



(51) International Patent Classification:

C12P 19/34 (2006.01) *C12N 9/12* (2006.01)
C12Q 1/68 (2006.01)

(21) International Application Number:

PCT/GB2016/050301

(22) International Filing Date:

10 February 2016 (10.02.2016)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

1502152.0 10 February 2015 (10.02.2015) GB

(71) Applicant: **NUCLERA NUCLEICS LTD** [GB/GB];
Stevenage Bioscience Catalyst, Gunnels Wood Road,
Stevenage, Hertfordshire SG1 2FX (GB).

(72) Inventors: **CHEN, Michael C**; c/o Nuclera Nucleics Ltd,
57 Cavendish Road, Cambridge Cambridgeshire CB1 3AE
(GB). **LAZAR, Radu A**; c/o Nuclera Nucleics Ltd, 57
Cavendish Road, Cambridge Cambridgeshire CB1 3AE
(GB). **HUANG, Jiahao**; c/o Nuclera Nucleics Ltd, 57 Cav-
endish Road, Cambridge Cambridgeshire CB1 3AE (GB).
MCINROY, Gordon R; c/o Nuclera Nucleics Ltd, 57

Cavendish Road, Cambridge Cambridgeshire CB1 3AE
(GB).

(74) Agents: **GIBSON, Mark** et al.; Sagittarius IP, Three
Globeside, Fieldhouse Lane, Marlow Buckinghamshire
SL7 1HZ (GB).

(81) Designated States (*unless otherwise indicated, for every
kind of national protection available*): AE, AG, AL, AM,
AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY,
BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM,
DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT,
HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR,
KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG,
MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM,
PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC,
SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN,
TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States (*unless otherwise indicated, for every
kind of regional protection available*): ARIPO (BW, GH,
GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ,
TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU,
TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE,
DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU,
LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK,

[Continued on next page]

(54) Title: NOVEL USE

(57) Abstract: The invention relates to the use of specific
terminal deoxynucleotidyl transferase (TdT) enzymes in a
method of nucleic acid synthesis, to methods of synthesizing
nucleic acids, and to the use of kits comprising said enzymes
in a method of nucleic acid synthesis. The invention also
relates to the use of terminal deoxynucleotidyl transferases
and 3'-blocked nucleotide triphosphates in a method of tem-
plate independent nucleic acid synthesis.

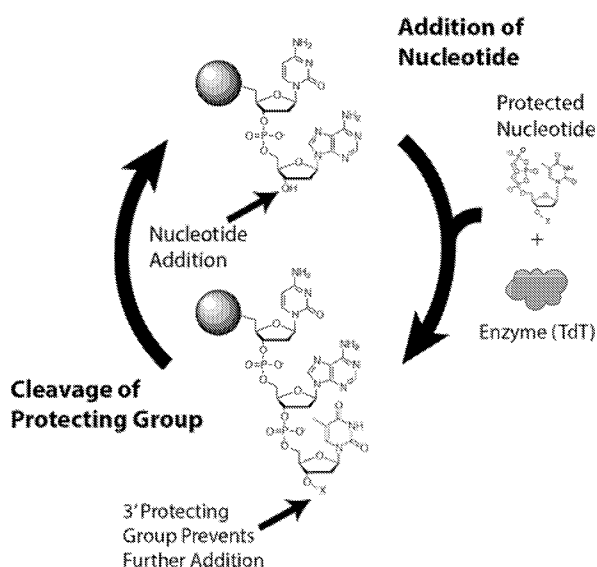


FIGURE 1



SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ,
GW, KM, ML, MR, NE, SN, TD, TG).

— *before the expiration of the time limit for amending the
claims and to be republished in the event of receipt of
amendments (Rule 48.2(h))*

Published:

— *with international search report (Art. 21(3))*

— *with sequence listing part of description (Rule 5.2(a))*

NOVEL USE

FIELD OF THE INVENTION

The invention relates to the use of specific terminal deoxynucleotidyl transferase (TdT) enzymes in a method of nucleic acid synthesis, to methods of synthesizing nucleic acids, and to the use of kits comprising said enzymes in a method of nucleic acid synthesis. The invention also relates to the use of terminal deoxynucleotidyl transferases and 3'-blocked nucleotide triphosphates in a method of template independent nucleic acid synthesis.

BACKGROUND OF THE INVENTION

Nucleic acid synthesis is vital to modern biotechnology. The rapid pace of development in the biotechnology arena has been made possible by the scientific community's ability to artificially synthesise DNA, RNA and proteins.

Artificial DNA synthesis – a £1 billion and growing market – allows biotechnology and pharmaceutical companies to develop a range of peptide therapeutics, such as insulin for the treatment of diabetes. It allows researchers to characterise cellular proteins to develop new small molecule therapies for the treatment of diseases our aging population faces today, such as heart disease and cancer. It even paves the way forward to creating life, as the Venter Institute demonstrated in 2010 when they placed an artificially synthesised genome into a bacterial cell.

However, current DNA synthesis technology does not meet the demands of the biotechnology industry. While the benefits of DNA synthesis are numerous, an oft-mentioned problem prevents the further growth of the artificial DNA synthesis industry, and thus the biotechnology field. Despite being a mature technology, it is practically impossible to synthesise a DNA strand greater than 200 nucleotides in length, and most DNA synthesis companies only offer up to 120 nucleotides. In comparison, an average protein-coding gene is of the order of 2000 – 3000 nucleotides, and an average eukaryotic genome numbers in the billions of nucleotides. Thus, all major gene synthesis companies today rely on variations of a 'synthesise and stitch' technique, where overlapping 40-60-mer fragments are synthesised and stitched together by PCR (see Young, L. et al. (2004) *Nucleic Acid Res.* 32, e59). Current methods offered by the gene synthesis industry generally allow up to 3 kb in length for routine production.

The reason DNA cannot be synthesised beyond 120-200 nucleotides at a time is due to the current methodology for generating DNA, which uses synthetic chemistry (i.e.,

phosphoramidite technology) to couple a nucleotide one at a time to make DNA. As the efficiency of each nucleotide-coupling step is 95.0 - 99.5% efficient, it is mathematically impossible to synthesise DNA longer than 200 nucleotides in acceptable yields. The Venter Institute illustrated this laborious process by spending 4 years and 20 million USD to synthesise the relatively small genome of a bacterium (see Gibson, D. G. et al. (2010) *Science* 329, 52–56).

Known methods of DNA sequencing use template-dependent DNA polymerases to add 3'-reversibly terminated nucleotides to a growing double-stranded substrate (see, Bentley, D. R. et al. (2008) *Nature* 456, 53–59). In the 'sequencing-by-synthesis' process, each added nucleotide contains a dye, allowing the user to identify the exact sequence of the template strand. Albeit on double-stranded DNA, this technology is able to produce strands of between 500-1000 bps long. However, this technology is not suitable for *de novo* nucleic acid synthesis because of the requirement for an existing nucleic acid strand to act as a template.

Various attempts have been made to use a terminal deoxynucleotidyl transferase for controlled *de novo* single-stranded DNA synthesis (see Ud-Dean, S. M. M. (2009) *Syst Synth Boil* 2, 67–73, US 5,763,594 and US 8,808,989). Uncontrolled *de novo* single-stranded DNA synthesis, as opposed to controlled, takes advantage of TdT's deoxynucleotide triphosphate (dNTP) 3' tailing properties on single-stranded DNA to create, for example, homopolymeric adaptor sequences for next-generation sequencing library preparation (see Roychoudhury R., et al. (1976) *Nucleic Acids Res* 3, 101–116 and WO 2003/050242). A reversible deoxynucleotide triphosphate termination technology needs to be employed to prevent uncontrolled addition of dNTPs to the 3'-end of a growing DNA strand. The development of a controlled single-stranded DNA synthesis process through TdT would be invaluable to *in situ* DNA synthesis for gene assembly or hybridization microarrays as it removes the need for an anhydrous environment and allows the use of various polymers incompatible with organic solvents (see Blanchard, A. P. (1996) *Biosens Bioelectron* 11, 687–690 and US 7,534,561).

However, TdT has not been shown to efficiently add nucleotide triphosphates containing 3'-O reversibly terminating moieties for building up a nascent single-stranded DNA chain necessary for a *de novo* synthesis cycle. A 3'-O reversibly terminating moiety would prevent a terminal transferase like TdT from catalysing the nucleotide transferase reaction between the 3'-end of a growing DNA strand and the 5'-triphosphate of an incoming nucleotide triphosphate. Data is presented herein which demonstrates that the widely commercially

available recombinant TdT sourced from calf thymus is unable to add 3'-O-terminated nucleotide triphosphates in a quantitative fashion (see Figure 3). In previous reports, the TdT specifically mentioned is recombinant TdT from calf thymus (see Ud-Dean, S. M. M. (2009) *Syst Synth Boll* 2, 67-73, US 5,763,594 and US 8,808,989) or uses a different
5 reversible terminating mechanism not located on the 3' end of the deoxyribose moiety (see US 8,808,989).

Most DNA and RNA polymerases contain highly selective sugar steric gates to tightly discriminate between deoxyribose and ribose nucleotide triphosphate substrates (see Joyce
10 C. M. (1997) *Proc Natl Acad Sci* 94, 1619-22). The result of this sugar steric gate is the enormous challenge of finding and/or engineering polymerases to accept sugar variants for biotechnology reasons, such as sequencing-by-synthesis (see Metzker M. L. (2010) *Nat Rev Genet* 11, 31-46 and US 8,460,910). The challenge of finding a polymerase that accepts a 3'-O reversibly terminating nucleotide is so large, various efforts have been made to create
15 reversible terminating nucleotides where the polymerase termination mechanism is located on the nitrogenous base of the terminating nucleotide (see Gardner, A. F. (2012) *Nucleic Acids Res* 40, 7404-15 and US 8,889,860).

There is therefore a need to identify terminal deoxynucleotidyl transferases that readily
20 incorporate 3'-O reversibly terminated nucleotides and modified said terminal deoxynucleotidyl transferases to incorporate 3'-O reversibly terminated nucleotides in a fashion useful for biotechnology and single-stranded DNA synthesis processes in order to provide an improved method of nucleic acid synthesis that is able to overcome the problems associated with currently available methods.

SUMMARY OF THE INVENTION

According to a first aspect of the invention, there is provided the use of a terminal deoxynucleotidyl transferase (TdT) enzyme comprising an amino acid sequence selected from either: (a) any one of SEQ ID NOS: 1 to 5 and 8 or a functional equivalent or fragment
30 thereof having at least 20% sequence homology to said amino acid sequence; or (b) a modified derivative of SEQ ID NO: 6; in a method of nucleic acid synthesis.

According to a second aspect of the invention, there is provided a method of nucleic acid
35 synthesis, which comprises the steps of:

- (a) providing an initiator sequence;

(b) adding a 3'-blocked nucleotide triphosphate to said initiator sequence in the presence of a terminal deoxynucleotidyl transferase (TdT) as defined in the first aspect of the invention;

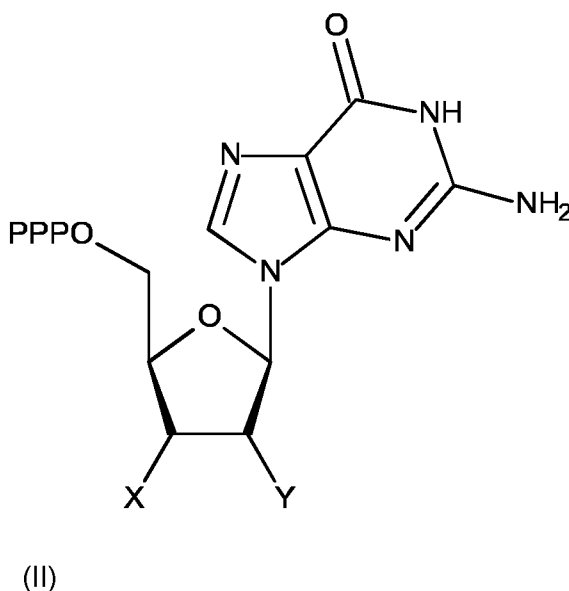
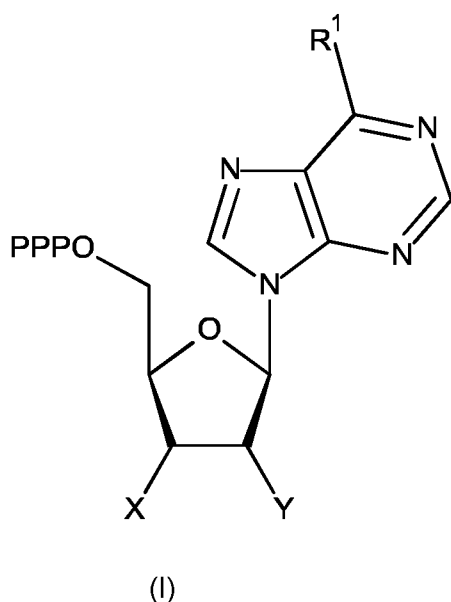
(c) removal of all reagents from the initiator sequence;

