

TEMPLATE-FREE ENZYMATIC SYNTHESIS OF POLYNUCLEOTIDES
USING POLY(A) AND POLY(U) POLYMERASES

5

BACKGROUND

10 [0001] Interest has arisen in supplementing or replacing chemically-based synthesis methods by enzymatically-based methods using template-free polymerases, such as, terminal deoxynucleotidyl transferase (TdT), because of the proven efficiency of such enzymes and the benefit of mild non-toxic reaction conditions, e.g. Ybert et al, International patent publication WO2015/159023; Hiatt et al, U.S. patent 5763594; Jensen et al, Biochemistry, 57: 1821-1832
15 (2018); and the like. Most approaches in enzyme-based synthesis are limited to DNA synthesis and require the use of reversibly blocked nucleoside triphosphates in order to obtain a desired sequence in the polynucleotide product. Unfortunately, natural TdTs incorporate such modified nucleoside triphosphates with reduced efficiency as compared to unmodified nucleoside triphosphates. Thus, a great deal of work has been directed to developing new TdT
20 variants with better incorporation efficiencies for modified nucleoside triphosphates to synthesize DNA, e.g. Champion et al, U.S. patent publication US2019/0211315; Ybert et al, International patent publication WO2017/216472, and the like.

[0002] In view of the above, the field of template-free enzymatically-based polynucleotide synthesis would be advanced if new template-free polymerases were available for synthesize
25 DNA and RNA with improved incorporation of reversibly blocked nucleoside triphosphates.

SUMMARY OF THE INVENTION

[0003] The present invention is directed to methods, kits and compositions for template-free enzymatic synthesis of polynucleotides using poly(A) and poly(U) polymerases and
30 variants thereof. In some embodiments, methods employing poly(A) and poly(U) polymerases are used to synthesize RNA products of a predetermined sequence. In other embodiments, methods employing poly(A) and poly(U) polymerases are used to synthesize DNA products of a predetermined sequence.

[0004] In some embodiments, the invention is directed to methods of synthesizing a polynucleotide having a predetermined sequence comprising the steps of: a) providing an initiator having a 3'-terminal nucleotide having a free 3'-hydroxyl; and b) repeating, until the polynucleotide is formed, cycles of (i) contacting under elongation conditions the initiator or
5 elongated fragments having free 3'-hydroxyls with a 3'- O-blocked-nucleoside triphosphate and a template-free polymerase so that the initiator or elongated fragments are 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; wherein the template-free polymerase is a poly(A) polymerase (PAP) or a
10 poly(U) polymerase. In further embodiments, the initiator may be attached to a support by a 5' end. In further embodiments, the support may be a solid support.

[0005] In some embodiments, the above method may include a step of cleaving the polynucleotide from the initiator.

[0006] In some embodiments, the polynucleotide is a poly-2'-deoxyribonucleotide and the
15 3'- O-blocked-nucleoside triphosphate is a 3'-O-blocked-2'-deoxyribonucleoside triphosphate. In further embodiments, the 3'- O-blocked-2'-deoxyribonucleoside triphosphate is a 3'-O-azidomethyl-2'-deoxyribonucleoside triphosphate or a 3'-O-amino-2'-deoxyribonucleoside triphosphate. In further embodiments, the polynucleotide is a polyribonucleotide and said 3'-O-blocked-nucleoside triphosphate is a 3'- O-blocked-ribonucleoside triphosphate. In some
20 embodiments, the 3'- O-blocked-ribonucleoside triphosphate is a 3'- O-azidomethyl-ribonucleoside triphosphate. In further embodiments, the 3'-azidomethyl-O-ribonucleoside triphosphate is selected from the group consisting of 3'-azidomethyl-O-adenosine triphosphate, 3'-azidomethyl-O-guanosine triphosphate, and 3'-azidomethyl-O-cytidine triphosphate, 3'-azidomethyl-O-uridine triphosphate.

25 In some embodiments, the poly(A) polymerase employed in the method of the invention is a poly(A) polymerase variant comprising an amino acid sequence that is at least 90 percent identical to an amino acid sequence selected from SEQ ID NO: 1, 2, 3, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26 or 27 with a substitution of methionine at position 310 with respect to SEQ ID NO: 1, at position 318 with respect to SEQ ID NO: 2,
30 at position 318 with respect to SEQ ID NO: 3, at position 316 with respect to SEQ ID NO: 8, at position 309 with respect to SEQ ID NO: 9, at position 316 with respect to SEQ ID NO: 10,

[0007] at position 272 with respect to SEQ ID NO: 11, at position 316 with respect to SEQ ID NO: 12, at position 307 with respect to SEQ ID NO: 13, at position 313 with respect to SEQ ID NO: 14, at position 312 with respect to SEQ ID NO: 15, at position 317 with respect to SEQ ID NO: 16, at position 316 with respect to SEQ ID NO: 17, at position 316 with respect to SEQ ID NO: 18, at position 312 with respect to SEQ ID NO: 19, at position 310 with respect to SEQ ID NO: 20, at position 309 with respect to SEQ ID NO: 21, at position 317 with respect to SEQ ID NO: 22, at position 314 with respect to SEQ ID NO: 23, at position 307 with respect to SEQ ID NO: 24, at position 315 with respect to SEQ ID NO: 25, at position 316 with respect to SEQ ID NO: 26, and at position 311 with respect to SEQ ID NO: 27. In particular, said substitution of said methionine at said position may be selected from F, Y, V, E or T. Said amino acid sequence selected from said SEQ ID NO: 1, 2, 3, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26 or 27 may further comprise a substitution of valine at position 234 with respect to SEQ ID NO: 1, at position 240 with respect to SEQ ID NO: 2, at position 240 with respect to SEQ ID NO: 3, at position 232 with respect to SEQ ID NO: 9, at position 240 with respect to SEQ ID NO: 10, at position 196 with respect to SEQ ID NO: 11, at position 240 with respect to SEQ ID NO: 12, at position 229 with respect to SEQ ID NO: 13, at position 236 with respect to SEQ ID NO: 14, at position 236 with respect to SEQ ID NO: 15, at position 241 with respect to SEQ ID NO: 16, at position 233 with respect to SEQ ID NO: 17, at position 240 with respect to SEQ ID NO: 18, at position 240 with respect to SEQ ID NO: 19, at position 234 with respect to SEQ ID NO: 20, at position 233 with respect to SEQ ID NO: 21, at position 237 with respect to SEQ ID NO: 22, at position 238 with respect to SEQ ID NO: 23, at position 231 with respect to SEQ ID NO: 24, at position 239 with respect to SEQ ID NO: 25, at position 240 with respect to SEQ ID NO: 26, and at position 235 with respect to SEQ ID NO: 27. In particular, said substitution of said valine at said position may be alanine or glycine.

