molecular weight of 100 or less. In other
15 embodiments, Z is an enzymatically cleavable organic moiety with or without heteroatoms having a molecular weight of 50 or less. In other embodiments, Z is an enzymatically cleavable ester group having a molecular weight of 200 or less. In other embodiments, Z is a phosphate group removable by a 3'-phosphatase. In some
20 embodiments, one or more of the following 3'-phosphatases may be used with the manufacturer's recommended protocols: T4 polynucleotide kinase, calf intestinal alkaline phosphatase, recombinant shrimp alkaline phosphatase (e.g. available from New England Biolabs, Beverly, MA).

[0026] In a further particular embodiment, the 3'-blocked nucleotide triphosphate is blocked by either a 3'-O-azidomethyl, 3'-O-NH₂ or 3'-O-allyl group.

[0027] In still other embodiments, 3'-O-blocking groups of the invention include 3'-O-methyl, 3'-O-(2-nitrobenzyl), 3'-O-allyl, 3'-O-amine, 3'-O-azidomethyl, 3'-O-tert-butoxy
25 ethoxy, 3'-O-(2-cyanoethyl), and 3'-O-propargyl.

[0028] In some embodiments, 3'-O- protection groups are electrochemically labile groups. That is, deprotection or cleavage of the protection group is accomplished by changing the electrochemical conditions in the vicinity of the protection group which result in cleavage. Such changes in electrochemical conditions may be brought about by changing or applying a
30 physical quantity, such as a voltage difference or light to activate auxiliary species which, in turn, cause changes in the electrochemical conditions at the site of the protection group, such

as an increase or decrease in pH. In some embodiments, electrochemically labile groups include, for example, pH-sensitive protection groups that are cleaved whenever the pH is changed to a predetermined value. In other embodiments, electrochemically labile groups include protecting groups which are cleaved directly whenever reducing or oxidizing conditions are changed, for example, by increasing or decreasing a voltage difference at the site of the protection group.

[0029] In some embodiments, enzymatic synthesis methods employ TdT variants that display increased incorporation activity with respect to 3'-O-modified nucleoside triphosphates. For example, such TdT variants may be produced using techniques described in Champion et al, U.S. patent 10435676, which is incorporated herein by reference. In some embodiments, a TdT variant is employed having an amino acid sequence at least 60 percent identical to SEQ ID NO: 2 and a substitution at a first arginine at position 207 and a substitution at a second arginine at position 325, or functionally equivalent residues thereof. In some embodiments, a terminal deoxynucleotidyl transferase (TdT) variant is employed that has an amino acid sequence at least sixty percent identical to an amino acid sequence selected from SEQ ID NO: 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 or 15 with a substitution of arginine ("first arginine") at position 207 with respect to SEQ ID NOs 2, 3, 4, 6, 7, 9, 12 and 13, at position 206 with respect to SEQ ID NO 5, at position 208 with respect to SEQ ID NOs 8 and 10, at position 205 with respect to SEQ ID NO 11, at position 216 with respect to SEQ ID NO 14 and at position 210 with respect to SEQ ID NO 15; and a substitution of arginine ("second arginine") at position 325 with respect to SEQ ID NOs 2, 9 and 13, at position 324 with respect to SEQ ID NOs 3 and 4, at position 320 with respect to SEQ ID NO 320, at position 331 with respect to SEQ ID NOs 6 and 8, at position 323 with respect to SEQ ID NO 11, at position 328 with respect to SEQ ID NOs 12 and 15, and at position 338 with respect to SEQ ID NO 14; or functionally equivalent residues thereof; wherein the TdT variant (i) is capable of synthesizing a nucleic acid fragment without a template and (ii) is capable of incorporating a 3'-O-modified nucleotide onto a free 3'-hydroxyl of a nucleic acid fragment. In some embodiments, the above percent identity value is at least 80 percent identity with the indicated SEQ ID NOs; in some embodiments, the above percent identity value is at least 90 percent identity with the indicated SEQ ID NOs; in some embodiments, the above percent identity value is at least 95 percent identity with the indicated SEQ ID NOs; in some

embodiments, the above percent identity value is at least 97 percent identity; in some embodiments, the above percent identity value is at least 98 percent identity; in some embodiments, the above percent identity value is at least 99 percent identity. As used herein, the percent identity values used to compare a reference sequence to a variant sequence do not include the expressly specified amino acid positions containing substitutions of the variant sequence; that is, the percent identity relationship is between sequences of a reference protein and sequences of a variant protein outside of the expressly specified positions containing substitutions in the variant. Thus, for example, if the reference sequence and the variant sequence each comprised 100 amino acids and the variant sequence had mutations at positions 25 and 81, then the percent homology would be in regard to sequences 1-24, 26-80 and 82-100.

