



(51) International Patent Classification:

C12Q 1/6806 (2018.01)

(21) International Application Number:

PCT/GB2021/052095

(22) International Filing Date:

12 August 2021 (12.08.2021)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

2012544.9

12 August 2020 (12.08.2020)

GB

(71) Applicant: NUCLERA NUCLEICS LTD [GB/GB]; 137 Cambridge Science Park, Milton Road, Cambridge CB4 0GD (GB).

(72) Inventors: MCINROY, Gordon Ross; c/o Nuclera Nucleics Ltd, 137 Cambridge Science Park, Milton Road, Cambridge CB4 0GD (GB). OST, Tobias William Barr; c/o Nuclera Nucleics Ltd, 137 Cambridge Science Park, Milton Road, Cambridge CB4 0GD (GB). LOVEDALE, Do-

minic; c/o Nuclera Nucleics Ltd, 137 Cambridge Science Park, Milton Road, Cambridge CB4 0GD (GB).

(74) Agent: STRATAGEM IPM LIMITED; Meridian Court, Comberton Road, Toft, Cambridge CB23 2RY (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, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, IT, JO, JP, KE, KG, KH, KN, KP, KR, KW, 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, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, 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,

(54) Title: METHODS RELATING TO DE NOVO ENZYMATIC NUCLEIC ACID SYNTHESIS

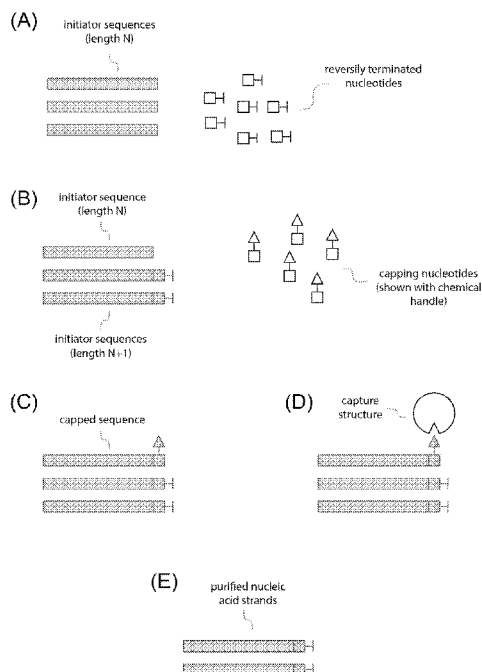


FIGURE 1

(57) Abstract: The invention relates to methods and compositions for the ensuring the quality of oligonucleotide synthesis during cycles of template independent terminal transferase extension. More specifically the invention relates to methods and kits containing a method for capping and removing un-extended strands that failed to correctly incorporate terminated nucleotides.

MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM,
TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW,
KM, ML, MR, NE, SN, TD, TG).

Published:

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

Methods relating to de novo enzymatic nucleic acid synthesis**FIELD OF THE INVENTION**

The invention relates to methods and compositions for the ensuring the quality of oligonucleotide synthesis during cycles of template independent terminal transferase extension. More specifically the invention relates to methods and kits for capping and removing un-extended nucleic acid strands.

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 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. Despite being a mature technology, it is highly challenging to synthesise a DNA strand greater than 200 nucleotides in length in viable yield, and most DNA synthesis companies only offer up to 120 nucleotides routinely. In comparison, an average protein-coding gene is of the order of 2000-3000 contiguous nucleotides, a chromosome is at least a million contiguous nucleotides in length and an average eukaryotic genome numbers in the billions of nucleotides. In order to prepare nucleic acid strands thousands of base pairs in length, 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 enzymatic copying and extension. Current methods generally allow up to 3 kb in length for routine production.

The reason DNA cannot be routinely synthesised beyond 120-200 nucleotides at a time is due to the current chemical methodology for generating DNA, which uses synthetic chemistry (i.e., phosphoramidite technology) to couple a nucleotide one at a time to make DNA. Even if the efficiency of each nucleotide-coupling step is 99% 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.