[0008] In further embodiments of the above composition, the amino acid sequence at least 90 percent identical to SEQ ID NO: 2 or 3 further comprises a substitution of alanine at position 410 with respect to SEQ ID NO: 2 or 3. In some embodiments, the substitution of alanine at position 410 is valine.

[0009] In one aspect, the invention is directed to methods and kits for synthesizing a polyribonucleotide having a predetermined sequence comprising the steps of: a) providing an

initiator having a 3'-terminal nucleotide having a free 3'-hydroxyl; and b) repeating cycles of (i) contacting under elongation conditions the initiator or elongated fragments having free 3'-O-hydroxyls with a 3'-O-blocked ribonucleoside triphosphate and a poly(A) polymerase so that the initiator or elongated fragments are elongated by incorporation of a 3'-O-blocked ribonucleoside 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. In some embodiments, methods of the invention further include a step of cleaving the polynucleotide from the initiator. In some embodiments, the 3'-O-blocked ribonucleoside triphosphate is a 3'-O-azidomethyl-ribonucleoside triphosphate.

In another aspect, the invention is directed to a kit for performing template-free synthesis of a polyribonucleotide having a predetermined sequence, the kit comprising a poly(A) polymerase, an initiator attached to a solid support, and 3'-O-blocked ribonucleoside triphosphate monomers. Said 3'-O-protected ribonucleoside triphosphate monomers may comprise one or more of 3'-O-azidomethyl-riboadenosine triphosphate, 3'-O-azidomethyl-riboguanosine triphosphate, 3'-O-azidomethyl-ribocytidine triphosphate, 3'-O-azidomethyl-ribothymidine triphosphate and 3'-O-amino-ribouridine triphosphate. Said poly(A) polymerase may comprise a poly(A) polymerase variant having an amino acid sequence that is at least 90 percent identical to SEQ ID NO: 1 and having a substitution at M310, or at least 90 percent identical to SEQ ID NO: 3 and having a substitution at M318; wherein the poly(A) polymerase variant is capable of (a) synthesizing a ribonucleic acid fragment without a template and (b) incorporating said 3'-O-azidomethyl-ribonucleoside triphosphate onto a ribonucleic acid fragment.

In another aspect, the invention is directed to poly(A) polymerase variants comprising an amino acid sequence that is at least 90 percent identical to an amino acid sequence selected from SEQ ID NO: 1, 2, 3, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26 or 27 with a substitution of methionine at position 310 with respect to SEQ ID NO: 1, at position 318 with respect to SEQ ID NO: 2, at position 318 with respect to SEQ ID NO: 3, at position 316 with respect to SEQ ID NO: 8, at position 309 with respect to SEQ ID NO: 9, at position 316 with respect to SEQ ID NO: 10, at position 272 with respect to SEQ ID NO: 11, at position 316 with respect to SEQ ID NO: 12, at position 307 with respect to SEQ ID NO: 13, at position 313 with respect to SEQ ID NO: 14, at position 312 with respect to SEQ

ID NO: 15, at position 317 with respect to SEQ ID NO: 16, at position 316 with respect to
 SEQ ID NO: 17, at position 316 with respect to SEQ ID NO: 18, at position 312 with respect
 to SEQ ID NO: 19, at position 310 with respect to SEQ ID NO: 20, at position 309 with
 respect to SEQ ID NO: 21, at position 317 with respect to SEQ ID NO: 22, at position 314
 5 with respect to SEQ ID NO: 23, at position 307 with respect to SEQ ID NO: 24, at position
 315 with respect to SEQ ID NO: 25, at position 316 with respect to SEQ ID NO: 26, and
 at position 311 with respect to SEQ ID NO: 27, wherein the PAP variant is capable of (a)
 synthesizing a ribonucleic acid fragment without a template and (b) incorporating a 3'-O-
 azidomethyl-ribonucleoside triphosphate onto a ribonucleic acid fragment or a 3'-O-
 10 azidomethyl-2'-deoxyribonucleoside triphosphate into a deoxyribonucleic acid fragment. In
 particular, said substitution of said methionine at said position may be selected from F, Y, V, E
 or T. Poly(A) polymerase variants may further comprise a substitution of valine at position
 234 with respect to SEQ ID NO: 1, at position 240 with respect to SEQ ID NO: 2, at position
 240 with respect to SEQ ID NO: 3, at position 232 with respect to SEQ ID NO: 9, at position
 15 240 with respect to SEQ ID NO: 10, at position 196 with respect to SEQ ID NO: 11, at
 position 240 with respect to SEQ ID NO: 12, at position 229 with respect to SEQ ID NO: 13,
 at position 236 with respect to SEQ ID NO: 14, at position 236 with respect to SEQ ID NO:
 15, at position 241 with respect to SEQ ID NO: 16, at position 233 with respect to SEQ ID
 NO: 17, at position 240 with respect to SEQ ID NO: 18, at position 240 with respect to SEQ
 20 ID NO: 19, at position 234 with respect to SEQ ID NO: 20, at position 233 with respect to
 SEQ ID NO: 21, at position 237 with respect to SEQ ID NO: 22, at position 238 with respect
 to SEQ ID NO: 23, at position 231 with respect to SEQ ID NO: 24, at position 239 with
 respect to SEQ ID NO: 25, at position 240 with respect to SEQ ID NO: 26, and at position 235
 with respect to SEQ ID NO: 27. In particular, said substitution of valine at said position may
 25 be alanine or glycine.