[0030] In regard to (ii), such 3'-O-modified nucleotide may comprise a 3'-O-NH₂-nucleoside triphosphate, a 3'-O-azidomethyl-nucleoside triphosphate, a 3'-O-allyl-nucleoside triphosphate, a 3'-O—(2-nitrobenzyl)-nucleoside triphosphate, or a 3'-O-propargyl-nucleoside triphosphate.

[0031] In some embodiments, TdT variants used in the invention have substitutions as shown in Table 1 or functionally equivalent residue positions in other TdTs.

Table 1

SEQ ID NO	Substitutions				
1	M192R/Q	C302G/R	R336L/N	R454P/N/A/V	E457N/L/T/S/K
2	M63R/Q	C173G/R	R207L/N	R325P/N/A/V	E328N/L/T/S/K
3	M63R/Q	C173G/R	R207L/N	R324P/N/A/V	E327N/L/T/S/K
4	M63R/Q	C173G/R	R207L/N	R324P/N/A/V	E327N/L/T/S/K
5	---	C172G/R	R206L/N	R320P/N/A/V	---
6	M63R/Q	C173G/R	R207L/N	R331P/N/A/V	E334N/L/T/S/K
7	M63R/Q	C173G/R	R207L/N	---	E328N/L/T/S/K
8	---	C174G/R	R208L/N	R331P/N/A/V	E334N/L/T/S/K
9	M73R/Q	C173G/R	R207L/N	R325P/N/A/V	E328N/L/T/S/K
10	M64R/Q	C174G/R	R208L/N	---	E329N/L/T/S/K
11	M61R/Q	C171G/R	R205L/N	R323P/N/A/V	E326N/L/T/S/K
12	M63R/Q	C173G/R	R207L/N	R328P/N/A/V	E331N/L/T/S/K
13	---	C173G/R	R207L/N	R325P/N/A/V	E328N/L/T/S/K
14	M63R/Q	C182G/R	R216L/N	R338P/N/A/V	E341N/L/T/S/K
15	M66R/Q	C176G/R	R210L/N	R328P/N/A/V	E331N/L/T/S/K

[0032] TdT variants of the invention as described above each comprise an amino acid sequence having a percent sequence identity with a specified SEQ ID NO, subject to the presence of indicated substitutions. In some embodiments, the number and type of sequence differences between a TdT variant of the invention described in this manner and the specified SEQ ID NO may be due to substitutions, deletion and/or insertions, and the amino acids substituted, deleted and/or inserted may comprise any amino acid. In some embodiments, such deletions, substitutions and/or insertions comprise only naturally occurring amino acids. In some embodiments, substitutions comprise only conservative, or synonymous, amino acid changes, as described in Grantham, Science, 185: 862-864 (1974). That is, a substitution of an amino acid can occur only among members of its set of synonymous amino acids. In some embodiments, sets of synonymous amino acids that may be employed are set forth in Table 2A.

Table 2A: Synonymous Sets of Amino Acids I

Amino Acid	Synonymous Set
Ser	Ser, Thr, Gly, Asn
Arg	Arg, Gln, Lys, Glu, His
Leu	Ile, Phe, Tyr, Met, Val, Leu
Pro	Gly, Ala, Thr, Pro
Thr	Pro, Ser, Ala, Gly, His, Gln, Thr
Ala	Gly, Thr, Pro, Ala
Val	Met, Tyr, Phe, Ile, Leu, Val
Gly	Gly, Ala, Thr, Pro, Ser
Ile	Met, Tyr, Phe, Val, Leu, Ile
Phe	Trp, Met, Tyr, Ile, Val, Leu, Phe
Tyr	Trp, Met, Phe, Ile, Val, Leu, Tyr
Cys	Cys, Ser, Thr
His	His, Glu, Lys, Gln, Thr, Arg
Gln	Gln, Glu, Lys, Asn, His, Thr, Arg
Asn	Asn, Gln, Asp, Ser
Lys	Lys, Glu, Gln, His, Arg
Asp	Asp, Glu, Asn
Glu	Glu, Asp, Lys, Asn, Gln, His, Arg
Met	Met, Phe, Ile, Val, Leu
Trp	Trp

[0033] In some embodiments, sets of synonymous amino acids that may be employed are set forth in Table 2B.