5 Known methods of DNA sequencing use template-dependent DNA polymerases to add 3'-reversibly terminated nucleotides to a growing double-stranded substrate. 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 can produce strands of between 500-1000 bps long. However, this technology is not suitable for de novo nucleic acid
10 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 de novo single-stranded DNA synthesis. Uncontrolled de novo single stranded DNA synthesis, as opposed to controlled, takes advantage of TdT's deoxynucleotide triphosphate (dNTP) 3' tailing properties on
15 single-stranded DNA to create, for example, homopolymeric adaptor sequences for next-generation sequencing library preparation. In controlled extensions, 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
20 microarrays as it removes the need for an anhydrous environment and allows the use of various polymers incompatible with organic solvents. However, naturally occurring TdTs have not been shown to efficiently add nucleoside triphosphates containing 3'-O-reversibly terminating moieties for building up a nascent single-stranded DNA chain necessary for a de novo synthesis cycle. Thus, engineered TdT enzymes have been created that are capable of efficient de novo DNA synthesis.

25 An aspect affecting synthesis quality is the efficiency of extension. Strands that fail to incorporate effectively fall out of phase on the following cycles. A capping mechanism for blocking and removing unextended strands is therefore an improvement.

30 The invention relates to methods and kits for the purification of nucleic acid strands generated by de novo enzymatic nucleic acid synthesis. The invention relates to a method of capping unreacted nucleic acid strands after each addition step, comprising the use of a template independent polymerase to install a capping moiety. The invention relates to a method of purifying products of de novo enzymatic nucleic acid synthesis comprising a capping step that employs a template

independent polymerase to install a chemical handle that can be used to remove capped nucleic acid strands during or after the synthesis process.

BRIEF DESCRIPTION OF THE FIGURES

5 **Figure 1:** A schematic showing a capping process. (A) Installation of 3'-reversibly terminators by a template independent polymerase. (B) Capping of strands that failed to undergo addition by means of a template independent polymerase installing a capping moiety. (C) Pool of initiator sequences that have undergone either addition of a reversible terminator or a capping moiety. (D) Exposure of the product pool to a capture structure, which will interact with the chemical handle present in the
10 capping moiety. (E) The purified product pool, with any strands bearing a chemical handle removed.

Figure 2: A schematic showing a capping process. (A) Installation of 3'-reversibly terminators by a template independent polymerase. (B) Capping of strands that failed to undergo addition by means of a template independent polymerase installing a blocking capping moiety. (C) Pool of initiator
15 sequences that have undergone either addition of a reversible terminator or a capping moiety. (D) Treatment of the pool with an appropriate deblocking solution. (E) The pool having repeated steps (A)-(D) five more times (six cycles in total). Strands that have a capping moiety installed are terminated by its blocking nature. The pool is exposed to a capture structure, leading to an interaction (chemical or affinity) with the chemical handle. (F) The purified product pool, with any
20 strands bearing a chemical handle removed.

Figure 3: A schematic showing a non-blocking capping process. (A) Installation of 3'-reversibly terminators by a template independent polymerase. (B) Capping of strands that failed to undergo addition by means of a template independent polymerase installing a non-blocking capping moiety.
25 (C) Pool of initiator sequences that have undergone either addition of a reversible terminator or a capping moiety. (D) Treatment of the pool with an appropriate deblocking solution. (E) The pool having repeated steps (A)-(D) five more times (six cycles in total). Strands that have a capping moiety installed are not terminated so may undergo further addition to full length. The pool is exposed to a capture structure, leading to an interaction (chemical or affinity) with the chemical handle. (F) The
30 purified product pool, with any strands bearing at least one chemical handle removed.

Figure 4: A schematic showing a non-blocking digestion-based capping process. (A) Installation of 3'-reversibly terminators by a template independent polymerase. (B) Capping of strands that failed to undergo addition by means of a template independent polymerase installing a non-blocking capping

moiety (e.g. a non-canonical nucleotide). (C) Pool of initiator sequences that have undergone either addition of a reversible terminator or a capping moiety. As the blocking moiety is non-blocking, it is possible for initiator sequences to be capped with one or more units. (D) Treatment of the pool with an appropriate deblocking solution. (E) The pool having repeated steps (A)-(D) five more times (six cycles in total). Strands that have a capping moiety installed are not terminated so may undergo further addition to full length. The pool is exposed to a digestion agent that recognizes the capping moiety and cleaves the DNA strand, for instance leaving a 3'-phosphate moiety. (F) The product pool after digestion. Any digested species are not viable in downstream processes, e.g. ligation or further extension.