[0010] In another aspect, the invention is directed to poly(U) polymerase (PUP) variants
 each comprising an amino acid sequence that is at least 90 percent identical to an amino acid
 sequence selected from SEQ ID NO: 4, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41,
 42, 43, 44, 45, 46 or 47 with a substitution of tyrosine at position 212 with respect to SEQ ID
 30 NO: 4, at position 189 with respect to SEQ ID NO: 28, at position 184 with respect to SEQ ID
 NO: 29, at position 227 with respect to SEQ ID NO: 30, at position 478 with respect to SEQ

ID NO: 31, at position 192 with respect to SEQ ID NO: 32, at position 186 with respect to SEQ ID NO: 33, at position 243 with respect to SEQ ID NO: 34, at position 196 with respect to SEQ ID NO: 35, at position 253 with respect to SEQ ID NO: 36, at position 284 with respect to SEQ ID NO: 37, at position 182 with respect to SEQ ID NO: 38, at position 187 with respect to SEQ ID NO: 39, at position 203 with respect to SEQ ID NO: 40, at position 224 with respect to SEQ ID NO: 41, at position 204 with respect to SEQ ID NO: 42, at position 337 with respect to SEQ ID NO: 43, at position 296 with respect to SEQ ID NO: 44, at position 291 with respect to SEQ ID NO: 45, at position 218 with respect to SEQ ID NO: 46, and at position 366 with respect to SEQ ID NO: 47, wherein the PUP variant is capable of (a) synthesizing a ribonucleic acid fragment without a template and (b) incorporating a 3'-O-azidomethyl-ribonucleoside triphosphate onto a ribonucleic acid fragment or a 3'-O-azidomethyl-2'-deoxyribonucleoside triphosphate into a deoxyribonucleic acid fragment. In particular, said substitution of said tyrosine at said position may be alanine or glycine.

[0011] The present invention advantageously overcomes the above problems in the field of enzymatic polynucleotide synthesis by providing template-free polymerases and 3'-O-modified ribonucleoside triphosphates that provide higher rates of ribonucleotide incorporation into a growing RNA fragment.

BRIEF DESCRIPTION OF THE DRAWINGS

[0012] Fig. 1 illustrates steps of an embodiment of the method of the invention.

[0013] Figs. 2A-2C shows data for extensions of from 1-5 ribonucleotides in accordance with the method of the invention.

[0014] Figs. 3A-3B show data for extensions with a single 3'-O-protected-rATP (3A) and of a polyguanylation using rGTP (3B) in presence of mutations M310F/Y/Q/A/V/E/T.

[0015] Fig. 4 is an electropherogram of incorporation products produced by *Schizosaccharomyces pombe* PUP. "No ext" means that the indicated column shows the band corresponding to the primer without any extension.

DETAILED DESCRIPTION OF THE INVENTION

[0016] 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.

[0017] 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.

[0018] The invention is directed to methods and compositions for synthesizing polynucleotides, either polyribonucleic acids or polydeoxyribonucleic acids, using a poly(A) polymerase (PAP), poly(U) polymerase (PUP) or both PAPs and PUPs in the same synthesis. In some embodiments, PAPs and/or PUPs are used to synthesize a polyribonucleic acid using 3'-O-reversibly protected-rNTP precursors, wherein a single PUP or PAP variant may be employed for coupling all ribonucleoside triphosphate monomers, or in alternative embodiments, wherein different PUPs and PAPs may be employed for coupling different kinds of ribonucleoside triphosphate monomers in the synthesis of a particular RNA. Likewise, in other embodiments, PAPs and/or PUPs be used to synthesize a polydeoxyribonucleic acid using 3'-O-reversibly protected-dNTP precursors, wherein a single PUP or PAP is employed for coupling all deoxyribonucleoside triphosphate (dNTP) monomers, or in an alternative embodiment, wherein different PUP and PAP polymerases may be employed for coupling different kinds of deoxyribonucleoside triphosphate monomers. In some embodiments, methods of the invention employ PAP and/or PUP variants that have been modified by genetic

engineering to improve efficiency of coupling 3'-O-blocked-ribonucleoside triphosphates and 3'-O-blocked-2'-deoxyribonucleoside triphosphates to growing polynucleotide chains in a synthesis.

[0019] Methods of the invention, whether for synthesizing RNA or DNA, comprise

- 5 repeated cycles of steps, such as are illustrated in Fig. 1, in which a predetermined ribonucleotide monomer for RNA synthesis (or 2'-deoxyribonucleotide monomer for DNA synthesis) is added in each cycle. Initiator polynucleotides (100) are provided, for example, attached to solid support (102), which have free 3'-hydroxyl groups (103). For synthesizing RNA, typically initiators are polyribonucleotides, and for synthesizing DNA, typically
- 10 initiators are polydeoxyribonucleotides. To the initiator polynucleotides (100) (or elongated initiator polynucleotides in subsequent cycles) are added a 3'-O-reversibly protected-rNTP (or 3'-O-reversibly protected-dNTP in case of DNA synthesis) and a PAP or PUP under conditions (104) effective for the enzymatic incorporation of the 3'-O-protected-rNTP (or 3'-O-protected dNTP) onto the 3' end of the initiator polynucleotides (100) (or elongated initiator
- 15 polynucleotides). This reaction produces elongated initiator polynucleotides whose 3'-hydroxyls are protected (106). If the elongated initiator polynucleotide contains a completed sequence, then the 3'-O-protection group is removed, or deprotected, and the desired sequence is cleaved from the original initiator polynucleotide. Such cleavage may be carried out using any of a variety of single strand cleavage techniques, for example, by inserting a cleavable
- 20 nucleotide at a predetermined location within the original initiator polynucleotide. An exemplary cleavable nucleotide may be a uracil nucleotide which is cleaved by uracil DNA glycosylase. Additionally a wide variety of cleavable linkages or cleavable nucleotides may be used for this purpose. In some embodiments, cleaving the desired polynucleotide leaves a natural free 5'-hydroxyl on a cleaved strand; however, in alternative embodiments, a cleaving
- 25 step may leave a moiety, e.g. a 5'-phosphate, that may be removed in a subsequent step, e.g. by phosphatase treatment. Cleaving steps may be carried out chemically, thermally, enzymatically or by photochemical methods. In some embodiments, cleavable nucleotides may be nucleotide analogs such as deoxyuridine or 8-oxo-deoxyguanosine that are recognized by specific glycosylases (e.g. uracil deoxyglycosylase followed by endonuclease VIII, and 8-
- 30 oxoguanine DNA glycosylase, respectively). In some embodiments, cleavage may be accomplished by providing initiators with a deoxyinosine as the penultimate 3' nucleotide,