Table 2B: Synonymous Sets of Amino Acids II

Amino Acid	Synonymous Set
Ser	Ser
Arg	Arg, Lys, His
Leu	Ile, Phe, Met, Leu
Pro	Ala, Pro
Thr	Thr
Ala	Pro, Ala
Val	Met, Ile Val
Gly	Gly
Ile	Met, Phe, Val, Leu, Ile
Phe	Met, Tyr, Ile, Leu, Phe
Tyr	Trp, Met
Cys	Cys, Ser
His	His, Gln, Arg
Gln	Gln, Glu, His
Asn	Asn, Asp
Lys	Lys, Arg
Asp	Asp, Asn
Glu	Glu, Gln
Met	Met, Phe, Ile, Val, Leu
Trp	Trp

Kits

[0034] The invention includes kits for carrying out methods of the invention. In some
5 embodiments, a kit of the invention comprises an initiator attached to a support by a 5' end
and having a deoxyinosine penultimate to a 3' end and free 3'-hydroxyl. In some
embodiments, a kit of the invention further includes an endonuclease V capable of cleaving an
initiator-polynucleotide conjugate 3' of a terminal nucleotide of the initiator. In some such
kits, the endonuclease V has a capture moiety to permit removal from a reaction mixture. In
10 some kits, such capture moiety is a His tag. In some embodiments, initiators of a kit have a 3'-
terminal sequence of 5'-dI-dT-3'. In some embodiments, initiators of a kit have a 3'-terminal
sequence of 5'-dI-dG-3'. In some embodiments, initiators of a kit have a 3'-terminal sequence
of 5'-dI-dA-3'. In some embodiments, initiators of a kit have a 3'-terminal sequence of 5'-dI-
dT-3', 5'-dI-dG-3', or 5'-dI-dA-3'. In some embodiments, such support is a solid support.
15 Such solid support may comprise beads, such as magnetic bead, a planar solid, such as a glass
slide, or a membrane, or the like. In some embodiments, a kit of the invention may further
include a template-free polymerase and 3'-O-blocked nucleoside triphosphates of one or more

of deoxyadenosine, deoxyguanosine, thymidine, deoxyuridine and deoxycytidine. In some kits, such template-free polymerase may be a TdT. In some embodiments, such TdT may be a TdT variant described herein. In some embodiments, a kit of the invention may further include a de-blocking agent which is capable of removing 3' blocking groups from
5 incorporated 3'-O-blocked nucleotides.

EXAMPLE

[0035] In this example, the efficiency of using deoxyinosine/endo V cleavage is compared to deoxyuridine/USER cleavage and the effects on cleavage of nucleotides adjacent to dI are
10 assessed. 5'-amino-poly(dT) oligonucleotides containing dI were coupled to carboxyl groups of magnetic beads using EDC in a conventional reaction. In all experiments, initiators comprised either (1) a 5'-10mer polyT segment followed by a deoxyinosine and 3' terminal dT, or (2) a 5'-10mer polyT segment followed by a terminal deoxyuridine. In some experiments, initiators were extended by a 20mer polyT segment followed by a final dA
15 labeled with a Cy5 dye, all using a TdT enzyme and 3'-O-NH₂-blocked nucleoside triphosphates (except for the labeled terminal dA). In other experiments, the initiators were extended by the indicated dinucleotide sequences followed by a 18mer poly(dT) and a final dA labeled with a Cy5 dye, all using a TdT enzyme and 3'-O-NH₂-blocked nucleoside triphosphates (except for the labeled terminal dA). After cleavage as indicated (USER or
20 Endo V), the cleaved labeled polynucleotides were analyzed by polyacrylamide gel electrophoresis.

[0036] Fig. 2 shows electrophoresis data comparing synthesis products of initiators having terminal deoxyuridines with synthesis products of initiators having penultimate deoxyinosines. The bands in the four ladders on the left of the gel corresponding to deoxyuridine initiators
25 show failure sequences that are significantly more intense than the corresponding bands from deoxyinosine initiators in the rightmost 10 ladders indicating that initiators with penultimate deoxyinosines result in more efficient synthesis than initiators with terminal deoxyuridines.

Definitions

[0037] Unless otherwise specifically defined herein, terms and symbols of nucleic acid chemistry, biochemistry, genetics, and molecular biology used herein follow those of standard treatises and texts in the field, e.g. Kornberg and Baker, DNA Replication, Second Edition (W.H. Freeman, New York, 1992); Lehninger, Biochemistry, Second Edition (Worth Publishers, New York, 1975); Strachan and Read, Human Molecular Genetics, Second Edition (Wiley-Liss, New York, 1999).