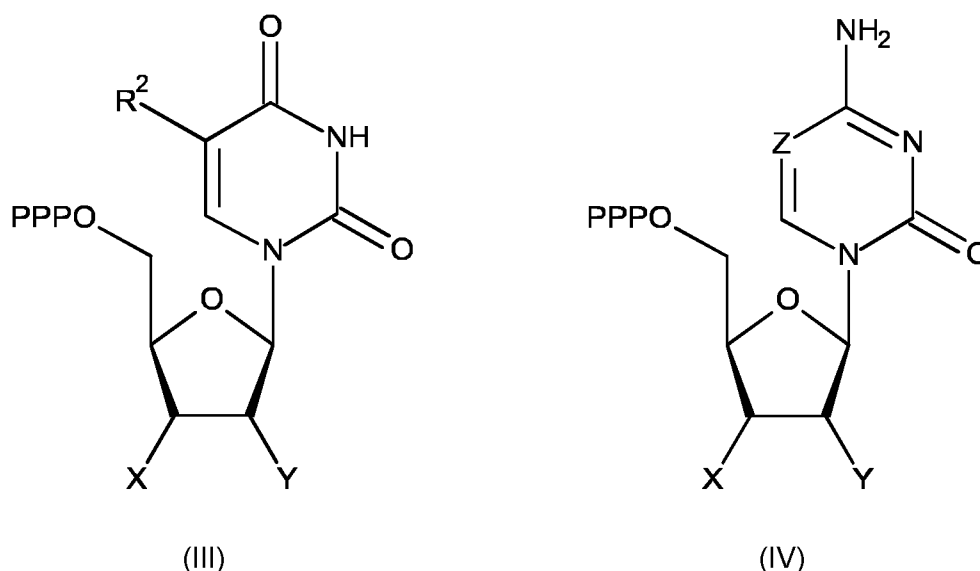
5 (d) cleaving the blocking group from the 3'-blocked nucleotide triphosphate in the presence of a cleaving agent;

(e) removal of the cleaving agent.

10 According to a further aspect of the invention, there is provided the use of a kit in a method of nucleic acid synthesis, wherein said kit comprises a TdT as defined in the first or second aspects of the invention optionally in combination with one or more components selected from: an initiator sequence, one or more 3'-blocked nucleotide triphosphates, inorganic pyrophosphatase, such as purified, recombinant inorganic pyrophosphatase from *Saccharomyces cerevisiae*, and a cleaving agent; further optionally together with instructions
15 for use of the kit in accordance with any of the methods defined herein.

20 According to a further aspect of the invention, there is provided the use of a 3'-blocked nucleotide triphosphate in a method of template independent nucleic acid synthesis, wherein the 3'-blocked nucleotide triphosphate is selected from a compound of formula (I), (II), (III) or (IV):





wherein

R^1 represents NR^aR^b , wherein R^a and R^b independently represent hydrogen or C_{1-6} alkyl,

5 R^2 represents hydrogen, C_{1-6} alkyl, C_{1-6} alkoxy, COH or COOH.

X represents C_{1-6} alkyl, NH_2 , N_3 or $-OR^3$,

R^3 represents C_{1-6} alkyl, CH_2N_3 , NH_2 or allyl,

Y represents hydrogen, halogen or hydroxyl, and

Z represents CR^4 or N, wherein R^4 represents hydrogen, C_{1-6} alkyl, C_{1-6} alkoxy, COH or COOH.

10

According to a further aspect of the invention, there is provided the use of inorganic pyrophosphatase in a method of nucleic acid synthesis.

BRIEF DESCRIPTION OF THE FIGURES

15 **Figure 1: Schematic of enzymatic DNA synthesis process.** Starting from the top of the diagram, an immobilised strand of DNA with a deprotected 3'-end is exposed to an extension mixture composed of TdT, a base-specific 3'-blocked nucleotide triphosphate, inorganic pyrophosphatase to reduce the buildup of inorganic pyrophosphate, and appropriate buffers/salts for optimal enzyme activity and stability. The protein adds one protected

20 nucleotide to the immobilised DNA strand (bottom of diagram). The extension mixture is then removed with wash mixture and optionally recycled. The immobilised (n+1) DNA strand is then washed with a cleavage mixture to cleave the 3'-protecting group, enabling reaction in the next cycle. In the cleavage mixture, denaturant may be present to disrupt any secondary structures. During this step, the temperature may be raised to assist in cleavage and disruption

25 of secondary structures. The immobilised DNA is treated with wash mixture to remove leftover

cleavage mixture. Steps 1-4 may be repeated with an appropriate nucleotide triphosphate until the desired oligonucleotide sequence is achieved.

Figure 2: TdT adds a 3'-irreversibly blocked nucleotide triphosphate to an initiator strand. (A) A single-stranded DNA initiator was incubated with 15 U *Bos taurus* TdT, required salts (50 mM potassium acetate, 20 mM tris acetate pH 7.9, 1 mM cobalt chloride), and 3'-O-Methyl dTTP at 37 °C for the indicated amount of time. The 3'-irreversibly blocked nucleotide triphosphate was at a concentration of 1 mM and the DNA initiator at 200 pM for a 1:5000 ratio to encourage nucleotide addition. The reaction was stopped with EDTA (0.5 M). (B) Similar to (A) with the exception that 3'-Azido dTTP was used as the 3'-blocked nucleotide triphosphate.

Figure 3: Various TdT orthologs add a 3'-reversibly protected (3'-O-azidomethyl) nucleotide triphosphate to an initiator strand up to 3.8-fold faster than *Bos taurus* TdT. (A) A single-stranded DNA initiator was incubated with TdT with the indicated species for 60 min in the above mentioned buffer and reaction conditions. Reactions were then analysed by capillary electrophoresis and a subset of chromatograms are shown. (B) Fraction of DNA initiator strand converted to N+1 species (addition of 3'-O-azidomethyl dTTP) in one hour. (C) Three TdT orthologs were reacted with 3'-O-azidomethyl dTTP and 3'-O-azidomethyl dCTP for 20 min. *Lepisosteus oculatus* and *Sarcophilus harrisii* TdT consistently perform better compared to *Bos taurus* TdT. The ability for TdT orthologs to add both dTTP and dCTP nucleotide analogs demonstrate control of DNA sequence specificity.

Figure 4: An engineered variant of *Lepisosteus oculatus* TdT shows improved activity over the wild-type *Lepisosteus oculatus* TdT. A DNA initiator strand was incubated with 1 mM 3'-reversibly blocked dNTP at 37°C for 20 minutes with required salts as described previously. Addition of the dNTP was measured by PAGE and the conversion relative to wild-type TdT was plotted.

Figure 5: An engineered variant of *Lepisosteus oculatus* shows vastly improved activity over the wild-type TdT (see Figure 3) and commercially available *Bos taurus* TdT. TdT was incubated with a DNA initiator and a 3'-reversibly blocked dNTP in a similar fashion as previously described. *Bos taurus* TdT was outperformed by an engineered variant of *Lepisosteus oculatus* TdT, as demonstrated in the PAGE gel.

Figure 6: Inorganic pyrophosphate is necessary for TdT-mediated, *de novo* sequence specific DNA synthesis. A nucleic acid initiator was incubated with and without *Saccharomyces cerevisiae* inorganic pyrophosphatase and *Bos taurus* TdT at 37 °C in 50 mM potassium acetate, 20 mM tris acetate pH 7.9, 1.5 mM cobalt chloride. 3'-O-azidomethyl dTTP (N₃Me-dTTP) was introduced at 1 mM and dideoxyTTP (ddTTP) was introduced at 100 µM. The reactions were analysed by PAGE. Without inorganic pyrophosphatase, strand
5 dismutation predominates and catastrophic loss of sequence-specificity occurs.

Figure 7: Simplified schematic representation of a column-based flow instrument used in DNA synthesis. A computer (302) controls two pumps and a solution
10 mixing chamber (311). Pump 1 (304) selectively pumps extension solution (301), wash solution (305) or cleavage solution (310) into the mixing chamber. Pump 2 (306) selectively pumps a single 3'-blocked nucleotide triphosphate (TP) solution containing either 3'-blocked A(adenine)TP (303), T(thymine)TP (307), G(guanine)TP (308), or C(cytosine)TP (309) into
15 the chamber. The computer controlled mixing chamber then passes appropriate solution ratios from pump 1 and pump 2 into a column based DNA synthesis chamber (312). A heating element (313) ensures that the DNA synthesis column remains at the necessary temperature for the synthesis process to take place. Upon exiting the DNA synthesis chamber, the reaction solution either enters a recycling vessel (314) for future use, a waste vessel (316) or moves
20 on to a polymerase chain reaction (PCR) step (315) for amplification of the resultant DNA. PCR completion leads to the final product (317).

DETAILED DESCRIPTION OF THE INVENTION

According to a first aspect of the invention, there is provided the use of a terminal
25 deoxynucleotidyl transferase (TdT) enzyme comprising an amino acid sequence selected from either: (a) any one of SEQ ID NOS: 1 to 5 and 8 or a functional equivalent or fragment thereof having at least 20% sequence homology to said amino acid sequence; or (b) a modified derivative of SEQ ID NO: 6; in a method of nucleic acid synthesis.

30 According to one particular aspect of the invention which may be mentioned, there is provided the use of a terminal deoxynucleotidyl transferase (TdT) enzyme comprising an amino acid sequence selected from either: (a) any one of SEQ ID NOS: 1 to 5 or a functional equivalent or fragment thereof having at least 20% sequence homology to said amino acid sequence; or (b) a modified derivative of SEQ ID NO: 6; in a method of nucleic acid
35 synthesis.

The present invention relates to the identification of never before studied terminal deoxynucleotidyl transferases that surprisingly have the ability to incorporate deoxynucleotide triphosphates with large 3'-O reversibly terminating moieties.

5 Since commercially available recombinant TdT sourced from calf thymus does not readily incorporate 3'-O reversibly terminated nucleotides, it is most unexpected that the present inventors have located a terminal deoxynucleotidyl transferase, which is a DNA polymerase, to accept a 3'-O reversibly terminating nucleotide, such as dNTPs modified with a 3'-O-azidomethyl.

10 Furthermore, the present invention relates to engineered terminal deoxynucleotidyl transferases, which achieve a substantial increase in incorporation rates of dNTPs containing 3'-O reversibly terminating moieties to be useful for the controlled *de novo* synthesis of single-stranded DNA.

15 As controlled *de novo* single-stranded DNA synthesis is an additive process, coupling efficiency is extremely important to obtaining practically useful yields of final single-stranded DNA product for use in applications such as gene assembly or hybridization microarrays. Thus, the present invention relates to the identification of TdT orthologs with the capability to
20 add 3'-O reversibly terminated nucleotides, and also an engineered variant of the TdT ortholog that adds 3'-O reversibly terminated nucleotides in a quantitative fashion that is practically useful for a single-stranded DNA synthesis process.

The use described herein has significant advantages, such as the ability to rapidly produce
25 long lengths of DNA while still maintaining high yields and without using any toxic organic solvents.

In one embodiment, the terminal deoxynucleotidyl transferase (TdT) enzyme comprises an amino acid sequence selected from any one of SEQ ID NOS: 1 to 5 and 8 or a functional
30 equivalent or fragment thereof having at least 20% sequence homology to said amino acid sequence.

In a further embodiment, the terminal deoxynucleotidyl transferase (TdT) enzyme comprises an amino acid sequence selected from any one of SEQ ID NOS: 1 to 5 or a functional
35 equivalent or fragment thereof having at least 20% sequence homology to said amino acid sequence.

In a further embodiment, the terminal deoxynucleotidyl transferase (TdT) enzyme comprises an amino acid sequence selected from SEQ ID NO: 1. The amino acid sequence of SEQ ID NO: 1 is the terminal deoxynucleotidyl transferase (TdT) sequence from *Sarcophilus harrisii* (UniProt: G3VQ55). *Sarcophilus harrisii* (also known as the Tasmanian devil) is a carnivorous marsupial of the family Dasyuridae, now found in the wild only on the Australian island state of Tasmania.

In a further embodiment, the terminal deoxynucleotidyl transferase (TdT) enzyme comprises an amino acid sequence selected from SEQ ID NO: 2. The amino acid sequence of SEQ ID NO: 2 is the terminal deoxynucleotidyl transferase (TdT) sequence from *Lepisosteus oculatus* (UniProt: W5MK82). *Lepisosteus oculatus* (also known as the spotted gar) is a primitive freshwater fish of the family Lepisosteidae, native to North America from the Lake Erie and southern Lake Michigan drainages south through the Mississippi River basin to Gulf Slope drainages, from lower Apalachicola River in Florida to Nueces River in Texas, USA.