which may be cleaved by endonuclease V at the 3' end of the initiator leaving a 5'-phosphate on the released polynucleotide. Further methods for cleaving single stranded polynucleotides are disclosed in the following references, which are incorporated by reference: U.S. Pat. Nos. 5,739,386, 5,700,642 and 5,830,655; and U.S. Patent Publication Nos. 2003/0186226 and 2004/0106728; and in Urdea and Horn, U.S. patent 5367066.

[0020] If the elongated initiator polynucleotide does not contain a completed sequence, then the 3'-O-protection groups are removed to expose free 3'-hydroxyls (103) and the elongated initiator polynucleotides are subjected to another cycle of ribonucleotide addition and deprotection.

[0021] In some embodiments, the method of synthesizing an oligoribonucleotide comprises the steps of (a) providing an initiator having a free 3'-hydroxyl; (b) reacting under extension conditions the initiator or an extension intermediate (i.e. an elongated initiator polynucleotide) having a free 3'-hydroxyl with a PAP in the presence of a 3'-O-blocked ribonucleoside triphosphate to produce a 3'-O-blocked extension intermediate; (c) deblocking the extension intermediate to produce an extension intermediate with a free 3'-hydroxyl; and (d) repeating steps (b) and (c) until the polyribonucleotide is synthesized. (The term "an extension intermediate" may also be referred to herein as an "elongation fragment" or an "elongated initiator polynucleotide".) In some embodiments, as noted above, 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 reacting may include a sub-step of removing unincorporated ribonucleoside 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.

[0022] In some embodiments, the method of synthesizing an oligoribonucleotide comprises the steps of (a) providing an initiator having a free 3'-hydroxyl; (b) reacting under extension conditions the initiator or an extension intermediate (i.e. an elongated initiator polynucleotide) having a free 3'-hydroxyl with a PAP or a PUP in the presence of a 3'-O-blocked ribonucleoside triphosphate to produce a 3'-O-blocked extension intermediate; (c) deblocking the extension intermediate to produce an extension intermediate with a free 3'-hydroxyl; and (d) repeating steps (b) and (c) until the polyribonucleotide is synthesized. (The

term “an extension intermediate” may also be referred to herein as an “elongation fragment” or an “elongated initiator polynucleotide”). In some embodiments, as noted above, 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 reacting may include a sub-step of removing unincorporated ribonucleoside 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.

[0023] In some embodiments, the method of synthesizing a polynucleotide comprises the steps of (a) providing an initiator having a free 3'-hydroxyl; (b) reacting under extension conditions the initiator or an extension intermediate (i.e. an elongated initiator polynucleotide) having a free 3'-hydroxyl with a PAP or a PUP 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 with a free 3'-hydroxyl; and (d) repeating steps (b) and (c) until the polynucleotide is synthesized. (The term “an extension intermediate” may also be referred to herein as an “elongation fragment” or an “elongated initiator polynucleotide”). In some embodiments, as noted above, 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 reacting may include a sub-step of removing unincorporated ribonucleoside 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.

[0024] 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'-exoribonuclease activity, e.g. RNase R (Epicentre). 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.

[0025] In some embodiments that comprise serial synthesis of oligoribonucleotides, 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 oligoribonucleotides will be present in equal molar amounts. This would not be the case where non-extended fragments are capped.

[0026] In some embodiments, reaction conditions for an extension or elongation step using PAP or PUP may comprising the following: Reaction conditions 1 (for primer+AM-rATP): 250 uM AM-rATP, 0.1 uM ATTO488-(rA)₅, 1 uM PAP, 1x ATP buffer (20 mM Tris-HCl, 0.6 mM MnCl₂, 0.02 mM EDTA, 0.1% BSA, 10% glycerol, 100 mM imidazole, pH 7-8), 37 C, 30 min. Reaction condition 2 (for primer+AM-rGTP): 250 uM rGTP, 0.1 uM ATTO488-(rA)₅, 1 uM PAP, 1x GTP buffer (0.6 mM MnCl₂, 0.1% BSA, 10 mM imidazole, pH 6), 37 C, 30 min. In the foregoing, "AM-rNTP" refers to 3'-azidomethyl-O-ribonucleoside triphosphate.

[0027] 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 references, such as Wuts, Green's Protection Groups in Organic Chemistry, 5th Edition (Wiley 2014). 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 group, palladium complexes can be used to cleave 3'-O-allyl group and 3'-O-propargyl group, or sodium nitrite can be used to cleave a 3'-O-amino group.

[0028] 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 in which two or more sequences are synthesized in the same reaction mixture. 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-propargyl
3'-O-NH ₂	3'-O-phosphate
3'-O-azidomethyl	3'-O-allyl, 3'-O-propargyl
3'-O-azidomethyl	3'-O-phosphate
3'-O-allyl, 3'-O-propargyl	3'-O-phosphate

[0029] 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 PAP or PUP. In one embodiment, the initiating fragment is a DNA or RNA 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 or agarose 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.

[0030] Many of the 3'-O-blocked rNTPs employed in the invention may be purchased from commercial vendors (e.g. Jena Bioscience, MyChemLabs, or the like) 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
5 Acids Research, 22: 4259-4267 (1994); Meng et al, J. Org. Chem., 14: 3248-3252 (3006); U.S. patent publication 2005/037991; Zavgorodny et al, Tetrahedron Letters, 32(51): 7593-7596 (1991).

[0031] In a further particular embodiments, the 3'-blocked nucleotide triphosphate is blocked by either 3'-O-propargyl, a 3'-O-azidomethyl, 3'-O-NH₂ or 3'-O-allyl group.