Figure 5: A schematic of an experiment demonstrating how capping with a moiety that enables subsequent digestion of capped strands is beneficial. The results from the experiment described in this figure are described in the examples and Figure 6. Briefly, a DNA initiator immobilised to a solid support is subjected to five cycles of *de novo* enzymatic DNA synthesis to generate a sequence, such as TAGCG and the 3' end is deprotected to leave a 3'-hydroxyl moiety. The strands are then subjected to capping with dUTP which generates a homopolymer tail of nucleobases which are vulnerable to enzymatic digestion, for example with uracil DNA glycosylase (UDG). The material is then split into two samples. One sample is extended with a further cycle of addition to incorporate a dC nucleobase, while the other sample undergoes a mock addition (without dCTP-ONH₂ nucleotide) and so does not experience addition. Each sample is then split further, whereby half the samples are treated with UDG and half are not. The experiment is analysed by next-generation DNA sequencing (NGS).

Figure 6: Results of Experiment described in Figure 5.

SUMMARY OF THE INVENTION

The invention relates to methods and kits for capping un-extended nucleic acid strands during a process of enzymatic nucleic acid synthesis.

Specifically the invention may include a method of template independent nucleic acid synthesis comprising

- a. providing an immobilized initiator oligonucleotide of length N bases;
- b. extending the immobilized initiator oligonucleotides with a 3'-O reversibly terminated nucleotide species and a template independent polymerase;

- c. extending any remaining immobilized initiator oligonucleotide of length N bases with a capping species having a moiety allowing for capture or digestion of the capped strands;
 - d. removing all reagents from the initiator sequence;
 - e. cleaving the 3'-O terminating group in the presence of a cleaving agent; and
 - 5 f. removing the cleaving agent by washing the immobilized initiator nucleic acid with wash solution,
- repeating steps b. to f. to add greater than one nucleotide to the above nucleic acid initiator sequence whilst capping the unextended oligonucleotide products; and
- g. capturing or digesting the capped strands.

10

Specifically the invention may include a method of template independent nucleic acid synthesis comprising,

- a. providing an immobilized initiator oligonucleotide of length N bases;
- b. extending the immobilized initiator oligonucleotides with a 3'-O reversibly terminated nucleotide species and a template independent polymerase;
- 15 c. extending any remaining unreacted immobilized initiator oligonucleotide of length N bases with a capping species having a moiety allowing for capture or digestion of the capped strands, wherein the capping species is a nucleotide;
- d. cleaving the 3'-O terminating group in the presence of a cleaving agent; and
- 20 repeating steps b. to d. to add greater than one nucleotide to the above nucleic acid initiator sequence whilst capping the unextended oligonucleotide products to produce an oligonucleotide product pool; and
- e. capturing or digesting the capped strands to remove them from the product pool.

25

DETAILED DESCRIPTION

Described herein are methods, kits, and compositions for the enzymatic installation of a chemical handle to strands that have failed to undergo enzymatically mediated addition of a reversibly terminated nucleotide and would otherwise introduce deletion errors into a nucleic acid product. In some embodiments, installation of the chemical handle caps the nucleic acid containing the error. In other embodiments, installation of the chemical handle does not cap the nucleic acid containing the error. In such cases, further synthesis can continue but the error containing nucleic acid may be removed in a downstream purification step. Nucleotides having a free 3'-OH moiety are efficiently incorporated, and can therefore be used to extend strands that have failed to be extended using a

30

terminated nucleotide. Such extension can incorporate multiple nucleotide tails, which can be captured or digested.

Disclosed is a method of template independent nucleic acid synthesis comprising,

- 5 a. providing an immobilized initiator oligonucleotide of length N bases;
- b. extending the immobilized initiator oligonucleotides with a 3'-O reversibly terminated nucleotide species and a template independent polymerase;
- c. extending any remaining unreacted immobilized initiator oligonucleotide of length N bases with a capping species having a moiety allowing for capture or digestion of the
- 10 capped strands;
- d. cleaving the 3'-O terminating group in the presence of a cleaving agent;
- repeating steps b. to d. to add greater than one nucleotide to the above nucleic acid initiator sequence whilst capping the unextended oligonucleotide products to produce an oligonucleotide product pool; and
- 15 e. capturing or digesting the capped strands to remove them from the product pool.