In a further embodiment, the terminal deoxynucleotidyl transferase (TdT) enzyme comprises an amino acid sequence selected from SEQ ID NO: 3. The amino acid sequence of SEQ ID NO: 3 is the terminal deoxynucleotidyl transferase (TdT) sequence from *Chinchilla lanigera* (NCBI Reference Sequence: XP_005407631.1; <http://www.ncbi.nlm.nih.gov/protein/533189443>). *Chinchilla lanigera* (also known as the long-tailed chinchilla, Chilean, coastal, common chinchilla, or lesser chinchilla), is one of two species of rodents from the genus *Chinchilla*, the other species being *Chinchilla chinchilla*.

In a further embodiment, the terminal deoxynucleotidyl transferase (TdT) enzyme comprises an amino acid sequence selected from SEQ ID NO: 4. The amino acid sequence of SEQ ID NO: 4 is the terminal deoxynucleotidyl transferase (TdT) sequence from *Otolemur garnettii* (UniProt: A4PCE6). *Otolemur garnettii* (also known as the northern greater galago, Garnett's greater galago or small-eared greater galago), is a nocturnal, arboreal primate endemic to Africa.

In a further embodiment, the terminal deoxynucleotidyl transferase (TdT) enzyme comprises an amino acid sequence selected from SEQ ID NO: 5. The amino acid sequence of SEQ ID NO: 5 is the terminal deoxynucleotidyl transferase (TdT) sequence from *Sus scrofa* (UniProt: F1SBG2). *Sus scrofa* (also known as the wild boar, wild swine or Eurasian wild pig) is a suid native to much of Eurasia, North Africa and the Greater Sunda Islands.

In a further embodiment, the terminal deoxynucleotidyl transferase (TdT) enzyme comprises an amino acid sequence selected from SEQ ID NO: 8. The amino acid sequence of SEQ ID NO: 8 is a variant of SEQ ID NO: 2 which has been engineered for improved activity by alteration of the amino acid sequence. Data are provided in Example 3 and Figure 4, which demonstrate the benefits of engineered variants, such as SEQ ID NO: 8, over the wild-type SEQ ID NO: 2.

In a further embodiment, the terminal deoxynucleotidyl transferase (TdT) enzyme comprises an amino acid sequence selected from SEQ ID NOS: 1, 2 or 8.

In a further embodiment, the terminal deoxynucleotidyl transferase (TdT) enzyme comprises an amino acid sequence selected from SEQ ID NO: 1 or 2. Data are provided in Example 2 and Figure 3, which demonstrate beneficial results over the natural, recombinant TdT enzyme from *Bos taurus*.

In an alternative embodiment, the terminal deoxynucleotidyl transferase (TdT) enzyme comprises an amino acid sequence selected from a modified derivative of SEQ ID NO: 6 (i.e. a non-natural, mutated derivative of SEQ ID NO: 6). The amino acid sequence of SEQ ID NO: 6 is the terminal deoxynucleotidyl transferase (TdT) sequence from *Bos taurus* (UniProt: P06526). *Bos taurus* (also known as cattle, or colloquially cows) are the most common type of large domesticated ungulates. They are a prominent modern member of the subfamily Bovinae, are the most widespread species of the genus *Bos*.

References herein to “TdT” refer to a terminal deoxynucleotidyl transferase (TdT) enzyme and include references to purified and recombinant forms of said enzyme. TdT is also commonly known as DNTT (DNA nucleotidylexotransferase) and any such terms should be used interchangeably.

References herein to a “method of nucleic acid synthesis” include methods of synthesising lengths of DNA (deoxyribonucleic acid) or RNA (ribonucleic acid) wherein a strand of nucleic acid (n) is extended by adding a further nucleotide (n+1). In one embodiment, the nucleic acid is DNA. In an alternative embodiment, the nucleic acid is RNA.

References herein to “method of DNA synthesis” refer to a method of DNA strand synthesis wherein a DNA strand (n) is extended by adding a further nucleotide (n+1). The method

described herein provides a novel use of the terminal deoxynucleotidyl transferases of the invention and 3'-reversibly blocked nucleotide triphosphates to sequentially add nucleotides in *de novo* DNA strand synthesis which has several advantages over the DNA synthesis methods currently known in the art.

5

Current synthetic methods for coupling nucleotides to form sequence-specific DNA have reached asymptotic length limits, therefore a new method of *de novo* DNA synthesis is required. Synthetic DNA synthesis methods also have the disadvantage of using toxic organic solvents and additives (e.g., acetonitrile, acetic anhydride, trichloroacetic acid, pyridine, etc.), which are harmful to the environment.

10

An alternative, enzymatic method of nucleic acid synthesis is desirable. Natural enzymes such as DNA polymerases are able to add 50,000 nucleotides before disassociation. However, DNA polymerases require a template strand, thereby defeating the purpose of *de novo* strand synthesis.

15

However, a DNA polymerase, called TdT, capable of template-independent DNA synthesis is found in vertebrates. Given a free 3'-end and nucleotide triphosphates, recombinant TdTs from *Bos taurus* and *Mus musculus* were shown to add ten to several hundred nucleotides onto the 3'-end of a DNA strand. As shown in a paper by Basu, M. *et al.* (*Biochem. Biophys. Res. Commun.* (1983) 111, 1105–1112) TdT will uncontrollably add nucleotide triphosphates to the 3'-end of a DNA strand. However, this uncontrolled addition is unsuitable for controlled *de novo* strand synthesis where a sequence-specific oligonucleotide is required. Thus, commercially available recombinant TdT is used primarily as a tool for molecular biologists to label DNA with useful chemical tags.

20

25

The present inventors have discovered several orthologs of *Bos taurus* TdT that, coupled with 3'-reversibly protected nucleotide triphosphates, are able to synthesise DNA in a controlled manner. *Bos taurus* TdT is not efficient at incorporating nucleotide triphosphates with 3'-protecting groups, likely due to steric issues in the TdT active site. Data is presented herein in Example 2 and Figure 3 which demonstrates that orthologs of *Bos taurus* TdT, such as the purified recombinant TdTs of the first aspect of the invention (in particular SEQ ID NOS: 1 and 2), are far more efficient at incorporating 3'-OH blocked nucleotide triphosphates, thereby enabling template-independent, sequence-specific synthesis of nucleic acid strands.

30

35

This enzymatic approach means that the method has the particular advantage of being able to produce DNA strands beyond the 120-200 nucleotide limit of current synthetic DNA synthesis methods. Furthermore, this enzymatic method avoids the need to use strong organic solvents which may be harmful to the environment.

5

It will be understood that the term 'functional equivalent' refers to the polypeptides which are different to the exact sequence of the TdTs of the first aspect of the invention, but can perform the same function, *i.e.*, catalyse the addition of a nucleotide triphosphate onto the 3'-end of a DNA strand in a template dependent manner.

10

In one embodiment, the terminal deoxynucleotidyl transferase (TdT) is a non-natural derivative of TdT, such as a functional fragment or homolog of the TdTs of the first aspect of the invention.

15

References herein to 'fragment' include, for example, functional fragments with a C-terminal truncation, or with an N-terminal truncation. Fragments are suitably greater than 10 amino acids in length, for example greater than 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 360, 370, 380, 390, 400, 410, 420, 430, 440, 450, 460, 470, 480, 490 or 500 amino acids in length.

20

It will be appreciated that references herein to "homology" are to be understood as meaning the percentage identity between two protein sequences, e.g.: SEQ ID NO: X and SEQ ID NO: Y, which is the sum of the common amino acids between aligned sequences SEQ ID NO: X and SEQ ID NO: Y, divided by the shorter length of either SEQ ID NO: X or SEQ ID NO: Y, expressed as a percentage.

25

In one embodiment, the terminal deoxynucleotidyl transferase (TdT) has at least 25% homology with the TdTs of the first aspect of the invention, such as at least 30%, 40%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98% or 99% homology.

30

According to a second aspect of the invention, there is provided a method of nucleic acid synthesis, which comprises the steps of:

- (a) providing an initiator sequence;

(b) adding a 3'-blocked nucleotide triphosphate to said initiator sequence in the presence of a terminal deoxynucleotidyl transferase (TdT) as defined in the first aspect of the invention;

(c) removal of all reagents from the initiator sequence;

5 (d) cleaving the blocking group from the 3'-blocked nucleotide triphosphate in the presence of a cleaving agent;

(e) removal of the cleaving agent.

10 In one embodiment, step (c) comprises removal of deoxynucleotide triphosphates and TdT, such as TdT. Thus, according to one particular aspect of the invention, there is provided a method of nucleic acid synthesis, which comprises the steps of:

(a) providing an initiator sequence;

15 (b) adding a 3'-blocked nucleotide triphosphate to said initiator sequence in the presence of a terminal deoxynucleotidyl transferase (TdT) as defined in the first aspect of the invention;

(c) removal of TdT;

(d) cleaving the blocking group from the 3'-blocked nucleotide triphosphate in the presence of a cleaving agent;

(e) removal of the cleaving agent.

20

It will be understood that steps (b) to (e) of the method may be repeated multiple times to produce a DNA or RNA strand of a desired length. Therefore, in one embodiment, greater than 1 nucleotide is added to the initiator sequence, such as greater than 5, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110 or 120 nucleotides are added to the initiator sequence by repeating
25 steps (b) to (e). In a further embodiment, greater than 200 nucleotides are added, such as greater than 300, 400, 500, 600, 700, 800, 900, 1000, 1250, 1500, 1750, 2000, 2250, 2500, 2750, 3000, 4000, 5000, 6000, 7000, 8000, 9000 or 10000 nucleotides.

3'-blocked nucleotide triphosphates

30 References herein to 'nucleotide triphosphates' refer to a molecule containing a nucleoside (*i.e.* a base attached to a deoxyribose or ribose sugar molecule) bound to three phosphate groups. Examples of nucleotide triphosphates that contain deoxyribose are: deoxyadenosine triphosphate (dATP), deoxyguanosine triphosphate (dGTP), deoxycytidine triphosphate (dCTP) or deoxythymidine triphosphate (dTTP). Examples of nucleotide triphosphates that
35 contain ribose are: adenosine triphosphate (ATP), guanosine triphosphate (GTP), cytidine triphosphate (CTP) or uridine triphosphate (UTP). Other types of nucleosides may be bound

to three phosphates to form nucleotide triphosphates, such as naturally occurring modified nucleosides and artificial nucleosides.

Therefore, references herein to '3'-blocked nucleotide triphosphates' refer to nucleotide triphosphates (e.g., dATP, dGTP, dCTP or dTTP) which have an additional group on the 3' end which prevents further addition of nucleotides, i.e., by replacing the 3'-OH group with a protecting group.

It will be understood that references herein to '3'-block', '3'-blocking group' or '3'-protecting group' refer to the group attached to the 3' end of the nucleotide triphosphate which prevents further nucleotide addition. The present method uses reversible 3'-blocking groups which can be removed by cleavage to allow the addition of further nucleotides. By contrast, irreversible 3'-blocking groups refer to dNTPs where the 3'-OH group can neither be exposed nor uncovered by cleavage.

There exist several documented reversible protecting groups, such as azidomethyl, aminoxy, and allyl, which can be applied to the method described herein. Examples of suitable protecting groups are described in *Greene's Protective Groups in Organic Synthesis*, (Wuts, P.G.M. & Greene, T.W. (2012) 4th Ed., John Wiley & Sons).

In one embodiment, the 3'-blocked nucleotide triphosphate is blocked by a reversible protecting group. In an alternative embodiment, the 3'-blocked nucleotide triphosphate is blocked by an irreversible protecting group.

Therefore, in one embodiment, the 3'-blocked nucleotide triphosphate is blocked by either a 3'-O-methyl, 3'-azido, 3'-O-azidomethyl, 3'-aminoxy or 3'-O-allyl group. In a further embodiment, the 3'-blocked nucleotide triphosphate is blocked by either a 3'-O-azidomethyl, 3'-aminoxy or 3'-O-allyl group. In an alternative embodiment, the 3'-blocked nucleotide triphosphate is blocked by either a 3'-O-methyl or 3'-azido group.

Cleaving agent

References herein to 'cleaving agent' refer to a substance which is able to cleave the 3'-blocking group from the 3'-blocked nucleotide triphosphate.