10 [0032] 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 ethoxy, 3'-O-(2-cyanoethyl), and 3'-O-propargyl.

15

Poly(A) Polymerases and Useful Variants

[0033] A wide variety of PAPs may be used with the method of the invention, including PAP variants that have been engineered for improved characteristics, such as, higher incorporation rates of 3'-O-protected-rNTPs (including for particular protection groups, such as, 3'-O-azidomethyl), greater stability and shelf life, thermostability, solubility, and the like.

20

In particular, a yeast PAP with a mutation at M310 (SEQ ID NO: 1), or a functionally equivalent residue in other PAPs, such as PAPs from various different species, shows improved incorporation of 3'-O-protected rNTPs with respect to a wildtype PAP. In some embodiments, a yeast PAP variant of the invention has an amino acid sequence of SEQ ID NO: 1 except for a substitution at M310. In some embodiments, such substitution is selected
25 from M310F/Y/V/E/T. In particular, substitutions M310F/Y allow the incorporation of 3'-O-amino-rATPs and substitutions M310V/E/T improve the rate of incorporation of 3'-O-protected-rGTPs. In other embodiments, a yeast PAP variant of the invention has an amino acid sequence with at least 90 percent identity of SEQ ID NO: 1 except for a substitution at M310.

30

[0034] PAP variants for use with the invention include those listed in Table 1 below. In some embodiments PAP variants of the invention comprise at least a substitution at the second

position indicated in Table 1. In other embodiments, embodiments of PAP variants of the invention comprise at least a substitution at the first position indicated in Table 1.

Table 1

PAP Variants: Positions of Substitutions

SEQ ID NO	Organism	First Position	Second Position
1	yeast	V234	M310
2	Myceliophthora	V240	M318
3	Thielavia	V240	M318
8	Pyronema	I237	M316
9	Tilletia	V232	M309
10	Clathrospora	V240	M316
11	Drechslerella	V196	M272
12	Magnaporthiopsis	V240	M316
13	Cryptococcus	V229	M307
14	Golovinomyces	V236	M313
15	Hortaea	V236	M312
16	Valsa	V241	M317
17	Wallemia	V233	M316
18	Xylaria	V240	M316
19	Chaetomium	V240	M312
20	Lachancea	V234	M310
21	Schizosaccharomyces	V233	M309
22	Exophiala	V237	M317
23	Scedosporium	V238	M314
24	Trichoderma	V231	M307
25	Aspergillus	V239	M315
26	Sodiomyces	V240	M316
27	Neohortaea	V235	M311

[0035] In some embodiments, a substitution at a first position as indicated in Table 1 is A or G (thus, for example, for SEQ ID NO:1, the substitution may be written V234A/G). In

10 some embodiments, a substitution at a second position as indicated in Table 1 is F, Y, V, E, or T (thus, for example, for SEQ ID NO: 1, the substitution may be written M310F/Y/V/E/T)

[0036] In some embodiments, a PAP variant of the invention has one or more of the substitutions of Table 1 and a percent identity value of at least 80 percent identity with the indicated SEQ ID NO; in some embodiments, the above percent identity value is at least 90

15 percent identity with the indicated SEQ ID NO; in some embodiments, the above percent

identity value is at least 95 percent identity with the indicated SEQ ID NO; 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.

- 5 **[0037]** In some embodiments, a thermostable PAP is employed so that the method may be practiced at a temperature that reduces or eliminates the formation of secondary structures in the RNA or DNA being synthesized. In some embodiments, the temperature range within which the highest incorporation rate occurs for the thermostable PAP is higher than 40°C. In some embodiments, the temperature range within which the highest incorporation rate occurs
- 10 for the thermostable PAP is higher than 50°C. In some embodiments, the temperature range within which the highest incorporation rate occurs for the thermostable PAP is between 40°C and 85°C. In some embodiments, the temperature range within which the highest incorporation rate occurs for the thermostable PAP is between 50°C and 85°C.

15 Poly(U) Polymerases and Useful Variants

- [0038]** As with PAPs, a wide variety of PUPs may be used with the method of the invention, including PUP variants that have been engineered for improved characteristics, such as, higher incorporation rates of 3'-O-protected-rNTPs (including for particular protection groups, such as, 3'-O-azidomethyl), greater stability and shelf life, thermostability, solubility,
- 20 and the like. PUP variants for use with the invention include those listed in Table 2 below. In some embodiments PUP variants of the invention comprise at least a substitution at the first position indicated in Table 2. In other embodiments, embodiments of PAP variants of the invention comprise at least a substitution at the second position indicated in Table 2.

25

Table 2

PUP Variants: Positions of Substitutions

SEQ ID NO	Organism	First Position	Second Position
4	<i>S. pombe</i>	Y212	H336
28	<i>T. brucei</i>	Y189	L303
29	<i>S. pombe</i>	Y184	H308
30	<i>T. boudieri</i>	Y227	H364
31	<i>D. stenobrocha</i>	Y478	H613

32	Phytomonas	Y192	L306
33	B. saltans	Y186	L326
34	A. deanei	Y243	L392
35	P. lactucaedebilis	Y196	H330
36	S. culicis	Y253	L392
37	B. meristosporus	Y284	H408
38	N. californiae	Y182	H310
39	Perkinsela	Y187	L394
40	S. complicate	Y203	H331
41	S. ochraceum	Y224	F349
42	G. androsaceus	Y204	Y332
43	T. equiperdum	Y337	L473
44	M. conica	Y296	H431
45	P. murina	Y291	H423
46	S. japonicus	Y218	H340
47	A. nigricans	Y366	H509

[0039] In some embodiments, a substitution at a first position as indicated in Table 2 is A or G (thus, for example, for SEQ ID NO:4, the substitution may be written Y212A/G). In

5 some embodiments, a substitution at a second position as indicated in Table 2 is F, Y, V, E, or T (thus, for example, for SEQ ID NO: 4, the substitution may be written H336F/Y/V/E/T)

[0040] In some embodiments, a PUP variant of the invention has one or more of the substitutions of Table 2 and a percent identity value of at least 80 percent identity with the indicated SEQ ID NO; in some embodiments, the above percent identity value is at least 90
10 percent identity with the indicated SEQ ID NO; in some embodiments, the above percent identity value is at least 95 percent identity with the indicated SEQ ID NO; 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.