The capping species can be nucleotides having a free 3'-OH group or a 3-H (dideoxy).

Nucleic acid synthesis involves the addition of a series of selected reversibly terminated building

20 blocks to a nucleic acid strand to build up a target sequence. For example, building the sequence "AATCG" requires five cycles of addition and deblocking. In the first cycle the base A must be installed, in the second A, the third T, the fourth C, and the fifth G. If any one or more of these additions fails, an incorrect sequence will be produced, for example "AATG", "ATCG", or "AAG". Typically, many cycles of synthesis are performed to make a DNA product of, for example, 50, 100,

25 or 200 nucleotides in length. With a 99% efficiency at each step, a 50-mer product is only expected to contain 60% of correct sequences, while 40% will have 1 or more errors. As the errors are largely stochastic, many sequences contain only one (31%), two errors (8%), or three (1%) errors. It is challenging to separate the produced nucleic acids as the properties of a 50 nucleotide product are very similar to those of a 49 nucleotide product with a single nucleotide deletion error present at a

30 random position in the sequence. One approach is to introduce a 'capping' step into the process. For instance, in phosphoramidite DNA synthesis, the growing strand is contacted with acetic anhydride after every addition step. Acetic anhydride reacts rapidly with the 5'-hydroxyl of the growing chemically synthesized DNA strand and prevents any further addition. The purification problem is no longer to separate a 50 nucleotide product from a 49 nucleotide product, but rather to separate a 50

nucleotide product from capped strands ranging from 1 to 49 nucleotides in length, which is feasible by methods such as liquid chromatography (LC) and polyacrylamide gel electrophoresis (PAGE). Acetic anhydride capping is largely incompatible with enzymatic DNA synthesis as the capping agent reacts rapidly with the typical solvent – water. There is thus a need for capping strategies that are compatible with de novo enzymatic nucleic acid synthesis.

Described herein are methods, kits, and compositions for the enzymatic capping of nucleic acid strands that have failed to undergo enzymatically mediated addition of a reversibly terminated nucleotide and would otherwise introduce deletion errors into a nucleic acid product. In some embodiments, the enzymatic capping technique introduces a chemical handle by which the capped products can be separated from uncapped products.

Thus, it can be said that the capping process may be blocking or non-blocking. A blocking capping process involves chain termination of the nucleic acid that failed to undergo addition. The chain terminated nucleic acid may or may not bear a chemical handle to facilitate downstream purification. A non-blocking capping process does not involve chain termination of the nucleic acid that failed to undergo addition. In this case the non-blocking capping process must involve the installation of a chemical handle that facilitates downstream purification. The downstream purification may involve affinity purification or targeted degradation.

The nucleic acid strands may be a plurality of nucleic acid strands which have been exposed to conditions for the enzymatic addition of a reversibly terminated nucleotide, a subset of which have failed to become reversibly terminated. The strands may be deoxyribonucleic acid (DNA), ribonucleic acid (RNA), or a combination of the two. The strands may have natural (e.g. phosphate) or unnatural (e.g. thiophosphate, morpholino) backbone linkages. The strands may be peptide nucleic acid or locked nucleic acid.

The nucleic acid initiator strands may be immobilized on a solid support. The solid support may be a surface, a gel, or beads. The solid support may be an inorganic, polymer, or glass surface. The solid support may be a biopolymer or synthetic polymer. The solid support may be paramagnetic beads, silica beads, or a resin.

The nucleic acid strands may include a cleavage site to enable cleavage from the solid support. The cleavage site may be a base or base sequence recognisable by an enzyme. A base recognised by an

enzyme, such as a glycosylase, may be removed to generate an abasic site which may be cleaved by chemical or enzymatic means. An example of such a glycosylase system includes the presence of a uracil base in the initiator sequence, which may be excised with 5 uracil DNA glycosylase (UDG) to leave an abasic site which may be cleaved with, for example, basic solutions, organic amines, or an endonuclease (such as endonuclease VIII), to release a nucleic acid bearing a 5'-phosphate into solution. A base sequence may be recognised and cleaved by a restriction enzyme.