The 3'-blocking groups described herein may all be quantitatively removed in aqueous solution with documented compounds which may be used as cleaving agents (for example, see: Wuts,

P.G.M. & Greene, T.W. (2012) 4th Ed., John Wiley & Sons; Hutter, D. *et al.* (2010) *Nucleosides Nucleotides Nucleic Acids* 29, 879–895; EP 1560838 and US 7,795,424).

In one embodiment, the cleaving agent is a chemical cleaving agent. In an alternative
5 embodiment, the cleaving agent is an enzymatic cleaving agent.

It will be understood by the person skilled in the art that the selection of cleaving agent is dependent on the type of 3'-nucleotide blocking group used. For instance, tris(2-carboxyethyl)phosphine (TCEP) can be used to cleave a 3'-O-azidomethyl group, palladium
10 complexes can be used to cleave a 3'-O-allyl group, or sodium nitrite can be used to cleave a 3'-aminoxy group. Therefore, in one embodiment, the cleaving agent is selected from: tris(2-carboxyethyl)phosphine (TCEP), a palladium complex or sodium nitrite.

In one embodiment, the cleaving agent is added in the presence of a cleavage solution
15 comprising a denaturant, such as urea, guanidinium chloride, formamide or betaine. The addition of a denaturant has the advantage of being able to disrupt any undesirable secondary structures in the DNA. In a further embodiment, the cleavage solution comprises one or more buffers. It will be understood by the person skilled in the art that the choice of buffer is dependent on the exact cleavage chemistry and cleaving agent required.

Initiator Sequences

References herein to an 'initiator sequence' refer to a short oligonucleotide with a free 3'-end which the 3'-blocked nucleotide triphosphate can be attached to. In one embodiment, the
25 initiator sequence is a DNA initiator sequence. In an alternative embodiment, the initiator sequence is an RNA initiator sequence.

References herein to a 'DNA initiator sequence' refer to a small sequence of DNA which the 3'-blocked nucleotide triphosphate can be attached to, *i.e.* DNA will be synthesised from the end of the DNA initiator sequence.

In one embodiment, the initiator sequence is between 5 and 50 nucleotides long, such as between 5 and 30 nucleotides long (*i.e.* between 10 and 30), in particular between 5 and 20 nucleotides long (*i.e.*, approximately 20 nucleotides long), more particularly 5 to 15 nucleotides long, for example 10 to 15 nucleotides long, especially 12 nucleotides long.

In one embodiment, the initiator sequence has the following sequence: 5'-CGTTAACATATT-3' (SEQ ID NO: 7).

In one embodiment, the initiator sequence is single-stranded. In an alternative embodiment, the initiator sequence is double-stranded. It will be understood by persons skilled in the art that a 3'-overhang (*i.e.*, a free 3'-end) allows for efficient addition.

In one embodiment, the initiator sequence is immobilised on a solid support. This allows TdT and the cleaving agent to be removed (in steps (c) and (e), respectively) without washing away the synthesised nucleic acid. The initiator sequence may be attached to a solid support stable under aqueous conditions so that the method can be easily performed via a flow setup.

In one embodiment, the initiator sequence is immobilised on a solid support via a reversible interacting moiety, such as a chemically-cleavable linker, an antibody/immunogenic epitope, a biotin/biotin binding protein (such as avidin or streptavidin), or glutathione-GST tag. Therefore, in a further embodiment, the method additionally comprises extracting the resultant nucleic acid by removing the reversible interacting moiety in the initiator sequence, such as by incubating with proteinase K.

In a further embodiment, the initiator sequence is immobilised on a solid support via a chemically-cleavable linker, such as a disulfide, allyl, or azide-masked hemiaminal ether linker. Therefore, in one embodiment, the method additionally comprises extracting the resultant nucleic acid by cleaving the chemical linker through the addition of tris(2-carboxyethyl)phosphine (TCEP) or dithiothreitol (DTT) for a disulfide linker; palladium complexes for an allyl linker; or TCEP for an azide-masked hemiaminal ether linker.

In one embodiment, the resultant nucleic acid is extracted and amplified by polymerase chain reaction using the nucleic acid bound to the solid support as a template. The initiator sequence could therefore contain an appropriate forward primer sequence and an appropriate reverse primer could be synthesised.

In an alternative embodiment, the immobilised initiator sequence contains at least one restriction site. Therefore, in a further embodiment, the method additionally comprises extracting the resultant nucleic acid by using a restriction enzyme.

The use of restriction enzymes and restriction sites to cut nucleic acids in a specific location is well known in the art. The choice of restriction site and enzyme can depend on the desired properties, for example whether 'blunt' or 'sticky' ends are required. Examples of restriction enzymes include: Alul, BamHI, EcoRI, EcoRII, EcoRV, HaeII, HgaI, HindIII, HinfI, NotI, PstI, PvuII, Sall, Sau3A, Scal, SmaI, TaqI and XbaI.

In an alternative embodiment, the initiator sequence contains at least one uridine. Treatment with uracil-DNA glycosylase (UDG) generates an abasic site. Treatment on an appropriate substrate with an apurinic/apyrimidinic (AP) site endonuclease will extract the nucleic acid strand.

Nucleic acid synthesis method

In one embodiment, the terminal deoxynucleotidyl transferase (TdT) of the invention is added in the presence of an extension solution comprising one or more buffers (e.g., Tris or cacodylate), one or more salts (e.g., Na⁺, K⁺, Mg²⁺, Mn²⁺, Cu²⁺, Zn²⁺, Co²⁺, etc., all with appropriate counterions, such as Cl⁻) and inorganic pyrophosphatase (e.g., the *Saccharomyces cerevisiae* homolog). It will be understood that the choice of buffers and salts depends on the optimal enzyme activity and stability.

The use of an inorganic pyrophosphatase helps to reduce the build-up of pyrophosphate due to nucleotide triphosphate hydrolysis by TdT. Therefore, the use of an inorganic pyrophosphatase has the advantage of reducing the rate of (1) backwards reaction and (2) TdT strand dismutation. Thus, according to a further aspect of the invention, there is provided the use of inorganic pyrophosphatase in a method of nucleic acid synthesis. Data is presented herein in Example 5 and Figure 6 which demonstrates the benefit of the use of inorganic pyrophosphatase during nucleic acid synthesis. In one embodiment, the inorganic pyrophosphatase comprises purified, recombinant inorganic pyrophosphatase from *Saccharomyces cerevisiae*.

In one embodiment, step (b) is performed at a pH range between 5 and 10. Therefore, it will be understood that any buffer with a buffering range of pH 5-10 could be used, for example cacodylate, Tris, HEPES or Tricine, in particular cacodylate or Tris.

In one embodiment, step (d) is performed at a temperature less than 99°C, such as less than 95°C, 90°C, 85°C, 80°C, 75°C, 70°C, 65°C, 60°C, 55°C, 50°C, 45°C, 40°C, 35°C, or 30°C. It will be understood that the optimal temperature will depend on the cleavage agent utilised. The

temperature used helps to assist cleavage and disrupt any secondary structures formed during nucleotide addition.

In one embodiment, steps (c) and (e) are performed by applying a wash solution. In one
5 embodiment, the wash solution comprises the same buffers and salts as used in the extension solution described herein. This has the advantage of allowing the wash solution to be collected after step (c) and recycled as extension solution in step (b) when the method steps are repeated.

10 In one embodiment, the method is performed within a flow instrument as shown in Figure 7, such as a microfluidic or column-based flow instrument. The method described herein can easily be performed in a flow setup which makes the method simple to use. It will be understood that examples of commercially available DNA synthesisers (e.g., MerMade 192E from BioAutomation or H-8 SE from K&A) may be optimised for the required reaction
15 conditions and used to perform the method described herein.

In one embodiment, the method is performed on a plate or microarray setup. For example, nucleotides may be individually addressed through a series of microdispensing nozzles using any applicable jetting technology, including piezo and thermal jets. This highly parallel process
20 may be used to generate hybridization microarrays and is also amenable to DNA fragment assembly through standard molecular biology techniques.

In one embodiment, the method additionally comprises amplifying the resultant nucleic acid. Methods of DNA/RNA amplification are well known in the art. For example, in a further
25 embodiment, the amplification is performed by polymerase chain reaction (PCR). This step has the advantage of being able to extract and amplify the resultant nucleic acid all in one step.

The template independent nucleic acid synthesis method described herein has the capability
30 to add a nucleic acid sequence of defined composition and length to an initiator sequence. Therefore, it will be understood by persons skilled in the art, that the method described herein may be used as a novel way to introduce adapter sequences to a nucleic acid library.

If the initiator sequence is not one defined sequence, but instead a library of nucleic acid
35 fragments (for example generated by sonication of genomic DNA, or for example messenger RNA) then this method is capable of *de novo* synthesis of 'adapter sequences' on every

fragment. The installation of adapter sequences is an integral part of library preparation for next-generation library nucleic acid sequencing methods, as they contain sequence information allowing hybridisation to a flow cell/solid support and hybridisation of a sequencing primer.

5

Currently used methods include single-stranded ligation, however this technique is limited because ligation efficiency decreases strongly with increasing fragment length. Consequently, current methods are unable to attach sequences longer than 100 nucleotides in length. Therefore, the method described herein allows for library preparation in an improved fashion to that which is currently possible.

10

Therefore, in one embodiment, an adapter sequence is added to the initiator sequence. In a further embodiment, the initiator sequence may be a nucleic acid from a library of nucleic acid fragments.

15

Kits

According to a further aspect of the invention, there is provided the use of a kit in a method of nucleic acid synthesis, wherein said kit comprises a TdT as defined in the first or second aspects of the invention optionally in combination with one or more components selected from: an initiator sequence, one or more 3'-blocked nucleotide triphosphates, inorganic pyrophosphatase, such as purified, recombinant inorganic pyrophosphatase from *Saccharomyces cerevisiae*, and a cleaving agent; further optionally together with instructions for use of the kit in accordance with any of the methods defined herein.

20

Suitably a kit according to the invention may also contain one or more components selected from the group: an extension solution, a wash solution and/or a cleaving solution as defined herein; optionally together with instructions for use of the kit in accordance with any of the methods defined herein.

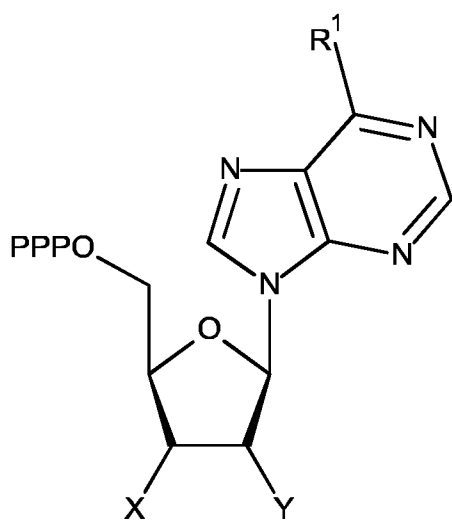
25

Use of 3'-blocked nucleotide triphosphates

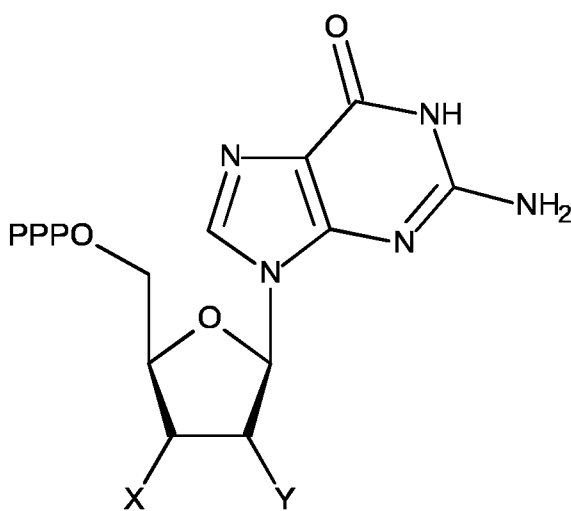
30

According to a further aspect of the invention, there is provided the use of a 3'-blocked nucleotide triphosphate in a method of template independent nucleic acid synthesis, wherein the 3'-blocked nucleotide triphosphate is selected from a compound of formula (I), (II), (III) or (IV):

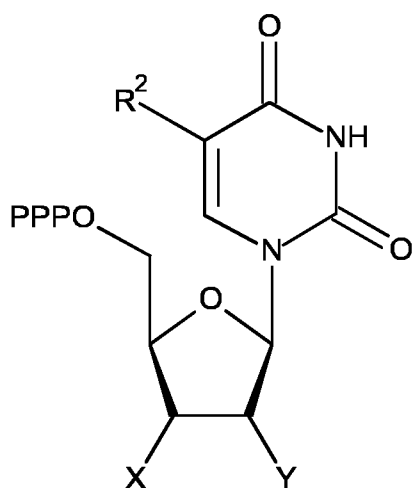
35



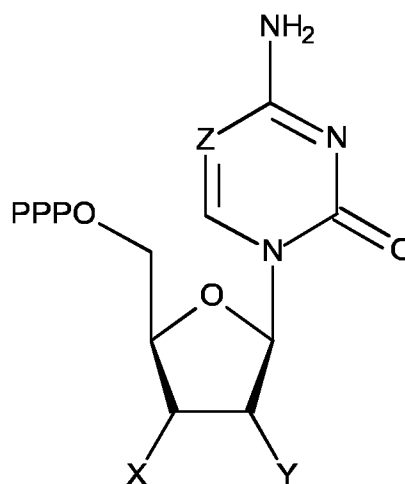
(I)



(II)



(III)



(IV)

5 wherein

R^1 represents NR^aR^b , wherein R^a and R^b independently represent hydrogen or C_{1-6} alkyl,

R^2 represents hydrogen, C_{1-6} alkyl, C_{1-6} alkoxy, COH or COOH.