[0038] “Functionally equivalent” in reference to amino acid positions in two or more different TdTs means (i) the amino acids at the respective positions play the same functional role in an activity of the TdTs, and (ii) the amino acids occur at homologous amino acid positions in the amino acid sequences of the respective TdTs. It is possible to identify positionally equivalent or homologous amino acid residues in the amino acid sequences of two or more different TdTs on the basis of sequence alignment and/or molecular modelling. In some embodiments, functionally equivalent amino acid positions belong to sequence motifs that are conserved among the amino acid sequences of TdTs of evolutionarily related species, e.g. genus, families, or the like. Examples of such conserved sequence motifs are described in Motea et al, Biochim. Biophys. Acta. 1804(5): 1151-1166 (2010); Delarue et al, EMBO J., 21: 427-439 (2002); and like references.

[0039] “Kit” refers to any delivery system for delivering materials or reagents for carrying out a method of the invention. In the context of reaction assays, such delivery systems include systems and/or compounds (such as dilutants, surfactants, carriers, or the like) that allow for the storage, transport, or delivery of reaction reagents (e.g., fluorescent labels, such as mutually quenching fluorescent labels, fluorescent label linking agents, enzymes, quenching agents, etc. in the appropriate containers) and/or supporting materials (e.g., buffers, written instructions for performing the assay etc.) from one location to another. For example, kits include one or more enclosures (e.g., boxes) containing the relevant reaction reagents and/or supporting materials. Such contents may be delivered to the intended recipient together or separately. For example, a first container may contain an enzyme for use in an assay, while a second or more containers contain mutually quenching fluorescent labels and/or quenching agents.

[0040] “Mutant” or “variant,” which are used interchangeably, refer to polypeptides derived from a natural or reference TdT polypeptide described herein, and comprising a modification or an alteration, i.e., a substitution, insertion, and/or deletion, at one or more positions. Variants may be obtained by various techniques well known in the art. In particular, examples of techniques for altering the DNA sequence encoding the wild-type protein, include, but are not limited to, site-directed mutagenesis, random mutagenesis, sequence shuffling and synthetic oligonucleotide construction. Mutagenesis activities consist in deleting, inserting or substituting one or several amino-acids in the sequence of a protein or in the case of the invention of a polymerase.

[0041] “Polynucleotide” or “oligonucleotide” are used interchangeably and each mean a linear polymer of nucleotide monomers or analogs thereof. Monomers making up polynucleotides and oligonucleotides are capable of specifically binding to a natural polynucleotide by way of a regular pattern of monomer-to-monomer interactions, such as Watson-Crick type of base pairing, base stacking, Hoogsteen or reverse Hoogsteen types of base pairing, or the like. Such monomers and their internucleosidic linkages may be naturally occurring or may be analogs thereof, e.g. naturally occurring or non-naturally occurring analogs. Non-naturally occurring analogs may include PNAs, phosphorothioate internucleosidic linkages, bases containing linking groups permitting the attachment of labels, such as fluorophores, or haptens, and the like. Whenever the use of an oligonucleotide or polynucleotide requires enzymatic processing, such as extension by a polymerase, ligation by a ligase, or the like, one of ordinary skill would understand that oligonucleotides or polynucleotides in those instances would not contain certain analogs of internucleosidic linkages, sugar moieties, or bases at any or some positions. Polynucleotides typically range in size from a few monomeric units, e.g. 5-40, when they are usually referred to as “oligonucleotides,” to several thousand monomeric units. Whenever a polynucleotide or oligonucleotide is represented by a sequence of letters (upper or lower case), such as "ATGCCTG," it will be understood that the nucleotides are in 5'→3' order from left to right and that "A" denotes deoxyadenosine, "C" denotes deoxycytidine, "G" denotes deoxyguanosine, and "T" denotes thymidine, “I” denotes deoxyinosine, “U” denotes uridine, unless otherwise indicated or obvious from context. Unless otherwise noted the terminology and atom numbering conventions will follow those disclosed in Strachan and Read, Human