The reversibly terminated nucleotide species may be deoxyribonucleotide (dNTP) species or ribonucleotide (rNTP) species. Furthermore, the nucleotide species may be modified on the sugar moiety or the nitrogenous base moiety. The modification may include a reversible terminator moiety on the sugar or nitrogenous base. The modification may include a reversible terminator moiety at the 3' position of the sugar.

The capping species may be deoxyribonucleotide (dNTP) species or ribonucleotide (rNTP) species. The nucleotide species may have a 3'-hydroxyl. The nucleotide species may be a dideoxynucleotide. The nucleotide species may be reversibly or permanently terminated. Furthermore, the nucleotide species may be modified on the sugar moiety or the nitrogenous base moiety. The modification may include a chemical handle that enables separation of capped nucleic acids from uncapped nucleic acids. Alternatively the nucleotide species may be a non-canonical nucleotide having a base such as uracil or 8-oxoguanine, such that the nucleotide allows specific strand cleavage using a suitable enzyme. The enzymes may include uracil DNA glycosylase. In an alternate embodiment, the non-canonical base is 8-oxoguanine, which is removed by formamidopyrimidine DNA glycosylase.

The enzymatic capping may be performed by a polymerase enzyme. Furthermore, the enzymatic capping may be performed by a template-independent polymerase. Template-independent polymerases may include polymerases from Family A and Family X, such as DNA Polymerase Theta or Terminal Deoxynucleotidyl Transferase. The polymerase enzyme may be naturally occurring or have a modified amino acid sequence; the modifications may have been brought about through protein engineering approaches. Polymerases may have been engineered to have an increased tolerance of 3'-O modifications. Polymerases may have been engineered to have an increased tolerance of modifications that introduce a chemical handle, either on the sugar or nitrogenous base.

The chemical handle is a moiety that may be recognized by a binding partner. Furthermore, the chemical handle is a moiety that forms half of a binding pair, and the chemical handle can be either half of the binding pair. The binding pair may be selected from a biological binding pair such as biotin/streptavidin or glutathione/GST. The binding pair may be selected from a chemical binding pair such as azide/alkyne, thiol/thiol, amine/carbonyl, or oxyamine/carbonyl. The binding pair may be selected from a chemical affinity handle such as a perfluorinated alkane/perfluorinated alkane silica gel. The half of the binding pair present in the cap is herein referred to as the chemical handle. The other half of the binding pair is herein referred to as the capture partner.

- 5 To facilitate removal of nucleic acid strands capped with a chemical handle, the other half of the binding pair (i.e. the half of the binding pair that is not present on the nucleic acid strand; the capture partner) may be immobilized on a solid support. The solid support may be a surface, a gel, or beads. The solid support may be an inorganic, polymer, or glass surface. The solid support may be a biopolymer or synthetic polymer. The solid support may be paramagnetic beads, silica beads, or a resin. Contacting a plurality of nucleic acid strands, some of which are capped with a chemical handle, with an immobilised capture partner enables the binding pair to interact. Separation of the solid support from the plurality of nucleic acid strands will remove capped nucleic acids from the original pool.
- 10 To facilitate removal of nucleic acid strands capped with a non-canonical nucleotide, the product pool can be digested with one or more enzymes. Treatment allows removal of the capped strands by digestion. The enzymes can be for example uracil DNA glycosylase or formamidopyrimidine DNA glycosylase. The treatment step can be the same step used to remove the initiator oligonucleotides from the solid support, or can be a further treatment step. Once cleaved, the shortened strands contain a 3'-phosphate moiety, unlike the full length strands. Thus the cleaved strands may be removed by steps such as ligation, where a 3'-phosphate is required. Thus the removal may be a step that uses the cleaved strands as substrates for further enzymatic steps.

- 20 Separation of capped and uncapped nucleic acids from a mixed pool of nucleic acids may be performed cycle by cycle after installation of a chemical handle. Preferably, separation of capped and uncapped nucleic acids from a mixed pool of nucleic acids may be performed after the final cycle of de novo enzymatic nucleic acid synthesis. Thus, a single separation step enables removal of all capped species with a chemical handle, regardless of in which synthesis cycle they were created.