X represents C_{1-6} alkyl, NH_2 , N_3 or $-OR^3$,

R^3 represents C_{1-6} alkyl, CH_2N_3 , NH_2 or allyl,

10 Y represents hydrogen, halogen or hydroxyl, and

Z represents CR^4 or N, wherein R^4 represents hydrogen, C_{1-6} alkyl, C_{1-6} alkoxy, COH or COOH.

References herein to a "template independent nucleic acid synthesis method" refer to a method of nucleic acid synthesis which does not require a template DNA/RNA strand, *i.e.* the

15 nucleic acid can be synthesised *de novo*.

In one embodiment, the nucleic acid is DNA. References herein to a “template independent DNA synthesis method” refer to a method of DNA synthesis which does not require a template DNA strand, *i.e.* the DNA can be synthesised *de novo*. In an alternative embodiment, the nucleic acid is RNA.

It will be understood that ‘PPP’ in the structures shown herein represents a triphosphate group.

References to the term ‘C₁₋₆ alkyl’ as used herein as a group or part of a group refers to a linear or branched saturated hydrocarbon group containing from 1 to 6 carbon atoms. Examples of such groups include methyl, ethyl, butyl, n-propyl, isopropyl and the like.

References to the term ‘C₁₋₆ alkoxy’ as used herein refer to an alkyl group bonded to oxygen via a single bond (*i.e.* R-O). Such references include those with straight and branched alkyl chains containing 1 to 6 carbon atoms, such as methoxy (or methyloxy), ethyloxy, n-propyloxy, iso-propyloxy, n-butyloxy and 2-methylpropyloxy.

References to the term ‘allyl’ as used herein refer to a substituent with the structural formula RCH₂-CH=CH₂, where R is the rest of the molecule. It consists of a methyl group (-CH₂-) attached to a vinyl group (-CH=CH₂).

References to the term ‘COOH’ or ‘CO₂H’ refer to a carboxyl group (or carboxy) which consists of a carbonyl (C=O) and a hydroxyl (O-H) group. References to the term ‘COH’ refer to a formyl group which consists of a carbonyl (C=O) group bonded to hydrogen.

The term ‘N₃’ (drawn structurally as -N=N⁺=N⁻) refers to an azido group.

In one embodiment, R^a and R^b both represent hydrogen (*i.e.* R¹ represents NH₂).

In an alternative embodiment, R^a represents hydrogen and R^b represents methyl (*i.e.* R¹ represents NHCH₃).

In one embodiment, R² represents hydrogen, methyl or methoxy. In a further embodiment, R² represents hydrogen. In an alternative embodiment, R² represents methyl. In a yet further alternative embodiment, R² represents methoxy.

In one embodiment, X represents $-OR^3$, and R^3 represents C_{1-6} alkyl, CH_2N_3 , NH_2 or allyl.

In an alternative embodiment, X represents C_{1-6} alkyl (such as methyl) or N_3 .

5 In one embodiment, Y represents hydrogen or hydroxyl.

In one embodiment, Y represents hydrogen.

In an alternative embodiment, Y represents hydroxyl.

10

In an alternative embodiment, Y represents halogen, such as fluorine.

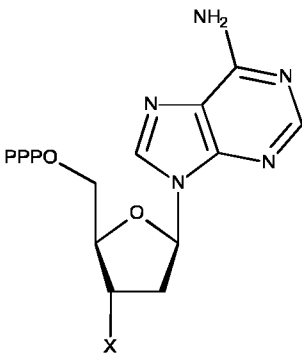
In one embodiment, Z represents N.

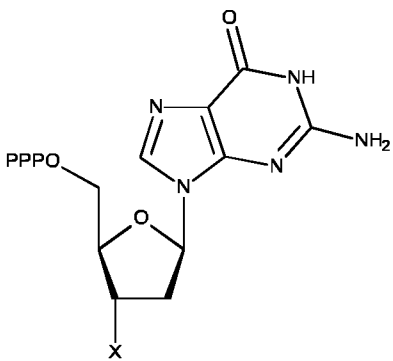
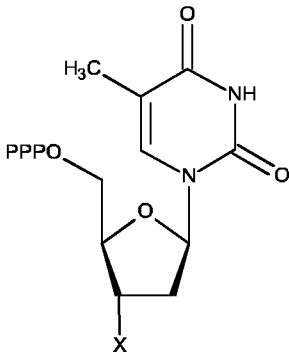
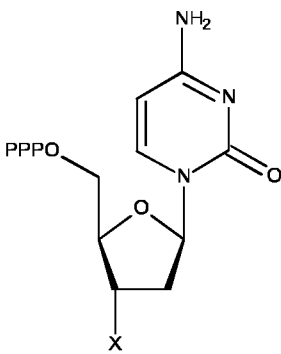
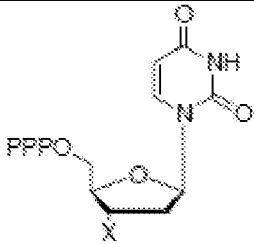
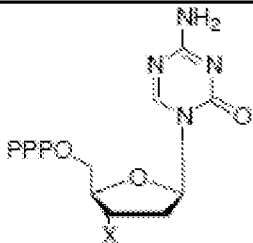
15 In an alternative embodiment, Z represents CR^4 .

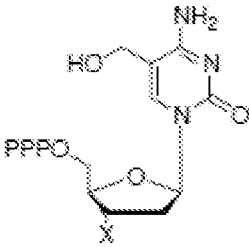
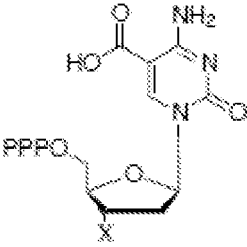
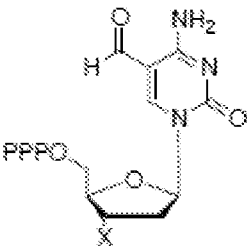
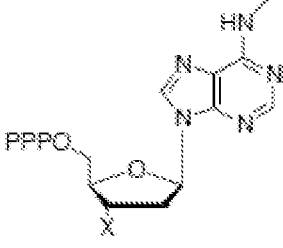
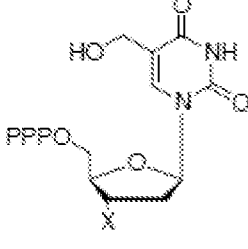
In one embodiment, R^4 represents C_{1-6} alkyl, C_{1-6} alkoxy, COH or COOH.

In a further embodiment, R^4 represents methoxy, COOH or COH. In a yet further embodiment,
 20 R^4 represents methoxy. In an alternative embodiment, R^4 represents COOH. In a yet further alternative embodiment, R^4 represents COH.

In one embodiment, the 3'-blocked nucleotide triphosphate is selected from:

Structure	Name	Example number
	Deoxyadenosine triphosphate	E1

	Deoxyguanosine triphosphate	E2
	Deoxythymidine triphosphate	E3
	Deoxycytidine triphosphate	E4
	2'-deoxy-uridine triphosphate	E5
	5-aza-2'-deoxy-cytidine-triphosphate	E6

	5-hydroxymethyl-deoxycytidine triphosphate	E7
	5-carboxy-deoxycytidine triphosphate	E8
	5-formyl-deoxycytidine triphosphate	E9
	N6-methyladenosine triphosphate	E10
	5-hydroxymethyl-deoxy-uridine triphosphate	E11

wherein 'X' is as defined hereinbefore.

The following studies and protocols illustrate embodiments of the methods described herein:

- 5 **COMPARATIVE EXAMPLE 1: Use of *Bos taurus* TdT to add 3'-irreversibly blocked nucleotide triphosphates to a DNA initiator**

A single-stranded DNA initiator (SEQ ID NO: 7) was incubated with 15 U *Bos taurus* TdT (Thermo Scientific), required salts (50 mM potassium acetate, 20 mM tris acetate pH 7.9, 1 mM cobalt chloride), and 3'-O-Methyl dTTP (TriLink) at 37°C for up to one hour. The 3'-irreversibly blocked nucleotide triphosphate was at a concentration of 1 mM and the DNA
5 initiator at 200 pM for a 1:5000 ratio to encourage nucleotide addition. The reaction was stopped with EDTA (0.5 M) at various intervals and the results are shown in **Figure 2(A)**.

The experiment was repeated with the exception that 3'-Azido dTTP was used as the 3'-irreversibly blocked nucleotide triphosphate instead of 3'-O-Methyl dTTP. The reaction was
10 stopped with EDTA (0.5 M) at various intervals and the results are shown in **Figure 2(B)**.

These studies show that the commercially available *Bos taurus* TdT adds irreversibly blocked nucleotides, specifically 3'-O-Methyl dTTP and 3'-Azido dTTP onto a DNA initiator strand (**Figure 2**). Furthermore, the addition of a 3'-blocked nucleotide triphosphate to a single-
15 stranded DNA initiator is completed to greater than 90% conversion of the n strand to the n+1 strand.

EXAMPLE 2: Use of TdT orthologs other than *Bos Taurus* TdT for controlled template-independent DNA synthesis

A single-stranded DNA initiator (SEQ ID NO: 7) was incubated with purified, recombinant TdT orthologs, required salts (50 mM potassium acetate, 20 mM tris acetate pH 7.9, 1.5 mM cobalt chloride, *Saccharomyces cerevisiae* inorganic pyrophosphatase), and 3'-O-azidomethyl dTTP at 37 °C for 60 min. The 3'-blocked nucleotide triphosphate was at a concentration of 1 mM and the DNA initiator at 200 nM and analysed via capillary electrophoresis as shown in **Figure**
20 **3(A)**. Reactions were quantified after 60 min and results are shown in **Figure 3(B)**.

The experiment was repeated with the exception that both 3'-O-azidomethyl dTTP and 3'-O-azidomethyl dCTP was used as the 3'-blocked nucleotide triphosphate. The reaction was stopped after 20 min and analysed via capillary electrophoresis. Quantified reactions are
30 shown in **Figure 3(C)**.

Bos taurus TdT only efficiently adds irreversibly blocked nucleotides, which is not useful for controlled, enzymatic single-stranded DNA synthesis. This example demonstrates that naturally occurring TdT orthologs other than *Bos taurus* TdT perform significantly better at
35 adding 3'-reversibly blocked nucleotide triphosphates. Better performance, which is judged by

the N+1 addition rate (**Figure 3**), results in longer achievable lengths and greater control of nucleic acid sequence specificity.

EXAMPLE 3: Engineered TdT orthologs demonstrate improved function over wild-type proteins

A single-stranded DNA initiator (SEQ ID NO: 7) was incubated with either a purified wild-type *Lepisosteus oculatus* TdT or a purified, recombinant engineered form of *Lepisosteus oculatus* TdT (SEQ ID NO: 8), required salts, cobalt chloride, *Saccharomyces cerevisiae* inorganic pyrophosphatase, and 3'-O-azidomethyl dTTP at 37 °C for 20 min.

The wild-type *Lepisosteus oculatus* TdT was outperformed by the engineered variant (SEQ ID NO: 8), as demonstrated by improved ability to convert the initiator strand of length n to a strand of length $n+1$, when supplied with 3'-reversibly blocked dCTP, dGTP or dTTP.

EXAMPLE 4: Use of engineered TdT orthologs for template-independent and sequence-specific DNA synthesis

A single-stranded DNA initiator (SEQ ID NO: 7) was incubated with a purified, recombinant engineered form of *Lepisosteus oculatus* TdT, required salts, cobalt chloride, *Saccharomyces cerevisiae* inorganic pyrophosphatase, and 3'-O-azidomethyl dTTP at 37 °C for 10 min.