15 **[0041]** In some embodiments, a thermostable PUP is employed so that the method may be practiced at a temperature that reduces or eliminates the formation of secondary structures in the RNA or DNA being synthesized. In some embodiments, the temperature range within which the highest incorporation rate occurs for the thermostable PUP is higher than 40°C. In some embodiments, the temperature range within which the highest incorporation rate occurs
20 for the thermostable PUP is higher than 50°C. In some embodiments, the temperature range

within which the highest incorporation rate occurs for the thermostable PUP is between 40°C and 85°C. In some embodiments, the temperature range within which the highest incorporation rate occurs for the thermostable PUP is between 50°C and 85°C.

5

Production of PAP and PUP Variants

[0042] Variants of the invention may be produced by mutating known reference or wild type PAP-coding or PUP-coding polynucleotides, then expressing it using conventional molecular biology techniques. For example, a desired gene or DNA fragment encoding a polypeptide of desired sequence may be assembled from synthetic fragments using conventional molecular biology techniques, e.g. using protocols described by Stemmer et al, Gene, 164: 49-53 (1995); Kodumal et al, Proc. Natl. Acad. Sci., 101: 15573-15578 (2004); or the like, or such gene or DNA fragment may be directly cloned from cells of a selected species using conventional protocols.

15 [0043] An isolated gene encoding a desired PAP or PUP variant may be inserted into an expression vector to give an expression vector which then may be used to make and express variant PAP or PUP proteins using conventional protocols. Vectors with the correct sequence may be transformed in E. coli producer strains.

[0044] The following procedures are described for PAP variants, but similar procedures may be applied by one of ordinary skill in the art to PUP variants. Transformed strains are cultured using conventional techniques to pellets from which PAP protein is extracted. For example, previously prepared pellets are thawed in 30 to 37°C water bath. Once fully thawed, pellets are resuspended in lysis buffer composed of 50mM tris-HCL (Sigma) pH 7.5, 150mM NaCl (Sigma), 0.5mM mercaptoethanol (Sigma), 5% glycerol (Sigma), 20mM imidazole (Sigma) and 1 tab for 100mL of protease cocktail inhibitor (Thermofisher). Careful resuspension is carried out in order to avoid premature lysis and remaining of aggregates. Resuspended cells are lysed through several cycles of French press, until full color homogeneity is obtained. Usual pressure used is 14,000psi. Lysate is then centrifuged for 1h to 1h30 at 10,000 rpm. Centrifugate is pass through a 0.2µm filter to remove any debris before column purification.

30

- [0045] PAP protein may be purified from the centrifugate in a one-step affinity procedure. For example, Ni-NTA affinity column (GE Healthcare) may be used to bind the PAP polymerases. Initially the column is washed and equilibrated with 15 column volumes of 50mM tris-HCL (Sigma) pH 7.5, 150mM NaCl (Sigma) and 20mM imidazole (Sigma). PAP polymerases are bound to the column after equilibration; then, a washing buffer, for example, composed of 50mM tris-HCL (Sigma) pH 7.5, 500mM NaCl (Sigma) and 20mM imidazole (Sigma), may be applied to the column for 15 column volumes. After such washing, the PAP polymerases are eluted with 50mM tris-HCL (Sigma) pH 7.5, 500mM NaCl (Sigma) and 0.5M imidazole (Sigma). Fractions corresponding to the highest concentration of PAP polymerases of interest are collected and pooled in a single sample. The pooled fractions are dialyzed against the dialysis buffer (20 mM Tris-HCl, pH 6.8, 200mM Na Cl, 50mM MgOAc, 100mM [NH₄]₂SO₄). The dialysate is subsequently concentrated with the help of concentration filters (Amicon Ultra-30, Merk Millipore). Concentrated enzyme is distributed in small aliquots, 50% glycerol final is added, and those aliquots are then frozen at -20°C and stored for long term. 5μL of various fraction of the purified enzymes are analyzed in SDS-PAGE gels.
- [0046] In some embodiments, a PAP variant may be operably linked to a linker moiety including a covalent or non-covalent bond; amino acid tag (e.g., poly-amino acid tag, poly-His tag, 6His-tag, or the like); chemical compound (e.g., polyethylene glycol); protein-protein binding pair (e.g., biotin-avidin); affinity coupling; capture probes; or any combination of these. The linker moiety can be separate from or part of a PAP variant. An exemplary His-tag for use with PAP variants of the invention is MASSHHHHHHSSGSENLYFQTGSSG- (SEQ ID NO: 5)). The tag-linker moiety does not interfere with the nucleotide binding activity, or catalytic activity of the PAP variant.
- [0047] The above processes, or equivalent processes, result in an isolated PAP or PUP variant that may be mixed with a variety of reagents, such as, salts, pH buffers, carrier compounds, and the like, that are necessary or useful for activity and/or preservation.

Measurement of Nucleotide Incorporation Activity

- [0048] The efficiency of nucleotide incorporation by variants of the invention may be measured by an extension, or elongation, assay, e.g. as described in Boule et al (cited below);

Bentolila et al (cited below); and Hiatt et al, U.S. patent 5808045, the latter of which is incorporated herein by reference. Briefly, in one form of such an assay, a fluorescently labeled oligonucleotide having a free 3'-hydroxyl is reacted under PAP extension conditions with a PAP variant to be tested for a predetermined duration in the presence of a reversibly blocked ribonucleoside triphosphate, after which the extension reaction is stopped and the amounts of extension products and unextended initiator oligonucleotide are quantified after separation by gel electrophoresis. By such assays, the incorporation efficiency of a PAP variant may be readily compared to the efficiencies of other variants or to that of wild type or reference PAPs, or other polymerases. In some embodiments, a measure of PAP variant efficiency may be a ratio (given as a percentage) of amount of extended product using the PAP variant over the amount of extended product using wild type PAP in an equivalent assay. The reagents are added in the tube, starting from water, and then in the order of Table 3.