References herein to 'nucleoside triphosphates' or 'nucleotides' 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 nucleoside triphosphates that contain deoxyribose are: deoxyadenosine triphosphate (dATP), deoxyguanosine triphosphate (dGTP), deoxycytidine triphosphate (dCTP) or deoxythymidine triphosphate (dTTP). Examples of nucleoside triphosphates that 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 nucleoside triphosphates, such as naturally occurring modified nucleosides and artificial nucleosides.

Therefore, references herein to '3'-blocked nucleoside triphosphates', '3'-reversibly terminated nucleoside triphosphates', or '3'-reversibly terminated nucleotides' refer to nucleoside triphosphates (e.g., dATP, dGTP, dCTP or dTTP) which have an additional group at the 3' position which prevents further addition of nucleotides, i.e., by replacing the 3'-OH group with a protecting group. The nitrogenous base of a blocked nucleotide may be a naturally occurring base or a chemically modified base.

It will be understood that references herein to '3'-block', '3'-blocking group', '3'-protecting group', and '3'-reversible terminator' refer to the group attached to the 3' position of the nucleoside triphosphate which prevents further nucleotide addition. De novo enzymatic nucleic acid synthesis uses reversible 3'-blocking groups (3'-reversible terminators) 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. An irreversible 3'-blocking group may be a capping agent.

The 3'-reversibly terminated nucleoside 5'-triphosphate can be blocked by any chemical group that can be unmasked to reveal a 3'-OH. The 3'-blocked nucleoside triphosphate can be blocked by a 3'-O-azidomethyl, 3'-aminoxy, 3'-O-allyl group, 3'-O-cyanoethyl, 3'-O-acetyl, 3'-O-nitrate, 3'-phosphate, 3'-O-acetyl levulinic ester, 3'-O-tert butyl dimethyl silane, 3'-O-trimethyl(silyl)ethoxymethyl, 3'-O-thiocarbamate, 3'-O-methylthiocarbamate, 3'-O-dimethylthiocarbamate, 3'-O-ortho-nitrobenzyl, and 3'-O-para-nitrobenzyl. The 3'-blocked nucleoside triphosphate can be blocked by 3'-O-azidomethyl or 3'-aminoxy.

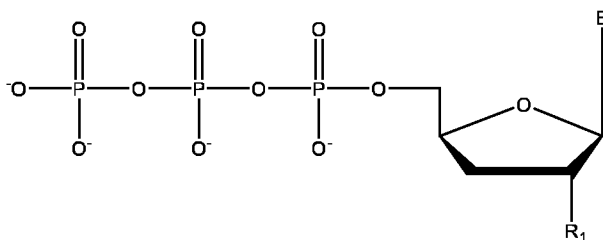
References herein to “3'-hydroxyl nucleoside triphosphates”, “3'-hydroxyl nucleotides”, or “free OH nucleotides” refer to species where the 3' position of the nucleoside triphosphate is present in an unmasked state and thus is permissible for further nucleotide addition. Examples of 3'-hydroxyl nucleotides include dATP, rGTP, biotin-16-(5-aminoallyl)-dUTP, and 5-methyl dCTP.

5

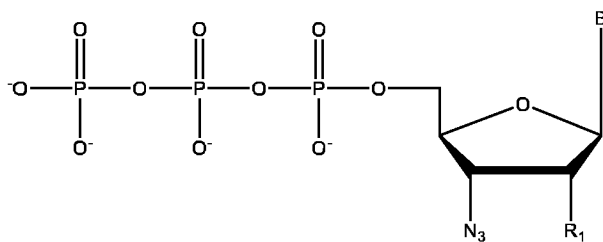
References herein to ‘capping species’ refer to a nucleotide that is employed as a capping group in de novo enzymatic nucleic acid synthesis. To meet the criteria for a capping species, a nucleotide needs to fulfil at least one of the three criteria below.

- 10 Capping criterion (1) is that the capping species prevents progression of de novo enzymatic nucleic acid synthesis (i.e. the capping species is part of a blocking capping process that causes chain termination). This criterion can be met by the capping species being a permanently terminating nucleotide or a reversibly terminated nucleotide that is orthogonal to the primary reversible terminator used in the synthesis process. A permanently terminated nucleotide is a nucleotide that
- 15 does not contain a 3'-hydroxyl or does not contain a moiety that will be readily unmasked to a 3'-hydroxyl during the synthesis process. A permanently terminated nucleotide may be a dideoxy nucleotide (e.g. compound I-A), a 3'-azido nucleotide (e.g. compound I-B), or a 3'-OMe nucleotide (e.g. compound I-C). A permanently terminated nucleotide may be an acyclic nucleotide (e.g. compound I-D).