Bos Taurus TdT was outperformed by an engineered variant of *Lepisosteus oculatus* TdT, as demonstrated by the denaturing PAGE gel in **Figure 5**. This example demonstrates an engineered form of *Lepisosteus oculatus* TdT incorporates 3'-reversibly blocked nucleotides (1) better than the wild-type *Lepisosteus oculatus* TdT, and (2) much better than *Bos Taurus* TdT (see **Figures 2-5**).

EXAMPLE 5: Use of inorganic pyrophosphatase to control nucleic acid sequence specificity

A single-stranded DNA initiator (SEQ ID NO: 7) was incubated with *Bos taurus* TdT at 37 °C for 60 min under the reaction buffer shown above with the exception that the concentration of *Saccharomyces cerevisiae* inorganic pyrophosphatase was varied. When dideoxyTTP (ddTTP) was used, the concentration of the NTP was 0.1 mM. Reactions were analysed by PAGE and are shown in **Figure 6**.

The inorganic pyrophosphatase studies with TdT demonstrate that TdT-mediated DNA synthesis must utilise an inorganic pyrophosphatase in order to control nucleic acid sequence specificity.

5 **EXAMPLE 6: Example DNA Synthesis method using TdT**

1. An immobilised single-stranded DNA initiator is exposed to the extension solution, which is composed of TdT; a base-specific 3'-blocked nucleotide triphosphate; inorganic pyrophosphatase (e.g., the *Saccharomyces cerevisiae* homolog); and any required buffers (e.g., tris(hydroxymethyl)aminomethane (Tris), cacodylate, or any buffer with a buffering range
10 between pH 5-10) and salts (e.g., Na⁺, K⁺, Mg²⁺, Mn²⁺, Cu²⁺, Zn²⁺, Co²⁺, etc., all with appropriate counterions, such as Cl⁻) and reacted at optimised concentrations, times and temperatures. The 3'-blocked nucleotide triphosphate will contain one of the nitrogenous bases adenine, guanine, cytosine or thymine.
- 15 2. The extension mixture is then removed with wash mixture and recycled. Wash mixture is the extension mixture without TdT and the 3'-blocked nucleotide triphosphate.
3. The immobilised (n+1) DNA strand is treated with cleavage mixture composed of an appropriate buffer, denaturant (e.g., urea, guanidinium chloride, formamide, betaine, etc.), and
20 cleavage agent (e.g., tris(2-carboxyethyl)phosphine (TCEP) to de-block a 3'-O-azidomethyl group; palladium complexes to de-block 3'-O-allyl group; or sodium nitrite to de-block the 3'-aminoxy group). The temperature can be raised up to 99°C to assist in cleavage and disruption of secondary structures. The optimal temperature depends on the cleavage agent utilised.
- 25 4. The immobilised deblocked (n+1) DNA strand is treated with wash mixture to remove the cleavage mixture.
5. Cycles 1-4 are repeated with the base-specific 3'-blocked nucleotide triphosphate until the desired oligonucleotide sequence is achieved.
30
6. Once the desired sequence is achieved, polymerase chain reaction with primers specific to the DNA product is used to directly "extract" and amplify the product.

CLAIMS

1. Use of a terminal deoxynucleotidyl transferase (TdT) enzyme comprising an amino acid sequence selected from either: (a) any one of SEQ ID NOS: 1 to 5 and 8 or a functional
5 equivalent or fragment thereof having at least 20% sequence homology to said amino acid sequence; or (b) a modified derivative of SEQ ID NO: 6; in a method of nucleic acid synthesis.
2. The use as defined in claim 1, wherein the terminal deoxynucleotidyl transferase
10 (TdT) enzyme comprises an amino acid sequence selected from any one of SEQ ID NOS: 1 to 5 and 8 or a functional equivalent or fragment thereof having at least 20% sequence homology to said amino acid sequence.
3. The use as defined in claim 1 or claim 2, wherein the terminal deoxynucleotidyl
15 transferase (TdT) enzyme comprises an amino acid sequence selected from SEQ ID NOS: 1, 2 or 8 or a functional equivalent or fragment thereof having at least 20% sequence homology to said amino acid sequence.
4. A method of nucleic acid synthesis, which comprises the steps of:
20 (a) providing an initiator sequence;
(b) adding a 3'-blocked nucleotide triphosphate to said initiator sequence in the presence of a terminal deoxynucleotidyl transferase (TdT) as defined in any one of claims 1 to 3;
(c) removal of all reagents from the initiator sequence;
25 (d) cleaving the blocking group from the 3'-blocked nucleotide triphosphate in the presence of a cleaving agent;
(e) removal of the cleaving agent.
5. The method as defined in claim 4, wherein greater than 1 nucleotide is added by
30 repeating steps (b) to (e).
6. The method as defined in claim 4 or claim 5, wherein the 3'-blocked nucleotide triphosphate is blocked by either a 3'-O-azidomethyl, 3'-aminoxy or 3'-O-allyl group.
- 35 7. The method as defined in any one of claims 4 to 6, wherein the terminal deoxynucleotidyl transferase (TdT) is added in the presence of an extension solution

comprising one or more buffers, such as Tris or cacodylate, one or more salts, and inorganic pyrophosphatase, such as purified, recombinant inorganic pyrophosphatase from *Saccharomyces cerevisiae*.

5 8. The method as defined in any one of claims 4 to 7, wherein step (b) is performed at a pH range between 5 and 10.

9. The method as defined in any one of claims 4 to 8, wherein the cleaving agent is a chemical cleaving agent.

10

10. The method as defined in any one of claims 4 to 9, wherein the cleaving agent is an enzymatic cleaving agent.

15

11. The method as defined in any one of claims 4 to 10, wherein the cleaving agent is selected from: tris(2-carboxyethyl)phosphine (TCEP), a palladium complex or sodium nitrite.

12. The method as defined in any one of claims 4 to 11, wherein the cleaving agent is added in the presence of a cleavage solution comprising a denaturant, such as urea, guanidinium chloride, formamide or betaine, and one or more buffers.

20

13. The method as defined in any one of claims 4 to 12, wherein step (d) is performed at a temperature less than 99°C.

25

14. The method as defined in any one of claims 4 to 13, wherein steps (c) and (e) are performed by applying a wash solution.

15. The method as defined in any one of claims 4 to 14, wherein the method is performed within a flow instrument, such as a microfluidic or column-based flow instrument.

30

16. The method as defined in any one of claims 4 to 14, wherein the method is performed within a plate or microarray setup.

35

17. The method as defined in any one of claims 4 to 16, wherein the initiator sequence is between 5 and 50 nucleotides long, such as between 10 and 30 nucleotides long, in particular approximately 20 nucleotides long.

18. The method as defined in any one of claims 4 to 17, wherein the initiator sequence is single-stranded or double-stranded.

5 19. The method as defined in any one of claims 4 to 18, wherein the initiator sequence is immobilised on a solid support.

20. The method as defined in claim 19, wherein the initiator sequence is immobilised via a reversible interacting moiety.

10

21. The method as defined in claim 20, which additionally comprises extracting the resultant nucleic acid by removing the reversible interacting moiety.

15

22. The method as defined in claim 20, wherein the reversible interacting moiety is a chemically-cleavable linker, such as a disulfide, allyl or azide-masked hemiaminal ether.

20

23. The method as defined in claim 22, which additionally comprises extracting the resultant nucleic acid by cleaving the chemically-cleavable linker, such as by the addition of tris(2-carboxyethyl)phosphine (TCEP), dithiothreitol (DTT) or a palladium complex.

24. The method as defined in any one of claims 4 to 19, wherein the initiator sequence contains at least one restriction site.

25

25. The method as defined in claim 24, which additionally comprises extracting the resultant nucleic acid by using a restriction enzyme.

26. The method as defined in any one of claims 4 to 25, wherein the initiator sequence contains at least one uridine.

30

27. The method as defined in any one of claims 4 to 26, which additionally comprises amplifying the resultant nucleic acid, such as by PCR.

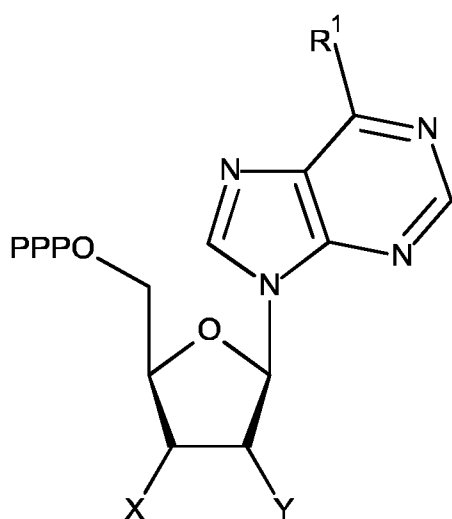
35

28. Use of a kit in a method of nucleic acid synthesis, wherein said kit comprises a TdT as defined in any one of claims 1 to 3, optionally in combination with one or more components selected from: an initiator sequence, one or more 3'-blocked nucleotide triphosphates, inorganic pyrophosphatase, such as purified, recombinant inorganic

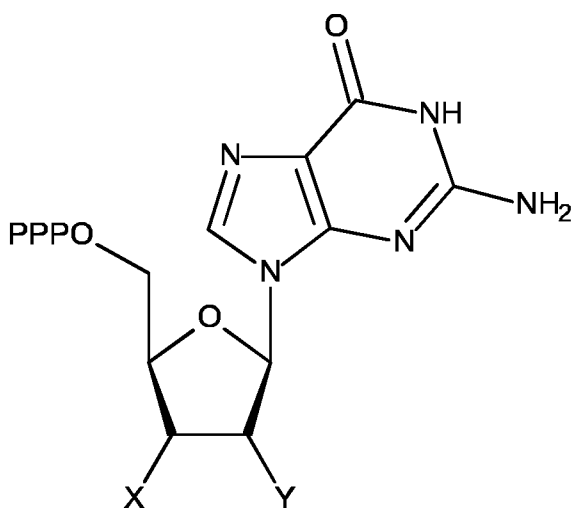
pyrophosphatase from *Saccharomyces cerevisiae*, and a cleaving agent; further optionally together with instructions for use of the kit in accordance with the method as defined in any one of claims 4 to 27.

- 5 29. The use as defined in claim 28, which additionally comprises one or more components selected from the group: an extension solution, a wash solution and/or a cleaving solution.

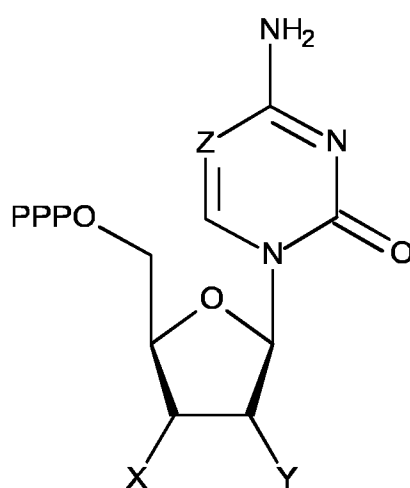
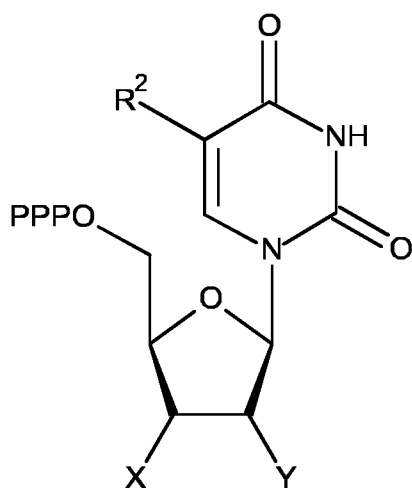
- 10 30. Use of a 3'-blocked nucleotide triphosphate in a method of template independent nucleic acid synthesis, wherein the 3'-blocked nucleotide triphosphate is selected from a compound of formula (I), (II), (III) or (IV):



(I)



(II)



(III)

(IV)

wherein

R¹ represents NR^aR^b, wherein R^a and R^b independently represent hydrogen or C₁₋₆ alkyl,

R² represents hydrogen, C₁₋₆ alkyl, C₁₋₆ alkoxy, COH or COOH.

5 X represents C₁₋₆ alkyl, NH₂, N₃ or -OR³,

R³ represents C₁₋₆ alkyl, CH₂N₃, NH₂ or allyl,

Y represents hydrogen, halogen or hydroxyl, and

Z represents CR⁴ or N, wherein R⁴ represents hydrogen, C₁₋₆ alkyl, C₁₋₆alkoxy, COH or COOH.

10 31. The use as defined in claim 30, wherein R^a and R^b both represent hydrogen or R^a represents hydrogen and R^b represents methyl.

32. The use as defined in claim 30 or claim 31, wherein R² represents hydrogen, methyl or methoxy.

15

33. The use as defined in any one of claims 30 to 32, wherein X represents -OR³ and R³ represents C₁₋₆ alkyl, CH₂N₃, NH₂ or allyl.