Molecular Genetics 2 (Wiley-Liss, New York, 1999). Usually polynucleotides comprise the four natural nucleosides (e.g. deoxyadenosine, deoxycytidine, deoxyguanosine, deoxythymidine for DNA or their ribose counterparts for RNA) linked by phosphodiester linkages; however, they may also comprise non-natural nucleotide analogs, e.g. including modified bases, sugars, or internucleosidic linkages. It is clear to those skilled in the art that where an enzyme has specific oligonucleotide or polynucleotide substrate requirements for activity, e.g. single stranded DNA, RNA/DNA duplex, or the like, then selection of appropriate composition for the oligonucleotide or polynucleotide substrates is well within the knowledge of one of ordinary skill, especially with guidance from treatises, such as Sambrook et al, Molecular Cloning, Second Edition (Cold Spring Harbor Laboratory, New York, 1989), and like references. Likewise, the oligonucleotide and polynucleotide may refer to either a single stranded form or a double stranded form (i.e. duplexes of an oligonucleotide or polynucleotide and its respective complement). It will be clear to one of ordinary skill which form or whether both forms are intended from the context of the terms usage.

[0042] “Primer” means an oligonucleotide, either natural or synthetic that is capable, upon forming a duplex with a polynucleotide template, of acting as a point of initiation of nucleic acid synthesis and being extended from its 3’ end along the template so that an extended duplex is formed. Extension of a primer is usually carried out with a nucleic acid polymerase, such as a DNA or RNA polymerase. The sequence of nucleotides added in the extension process is determined by the sequence of the template polynucleotide. Usually primers are extended by a DNA polymerase. Primers usually have a length in the range of from 14 to 40 nucleotides, or in the range of from 18 to 36 nucleotides. Primers are employed in a variety of nucleic amplification reactions, for example, linear amplification reactions using a single primer, or polymerase chain reactions, employing two or more primers. Guidance for selecting the lengths and sequences of primers for particular applications is well known to those of ordinary skill in the art, as evidenced by the following references that are incorporated by reference: Dieffenbach, editor, PCR Primer: A Laboratory Manual, 2nd Edition (Cold Spring Harbor Press, New York, 2003).

[0043] “Sequence identity” refers to the number (or fraction, usually expressed as a percentage) of matches (e.g., identical amino acid residues) between two sequences, such as two polypeptide sequences or two polynucleotide sequences. The sequence identity is

determined by comparing the sequences when aligned so as to maximize overlap and identity while minimizing sequence gaps. In particular, sequence identity may be determined using any of a number of mathematical global or local alignment algorithms, depending on the length of the two sequences. Sequences of similar lengths are preferably aligned using a global alignment algorithm (e.g. Needleman and Wunsch algorithm; Needleman and Wunsch, 1970) which aligns the sequences optimally over the entire length, while sequences of substantially different lengths are preferably aligned using a local alignment algorithm (e.g. Smith and Waterman algorithm (Smith and Waterman, 1981) or Altschul algorithm (Altschul et al., 1997; Altschul et al., 2005)). Alignment for purposes of determining percent amino acid sequence identity can be achieved in various ways that are within the skill in the art, for instance, using publicly available computer software available on internet web sites such as <http://blast.ncbi.nlm.nih.gov/> or <http://www.ebi.ac.uk/Tools/emboss/>. Those skilled in the art can determine appropriate parameters for measuring alignment, including any algorithm needed to achieve maximal alignment over the full length of the sequences being compared.

For purposes herein, % amino acid sequence identity values refer to values generated using the pair wise sequence alignment program EMBOSS Needle, that creates an optimal global alignment of two sequences using the Needleman-Wunsch algorithm, wherein all search parameters are set to default values, i.e. Scoring matrix = BLOSUM62, Gap open = 10, Gap extend = 0.5, End gap penalty = false, End gap open = 10 and End gap extend = 0.5.

[0044] A “substitution” means that an amino acid residue is replaced by another amino acid residue. Preferably, the term “substitution” refers to the replacement of an amino acid residue by another selected from the naturally-occurring standard 20 amino acid residues, rare naturally occurring amino acid residues (e.g. hydroxyproline, hydroxylysine, allohydroxylysine, 6-N-methyllysine, N-ethylglycine, N-methylglycine, N-ethylasparagine, allo-isoleucine, N-methylisoleucine, N-methylvaline, pyroglutamine, aminobutyric acid, ornithine, norleucine, norvaline), and non-naturally occurring amino acid residue, often made synthetically, (e.g. cyclohexyl-alanine). Preferably, the term “substitution” refers to the replacement of an amino acid residue by another selected from the naturally-occurring standard 20 amino acid residues. The sign “+” indicates a combination of substitutions. The amino acids are herein represented by their one-letter or three-letters code according to the following nomenclature: A: alanine (Ala); C: cysteine (Cys); D: aspartic acid (Asp); E:

glutamic acid (Glu); F: phenylalanine (Phe); G: glycine (Gly); H: histidine (His); I: isoleucine (Ile); K: lysine (Lys); L: leucine (Leu); M: methionine (Met); N: asparagine (Asn); P: proline (Pro); Q: glutamine (Gln); R: arginine (Arg); S: serine (Ser); T: threonine (Thr); V: valine (Val); W: tryptophan (Trp) and Y: tyrosine (Tyr). In the present document, the following terminology is used to designate a substitution: L238A denotes that amino acid residue (Leucine, L) at position 238 of the parent sequence is changed to an Alanine (A). A132V/I/M denotes that amino acid residue (Alanine, A) at position 132 of the parent sequence is substituted by one of the following amino acids: Valine (V), Isoleucine (I), or Methionine (M). The substitution can be a conservative or non-conservative substitution. Examples of conservative substitutions are within the groups of basic amino acids (arginine, lysine and histidine), acidic amino acids (glutamic acid and aspartic acid), polar amino acids (glutamine, asparagine and threonine), hydrophobic amino acids (methionine, leucine, isoleucine, cysteine and valine), aromatic amino acids (phenylalanine, tryptophan and tyrosine), and small amino acids (glycine, alanine and serine).

[0045] This disclosure is not intended to be limited to the scope of the particular forms set forth, but is intended to cover alternatives, modifications, and equivalents of the variations described herein. Further, the scope of the disclosure fully encompasses other variations that may become obvious to those skilled in the art in view of this disclosure. The scope of the present invention is limited only by the appended claims.

CLAIMS

1. A method of synthesizing a polynucleotide having a predetermined sequence, the method comprising the steps of:

5 a) providing an initiator having a 3'-penultimate deoxyinosine and a 3'-terminal nucleotide having a free 3'-hydroxyl;

b) repeating cycles of (i) contacting under elongation conditions the initiator or elongated fragments having free 3'-O-hydroxyls with a 3'-O-blocked nucleoside triphosphate and a template-independent DNA polymerase so that the initiator or elongated fragments are
10 elongated by incorporation of a 3'-O-blocked nucleoside triphosphate to form 3'-O-blocked elongated fragments, and (ii) deblocking the elongated fragments to form elongated fragments having free 3'-hydroxyls, until the polynucleotide is formed;

c) treating the polynucleotide with an endonuclease V activity to cleave the polynucleotide from the initiator.

15

2. The method of claim 1, wherein said endonuclease V activity is provided by a prokaryotic endonuclease V.

3. The method of claim 2, wherein said prokaryotic endonuclease V is an E. coli
20 endonuclease V.

4. The method of claim 2 or 3, further comprising a step of removing said prokaryotic endonuclease V from said cleaved polynucleotide.

25 5. The method of anyone of claims 1 to 4, wherein said template-independent DNA polymerase is a terminal deoxynucleotidyl transferase.

6. The method of anyone of claims 1 to 5, wherein said initiator is attached to a support by a 5' end.

30

7. The method of claim 6, wherein said support is a solid support.

8. The method of anyone of claims 1 to 7, wherein said initiator has a 3'-terminal sequence of 5'-dI-dT-3'.

5 9. The method of anyone of claims 1 to 8, wherein said polynucleotide cleaved from said initiator has a 5'-monophosphate.

10. A kit for enzymatically synthesizing a polynucleotide comprising an initiator attached to a support by a 5' end and having a 3'-penultimate deoxyinosine and a 3'-terminal nucleotide
10 having a free 3'-hydroxyl.

11. The kit of claim 10 further comprising an endonuclease V capable of cleaving an initiator-polynucleotide conjugate 3' of a terminal nucleotide of the initiator.

15 12. The kit of claims 10 or 11, further comprising said initiator with a 3'-terminal sequence of 5'-dI-dT-3'.

13. The kit of anyone claims 10 to 12, wherein said support is a solid support.

20 14. The kit of anyone of claims 10 to 13, further comprising a template-free polymerase and 3'-O-blocked nucleoside triphosphate for one or more of deoxyadenosine, deoxyguanosine, thymidine, deoxyuridine and deoxycytidine.

15. The kit of anyone of claims 10 to 14, further comprising a de-blocking agent.

25

16. The kit of claim 14 wherein said 3'-O-blocked nucleoside triphosphates are 3'-O-NH₂-nucleoside triphosphates.