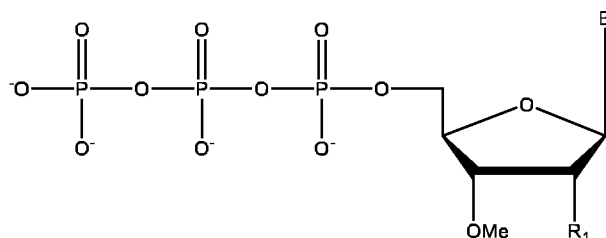
20



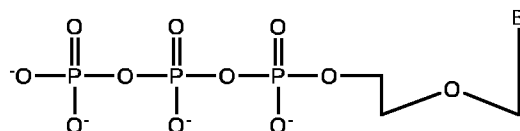
I-A



I-B



I-C



I-D

5

B is a nitrogenous base. A nitrogenous base may be a naturally occurring base such as adenine, cytosine, guanine, thymine, or uracil. A nitrogenous base may be an unnatural or modified base such as diaminopurine, 6-azidoadenine, 4-azidocytosine, 2-azidoguanine, 5-fluorocytosine, 5-hydroxybutynluracil, or 5-aminoallyl(biotin-16)uracil. Where the base is a non-canonical nucleotide such as uracil or 8-oxo-guanine, the incorporation of the non-canonical base within the strand allows digestion of the strands once fully extended, thereby removing the capped products by digestion from the product pool.

10

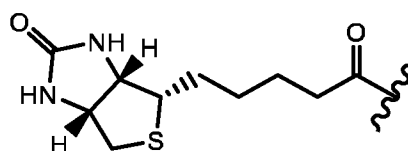
R_1 may be selected from H, OH, F, OMe, and $OCH_2CH_2OCH_3$.

15

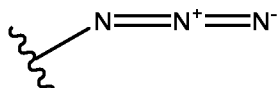
Capping criterion (2) is that the capping species installs a chemical handle that is one half of a binding pair. Capping species meeting this criterion do not necessarily prevent the capped nucleic acid strand from participating in further de novo enzymatic nucleic acid synthesis (thus this criterion can apply to both blocking and non-blocking capping processes). However, the installation of a chemical handle enables removal of strands that have experienced a failed addition. The installed chemical handle may be part of a biochemical binding pair or a chemical binding pair. The chemical handle may be present on the sugar moiety or the nitrogenous base moiety of the capping nucleotide species. The capping species may comprise glutathione, imidazole, or a peptide sequence such as a FLAG tag, Myc tag, or strep tag. The capping species may comprise a biotin (II-A), an azide (II-B), or an alkyne (II-C and II-D). The capping species may comprise a fluoruous tag, e.g. $-C_nF_{(2n+1)}$.

20

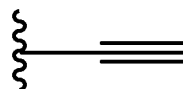
25



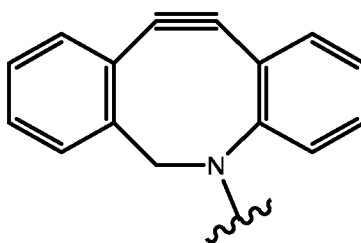
II-A



II-B



II-C



II-D

Capping criterion (3) is that the capping species installs a cleavable nucleotide within the failed strands. Thus a capping step includes a nucleotide having a free 3'OH. The nucleotide can be a non-canonical nucleotide. Where the base is a non-canonical nucleotide such as uracil or 8-oxo-guanine, the incorporation of the non-canonical base within the strand allows digestion of the strands once fully extended, thereby removing the failed capped products by digestion from the product pool. There the strands are immobilized to the support via a non-canonical nucleotide, the act of removing the strands from the support also acts to cleave the failed strands, thus aiding purification.