34. The use as defined in any one of claims 30 to 33, wherein Y represents hydrogen.

20

35. The use as defined in any one of claims 30 to 34, wherein Z represents CR⁴ and wherein R⁴ represents methoxy, COH or COOH.

25 36. The use as defined in any one of claims 30 to 35, wherein the 3'-blocked nucleotide triphosphate is selected from E1 to E11.

37. Use of inorganic pyrophosphatase in a method of nucleic acid synthesis.

30 38. The use as defined in claim 37, wherein the inorganic pyrophosphatase is purified, recombinant inorganic pyrophosphatase from *Saccharomyces cerevisiae*.

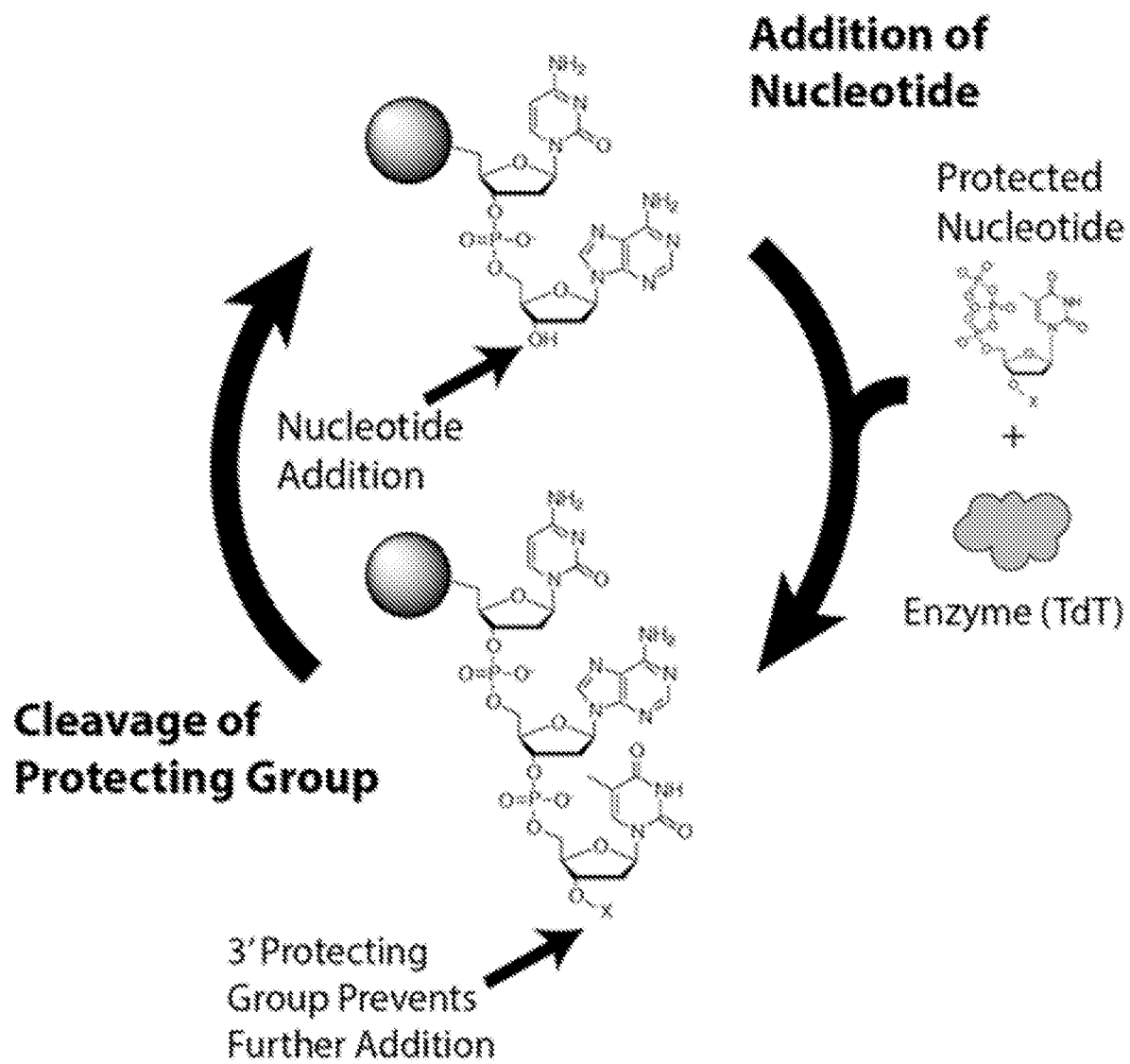


FIGURE 1

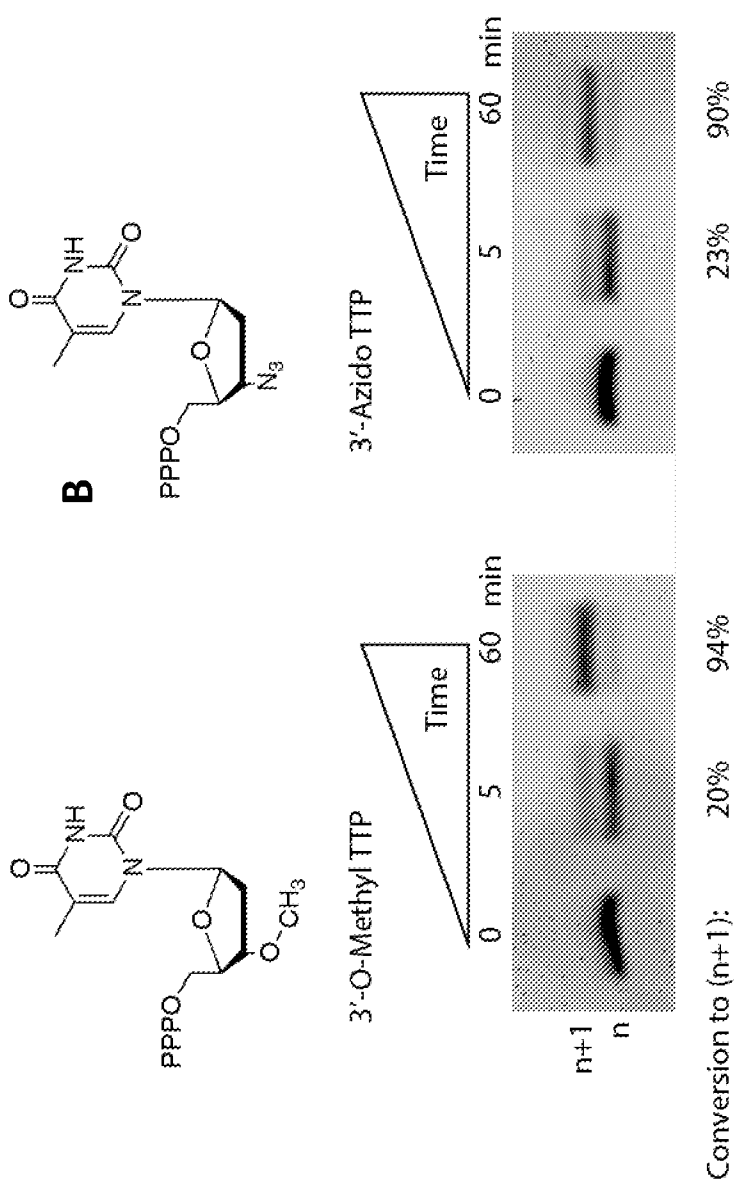


FIGURE 2

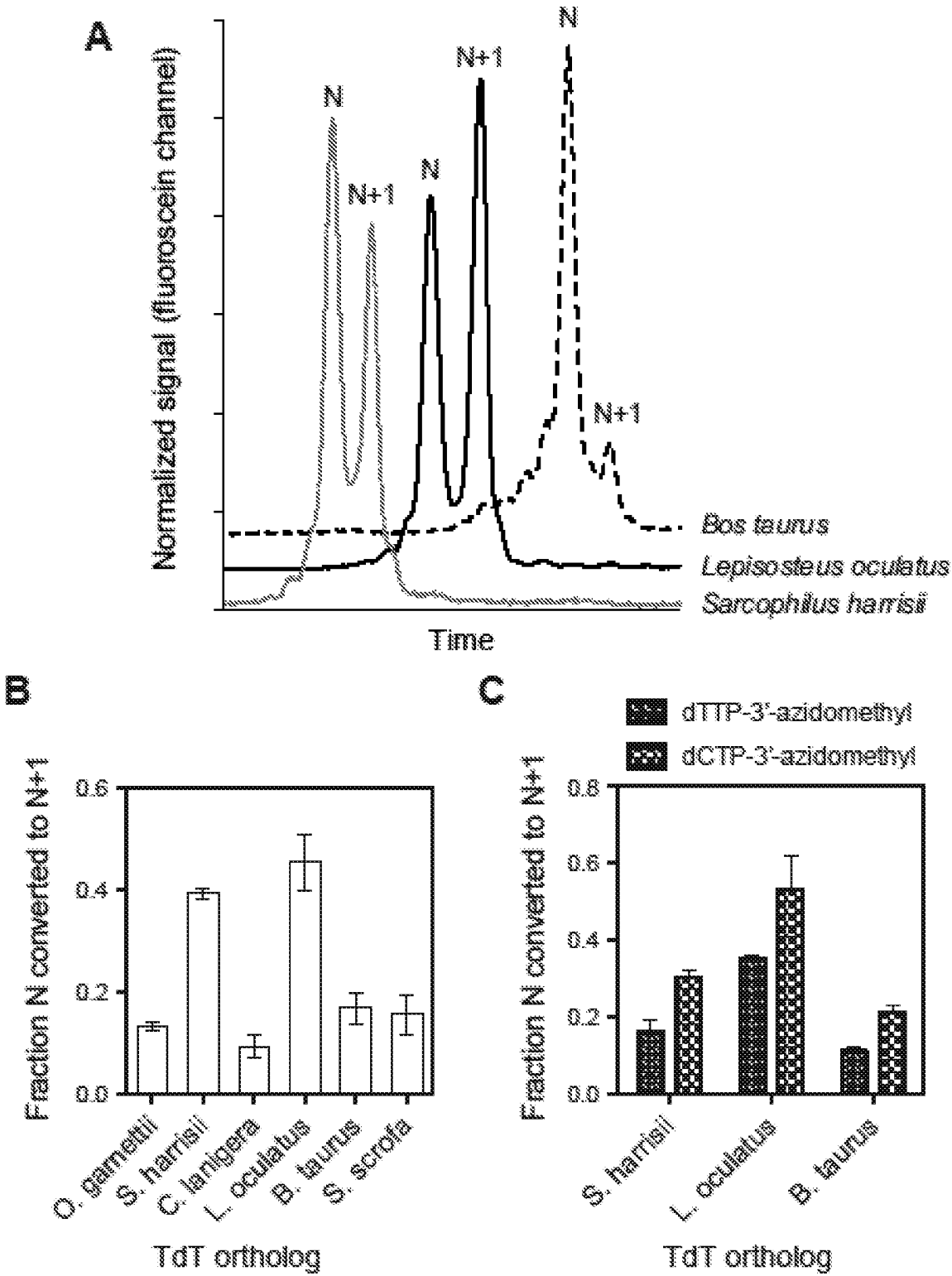


FIGURE 3

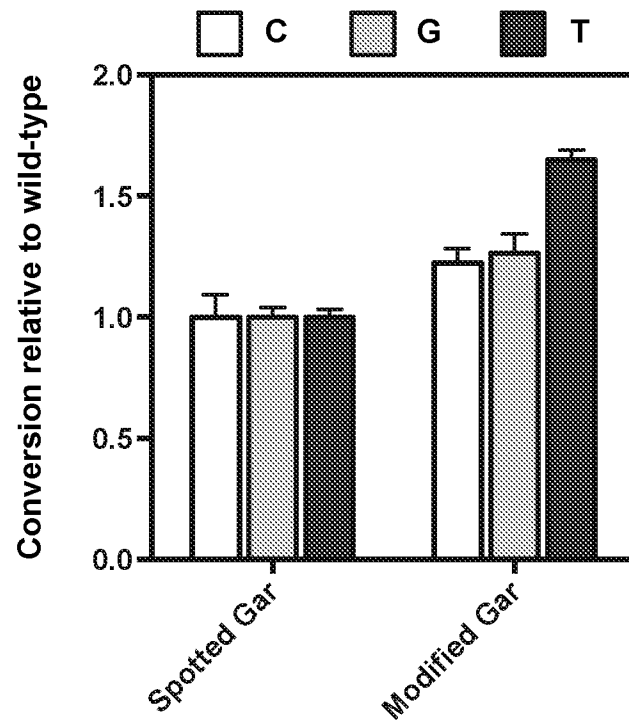


FIGURE 4

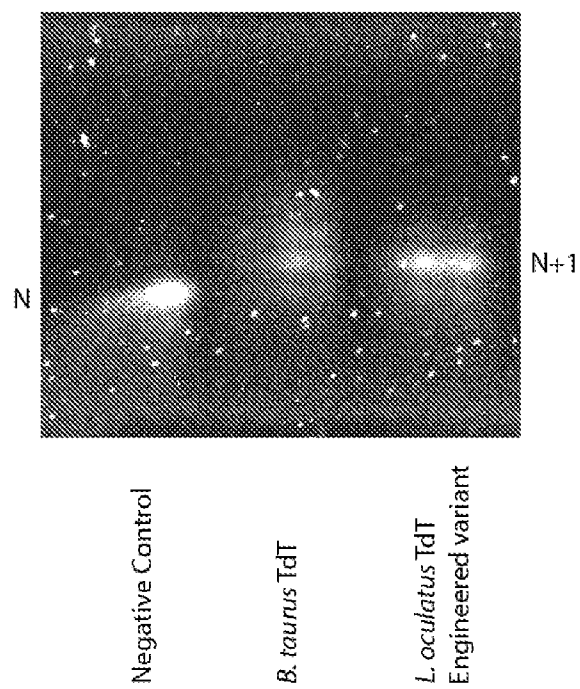


FIGURE 5

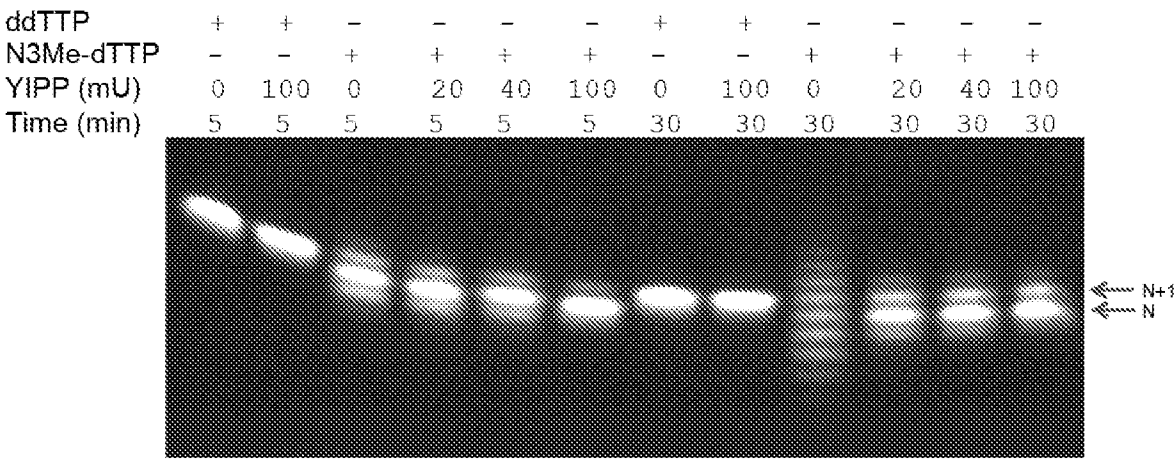


FIGURE 6

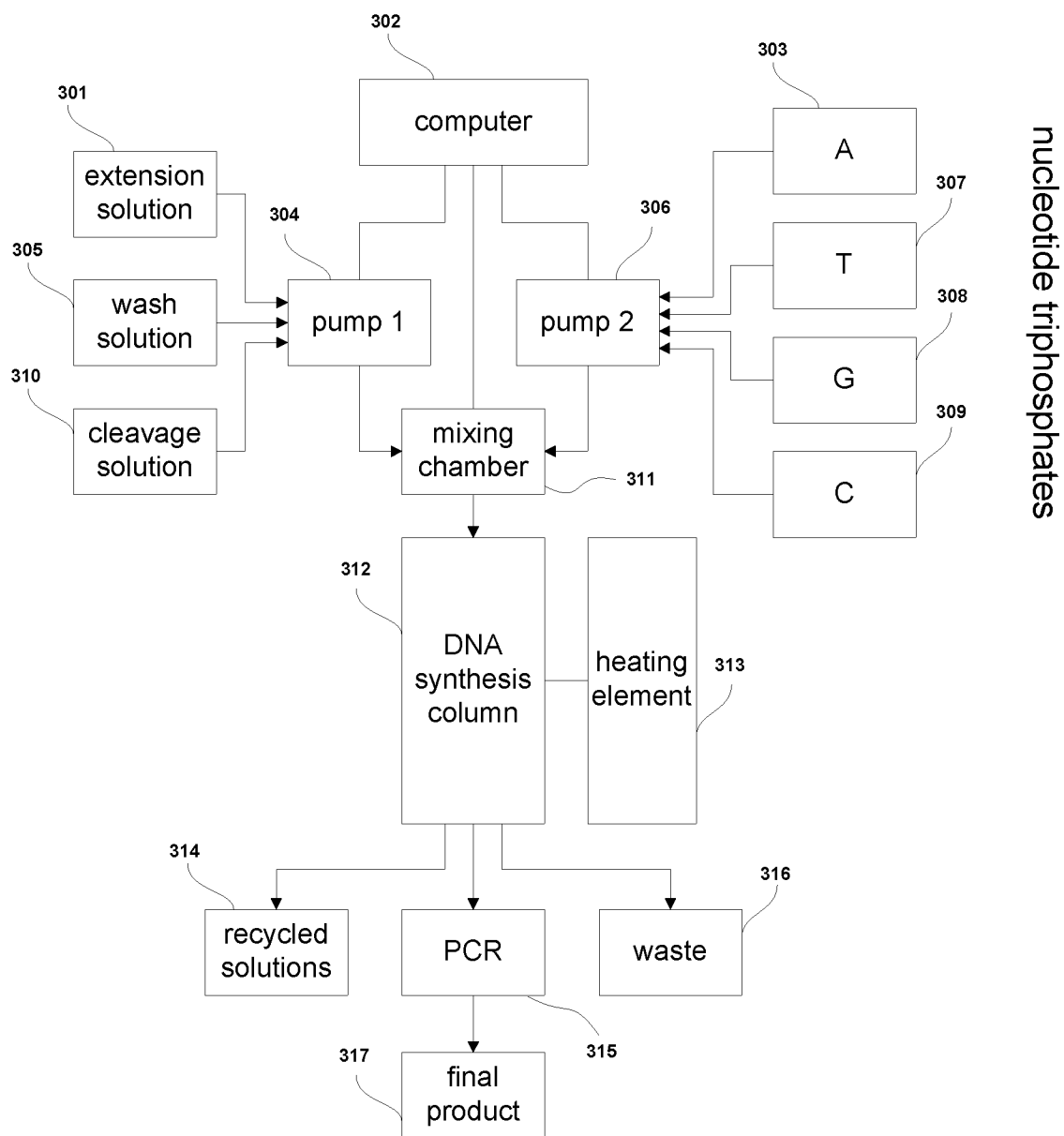


FIGURE 7

INTERNATIONAL SEARCH REPORT

International application No
PCT/GB2016/050301

A. CLASSIFICATION OF SUBJECT MATTER

INV. C12P19/34 C12Q1/68 C12N9/12
ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

C12P C12Q C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EPO-Internal, Sequence Search

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	DATABASE UniProt [Online] 16 November 2011 (2011-11-16), "SubName: Full=Uncharacterized protein {ECO:0000313 Ensembl:ENSSHAP000000005310};" , XP002756023, retrieved from EBI accession no. UNIPROT:G3VQ55 Database accession no. G3VQ55 sequence ----- -/--	1-29



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

5 April 2016

Date of mailing of the international search report

28/06/2016

Name and mailing address of the ISA/

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040,
Fax: (+31-70) 340-3016

Authorized officer

López García, F

INTERNATIONAL SEARCH REPORT

International application No
PCT/GB2016/050301

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	S. M. MINHAZ UD-DEAN: "A theoretical model for template-free synthesis of long DNA sequence", SYSTEMS AND SYNTHETIC BIOLOGY, vol. 2, no. 3-4, 1 December 2008 (2008-12-01), pages 67-73, XP055262384, NL ISSN: 1872-5325, DOI: 10.1007/s11693-009-9023-x Abstract; Figs. 1, 2. -----	1-29
X	US 8 808 989 B1 (EFCAVITCH J WILLIAM [US] ET AL) 19 August 2014 (2014-08-19) cited in the application	1-3
Y	Abstract; col. 18, l. 47-col. 19, l. 14; Fig. 13, claim 1. -----	1-29
A	WO 03/050242 A2 (RUBICON GENOMICS INC [US]; MAKAROV VLADIMIR L [US]; SLEPTSOVA IRINA [U] 19 June 2003 (2003-06-19) cited in the application Abstract; paragraphs [0106]-[0108]; example 13. -----	1-29
A	BEABEALASHVILLI R S ET AL: "Nucleoside 5'-triphosphates modified at sugar residues as substrates for calf thymus terminal deoxynucleotidyl transferase and for AMV reverse transcriptase", BIOCHIMICA ET BIOPHYSICA ACTA . GENE STRUCTURE AND EXPRESSION, ELSEVIER, AMSTERDAM, NL, vol. 868, no. 2-3, 13 November 1986 (1986-11-13), pages 136-144, XP025209880, ISSN: 0167-4781, DOI: 10.1016/0167-4781(86)90016-3 [retrieved on 1986-11-13] Abstract. -----	1-29