Table 3

Extension Activity Assay Reagents

Reagent	Concentration	Volume
H ₂ O	-	12 μ L
Activity buffer	10x	2 μ L
rNTP	250 μ M	2 μ L
Purified enzyme	20 μ M	2 μ L
Fluorescent primer	500 nM	2 μ L

[0049] The product of the assay may be analyzed by conventional polyacrylamide gel electrophoresis. For example, products of the above assay may be analyzed in a 16 percent polyacrylamide denaturing gel (Bio-Rad). Gels are made just before the analysis by pouring polyacrylamide inside glass plates and let it polymerize. The gel inside the glass plates is mounted on an adapted tank filed with TBE buffer (Sigma) for the electrophoresis step. The samples to be analyzed are loaded on the top of the gel. A voltage of 500 to 2,000V is applied between the top and bottom of the gel for 3 to 6h at room temperature. After separation, gel fluorescence is scanned using, for example, a Typhoon scanner (GE Life Sciences). The gel image is analyzed using ImageJ software (imagej.nih.gov/ij/), or its equivalent, to calculate the percentage of incorporation of the modified nucleotides.

Kits

[0050] The invention includes kits for carrying out methods of the invention. In some
5 embodiments, a kit of the invention comprises a poly(A) polymerase and 3'-O-blocked
ribonucleoside triphosphates of one or more of adenosine, guanosine, uridine and cytidine. In
other embodiments, such kit may further include a poly(U) polymerase. In some
embodiments, a kit of the invention comprises a poly(A) polymerase and 3'-O-blocked-2'-
deoxyribonucleoside triphosphates of one or more of deoxyadenosine, deoxyguanosine,
10 deoxythymidine and deoxycytidine. In other embodiments, such kit may further include a
poly(U) polymerase. In some embodiments, kits of the invention may include an initiator with
a free 3'-hydroxyl attached to a support by a 5' end. In some embodiments, such support is a
solid support. Such solid support may comprise beads, such as magnetic beads or agarose
beads, a planar solid, such as a glass slide, or a membrane, or the like. In some embodiments,
15 a kit of the invention may further include a de-blocking agent which is capable of removing 3'
blocking groups from incorporated 3'-O-blocked nucleotides. In some embodiments, a kit
may include an M310 mutant of a yeast PAP, such as, the yeast PAP of SEQ ID NO: 1, or a
functionally equivalent residue to M310 in a PAP from a different species. In some
embodiments, a kit may include a PAP variant comprising an amino acid sequence that is at
20 least 90 percent identical to an amino acid sequence selected from SEQ ID NO: 1, 2 or 3 with
a substitution of methionine at position 310 with respect to SEQ ID NO: 1, or methionine at
position 318 with respect to SEQ ID NOs 2 and 3, wherein the variant capable of (a)
synthesizing a ribonucleic acid fragment without a template and (b) incorporating a 3'-O-
azidomethyl-ribonucleoside triphosphate onto a ribonucleic acid fragment or a 3'-O-
25 azidomethyl-2'-deoxyribonucleoside triphosphate into a deoxyribonucleic acid fragment.

[0051] In some embodiments, in PAP variants of the foregoing kits the substitution of the
methionine at position 310 of SEQ ID NO: 1 or the methionine at position 318 of SEQ ID NOs
2 or 3 is selected from M310F/Y/V/E/T. In some embodiments, in PAP variants of the
foregoing kits there is a substitution of valine at position 234 with respect to SEQ ID NO: 1
30 and valine at position 240 with respect to SEQ ID NOs 2 and 3. In some embodiments, in
PAP variants of the foregoing kits, the substitution of valine at position 234 of SEQ ID NO: 1

and position 240 with respect to SEQ ID NOs 2 and 3 is alanine. In some embodiments, in PAP variants of the foregoing kits, there is a substitution of alanine at position 410 with respect to SEQ ID NOs 2 or 3.

[0052] In some embodiments, a kit of the invention may comprise a PAP variant having an amino acid sequence at least 80 percent identical to a sequence selected from SEQ ID NOs 1-3 and 8-27 of Table 1 which has a substitution in the first position or a substitution in the second position or a substitution in both positions of such selected sequence as indicated in Table 1. In some embodiments, such percent identity of the selected sequence is at least 90 percent identity. In some embodiments, a kit of the invention may comprise a PUP variant having an amino acid sequence at least 80 percent identical to a sequence selected from SEQ ID NOs 4 and 28-47 of Table 2 which has a substitution in the first position or a substitution in the second position or a substitution in both positions of such selected sequence as indicated in Table 2. In some embodiments, such percent identity of the selected sequence is at least 90 percent identity.

Example 1

One-to-Five Ribonucleotide Additions to Initiator Polynucleotides

Immobilized on a Solid Support

[0053] 2-3 cycle extension of Cp-RNA primer with 3'-O-azidomethyl-ribonucleoside triphosphates (AM-NTPs) in presence of PAP (Thermo). Experimental procedure: 1) preparation of Cp-beads: 100 uL Cp-beads (~1 uM concentration of Cp primer on bead) + 1 uL 100 uM SynRDA_DNA-SynRDA_RNA; incubate for 30 min at room temp, wash beads with 3x200 uL binding buffer (BB), (200 mM cac, LiCl), resuspend beads in 25 uL BB; 2) +1 addition reaction: 250 uM 3'-O-terminated nucleotide; 3.2 uM Cp-SynRDA_DNA-SynRDA_RNA; 1x PAP_yeast Thermo reaction buffer; 100 U/uL PAP_yeast (Thermo); 20 uL final volume; in Eppendorf tubes; 30 min, 37°C, thermomixer, 1500 rpm, 3) wash +1 product: 3X 200 uL BB 4) TCEP deprotection: resuspend beads in 50 uL 200 uM TCEP, pH 7.0 (dilution from freshly opened ampula), 37C, 15 min, thermomixer, 1500 rpm; 5) wash deprotected +1 product: 3x 200 uL binding buffer, 6) +2/deprotect/+3 addition: see 1)-5), 8) Elution + gel: add 20 uL B-blue, vortex, elute, apply on gel. The Cp-RNA beads used were

prepared the following way: i) covalent functionalization of commercial Dynabeads M-270 that display a terminal carboxylic acid group (purchased from Thermo Fisher) with a DNA oligo (TCATTTCACTCTCACA-NH₂)(SEQ ID NO: 6). These beads are named Cp-beads. Cp-beads were then incubated with a DNA-RNA hybrid primer of sequence ATTO488-
 5 TGTGAGAGTGAAATGAGGrUrGrUrGrArGrArGrUr-GrArArArUrGrArGrG (named SynRDA_DNA-SynRDA_RNA) (SEQ ID NO: 7)

[0054] Results are shown in Fig. 2A-2C. This data show: i) with 3'-O-AM-rATP and 3'-O-AM-rUTP at least 5 cycles of addition/deprotection can be achieved with same nucleotide; ii) at least 2 cycles of addition/deprotection can be achieved with two different 3'-O-AM-
 10 rNTPs (A and U); and iii) at least >95% deprotection of 3'-O-AM-protected RNA can be achieved.

15

Example 2

Effect of PAP M310 Mutations on Activities AM-rATP

Elongation and Riboguanylation

[0055] In this experiment, the effects of mutations in PAP position M310 on activities of AM-rATP elongation of an RNA primer (Fig. 3A) and riboguanylation of a RNA primer (Fig.
 20 3B) are examined. AM-rATP elongation of an RNA primer (Figure 3A) was carried out in presence of 0.1 uM ATTO488-(rA)₅ primer, 0.5 uM PAP mutant, 250 uM AM-rATP in elongation AM-rATP buffer (20 mM Tris-HCl, 0.6 mM MnCl₂, 0.02 mM EDTA, 0.1% BSA, 10% glycerol, 100 mM imidazole, pH 7-8) for 30 min. at 37C in a total reaction volume of 20 uL. Riboguanylation of an RNA primer (Figure 3B) was carried out in presence of 0.1 uM
 25 ATTO488-(rA)₅ primer, 0.5 uM PAP mutant, 250 uM AM-rATP in rGTP polymerization buffer (0.6 mM MnCl₂, 0.1% BSA, 10 mM imidazole, pH 6.2) for 30 min. at 37C in a total reaction volume of 20 uL.

30

Example 3

Effect of Thielavia PAP Variant on AM-rNTP Incorporation

[0056] Thielavia PAP (SEQ ID NO: 3) was engineered to have three substitutions (V240A, M318T and A410V) using conventional techniques, after which the modified polypeptide was expressed and purified. The ability of wild type Thielavia PAP and the mutated PAP to incorporate AM-NTPs was compared by performing the assay described in Example 2 except for the following conditions: 0.1 μ M Atto488-(rA)₇, 3.0 μ M PAP, 250 μ M AM-NTP. The reaction results are given in Table 4 below.

Table 4

PAP	Melting temp (degrees C)	Purity +1 product (%)AM-ATP	Purity +1 product (%)AM-UTP	Purity +1 product (%)AM-GTP	Purity +1 product (%)AM-CTP
Thelavia variant	46	100	100	97	95
Thelavia wild type	48	100	100	23	31

The data shows that the Thelavia PAP variant incorporates AM-rGTP and AM-rCTP monomers much more efficiently than the corresponding wild type PAP.

Example 4

Incorporation of AM-NTPs by PUP

[0057] In this experiment, schizosaccharomyces pombe PUP (SEQ ID NO: 4) was used to incorporate AM-NTPs into an initiator using the experimental protocol of Example 3.

The extension products were separated by gel electrophoresis to give the electropherogram shown in Fig. 4.

Definitions

[0058] Amino acids are represented by either 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).

[0059] “Functionally equivalent” in reference to a substituted residue means the

5 substituted residue of a variant PAP has an identical functional role as a residue in a sequence of another PAP having a sequence homologous to SEQ ID NO: 1. Functionally equivalent residues may be identified by using sequence alignments, for example, using the Mutalin line alignment software (<http://multalin.toulouse.inra.fr/multalin/multalin.html>; 1988, Nucl. Acids Res., 16 (22), 25 10881-10890). After alignment, the functionally equivalent residues are at
10 homologous positions on the different sequences considered. Sequence alignments and identification of functionally equivalent residues may be determined between any PAP and their natural variants, including inter-species.

[0060] “Isolated” in reference to protein means such a compound which has been

identified and separated and/or recovered from a component of its natural environment or from
15 a heterogeneous reaction mixture. Contaminant components of a natural environment or reaction mixture are materials which would interfere with a protein’s function, and may include enzymes, hormones, and other proteinaceous or nonproteinaceous solutes. In some embodiments, a protein of the invention is purified (1) to greater than 95% by weight of protein as determined by the Lowry method, and most preferably more than 99% by weight,
20 (2) to a degree sufficient to obtain at least 15 residues of N-terminal or internal amino acid sequence by use of a spinning cup sequenator, or (3) to homogeneity by SDS-PAGE under reducing or nonreducing conditions using Coomassie blue or, preferably, silver stain. When manufactured by recombinant methodologies, an isolated protein of the invention may include the protein of the invention in situ within recombinant cells since at least one component of the
25 protein’s natural environment will not be present. Ordinarily, an isolated protein of the invention is prepared by at least one purification step.

[0061] “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
30 the storage, transport, or delivery of reaction reagents (e.g., PAP enzymes, protected rNTP monomers, 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, while a second or more
5 containers contain rNTP monomers, buffers, solid supports with initiators, or the like.

[0062] “Mutant” or “variant,” which are used interchangeably, refer to polypeptides derived from SEQ ID NO: 1 and comprising a modification or an alteration, i.e., a substitution, insertion, and/or deletion, at one or more positions and having both a template-free polymerase activity and ability to incorporate one or more reversibly blocked nucleoside triphosphate
10 precursors. The 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 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
15 invention of a polymerase. The following terminology is used to designate a substitution: L238A denotes that amino acid residue (Leucine, L) at position 238 of a reference, or wild type, 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
20 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
25 and serine).

[0063] “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
30 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. Corresponding ribonucleotides may be designated "rA", "rC", "rG", and "rT." 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.

- [0064] “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).
- [0065] “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.

[0066] “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).

[0067] 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
5 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.