Y	DATABASE UniProt [Online] 16 April 2014 (2014-04-16), "SubName: Full=Uncharacterized protein {ECO:0000313 Ensembl:ENSLOCP00000008791};" , XP002756024, retrieved from EBI accession no. UNIPROT:W5MK82 Database accession no. W5MK82 sequence ----- -/--	1-29

INTERNATIONAL SEARCH REPORT

International application No
PCT/GB2016/050301

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	DATABASE UniProt [Online] 14 December 2011 (2011-12-14), "SubName: Full=DNA nucleotidylexotransferase {ECO:0000313 EMBL:EHB07806.1}";", XP002756025, retrieved from EBI accession no. UNIPROT:G5BEU5 Database accession no. G5BEU5 sequence	1-29
Y	----- DATABASE UniProt [Online] 15 May 2007 (2007-05-15), "SubName: Full=Deoxynucleotidyltransferase, terminal {ECO:0000313 EMBL:BAF51681.1}; SubName: Full=Uncharacterized protein {ECO:0000313 Ensembl:ENSOGAP00000013496}";", XP002756026, retrieved from EBI accession no. UNIPROT:A4PCE6 Database accession no. A4PCE6 sequence	1-29
Y	----- DATABASE UniProt [Online] 3 May 2011 (2011-05-03), "SubName: Full=Uncharacterized protein {ECO:0000313 Ensembl:ENSSSCP00000011193}";", XP002756027, retrieved from EBI accession no. UNIPROT:F1SBG2 Database accession no. F1SBG2 sequence	1-29
Y	----- DATABASE UniProt [Online] 1 January 1988 (1988-01-01), "RecName: Full=DNA nucleotidylexotransferase; EC=2.7.7.31; AltName: Full=Terminal addition enzyme; AltName: Full=Terminal deoxynucleotidyltransferase; Short=TDI; Short=Terminal transferase;", XP002756028, retrieved from EBI accession no. UNIPROT:P06526 Database accession no. P06526 sequence -----	1-29

INTERNATIONAL SEARCH REPORT

International application No.
PCT/GB2016/050301

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fees, this Authority did not invite payment of additional fees.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

1-29

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- ☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- ☐ No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. claims: 1-29

Methods of nucleic acid synthesis characterized by the deoxynucleotidyl transferases of sequences 1-6, 8.

2. claims: 30-36

Methods of nucleic acid synthesis characterized by the use of 3'-blocked nucleotide triphosphates I-IV.

3. claims: 37, 38

Methods of nucleic acid synthesis characterized by the use of inorganic pyrophosphatases

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/GB2016/050301

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 8808989	B1	19-08-2014	AU 2014247928 A1 29-10-2015
		CA 2908690 A1 09-10-2014	
		EP 2796552 A2 29-10-2014	
		US 8808989 B1 19-08-2014	
		US 2014363851 A1 11-12-2014	
		WO 2014162307 A2 09-10-2014	
		WO 2014165864 A2 09-10-2014	

WO 03050242	A2	19-06-2003	AU 2002359436 A1 23-06-2003
		EP 1451365 A2 01-09-2004	
		JP 2005535283 A 24-11-2005	
		US 2003143599 A1 31-07-2003	
		US 2010145037 A1 10-06-2010	
		US 2015087528 A1 26-03-2015	
		WO 03050242 A2 19-06-2003	