The term capping is used herein to signify that a modification is introduced to the failed strands allowing subsequent removal of the strands having the modification. The capped strands may in fact be longer than the desired strands, but the capped strands allow digestion such that the digestion products are shorter than the desired full-length strands. The term capping can be considered as 'marking the failed strands for subsequent removal' rather than preventing any further extensions.

In one embodiment, the polymerase used in the capping process is a terminal deoxynucleotidyl transferase (TdT). In one embodiment, the template independent polymerase is a modified TdT. In one embodiment, the TdT is modified to have improved ability to incorporate 3'-reversibly

terminated nucleotides. In another embodiment, the TdT is modified to have improved ability to incorporate permanently terminated nucleotides. In another embodiment, the TdT is modified to have improved ability to incorporate nucleotides bearing a chemical handle.

- 5 Error-containing products will bear at least one chemical handle or cleavable nucleotide and thus bind to the immobilised capture partner or be cleaved to shorter strands. This affinity purification may take place in a continuous flow system, such as in a column or a microfluidic device. The purification may take place in on a digital microfluidic device, which may employ electrowetting on dielectric phenomena to move droplets. The solid support may itself be held in place by a frit or by a
- 10 magnetic field. For example, if the capture partner is immobilised on superparamagnetic particles, a magnet or magnet array may be used to isolate the capture partner from solution, thus enabling recovery of the solution phase full-length synthesis product.

Disclosed is a method of enzymatic nucleic acid synthesis which includes:

- 15 a) providing a mixture comprising a 3'-reversibly terminated nucleotide and a template independent polymerase;
- b) providing an immobilized nucleic acid initiator sequence;
- c) adding a blocked nucleoside triphosphate to said initiator sequence through contacting the components of (a) and (b);
- 20 d) removing all addition reagents from the immobilized nucleic acid initiator sequence;
- e) capping any initiator sequences to which addition did not occur in (c), wherein the capping process involves contacting the initiator sequences with a capping species and a template independent polymerase.

25 Disclosed is a method of enzymatic nucleic acid synthesis which includes:

- a) providing a mixture comprising a reversibly terminated nucleotide and a template independent polymerase;
- b) providing an immobilized nucleic acid initiator sequence;
- c) adding a blocked nucleoside triphosphate to said initiator sequence through contacting
- 30 the components of (a) and (b);
- d) capping any initiator sequences to which addition did not occur in (c), wherein the capping process involves contacting the initiator sequences with a capping species bearing a chemical handle and a template independent polymerase;
- e) cleaving the blocking group from the 3'-reversibly terminated nucleotide in the presence
- 35 of a cleaving agent;

- f) repeating steps (c) to (e) in order to synthesise strands of a desired sequence;
- g) purifying full length synthesised strands, wherein purification involves contacting the synthesised strands with an appropriate immobilized capture partner that will interact with the capping chemical handle.

5

Specifically the invention may include a method of template independent nucleic acid synthesis comprising,

- a. providing an immobilized initiator oligonucleotide of length N bases;
 - b. extending the immobilized initiator oligonucleotides with a 3'-O reversibly terminated nucleotide species and a template independent polymerase;
 - c. extending any remaining unreacted immobilized initiator oligonucleotide of length N bases with a capping species having a moiety allowing for capture or digestion of the capped strands, wherein the capping species is a dUTP or UTP;
 - d. cleaving the 3'-O terminating group in the presence of a cleaving agent; and
- repeating steps b. to d. to add greater than one nucleotide to the above nucleic acid initiator sequence whilst capping the unextended oligonucleotide products to produce an oligonucleotide product pool; and
- e. digesting the capped strands to remove strands containing internal uracil bases from the product pool.

20

The removal/purification may comprise separation by physical properties (e.g. PAGE or chromatography). The removal may comprise affinity or covalent binding to a capture structure.

The polymerase employed for de novo enzymatic synthesis using 3'-O reversibly terminated nucleotides may be a template-independent polymerase. Template-independent polymerases may include those from Family A and Family X, such as DNA Polymerase Theta or terminal deoxynucleotidyl transferase (TdT).

The polymerase enzyme may be naturally occurring or have a modified amino acid sequence; the modifications may have been brought about through protein engineering approaches. The polymerase may have been engineered for improved properties pertaining to controlled de novo enzymatic nucleic acid synthesis. Polymerases may have been engineered to have an increased tolerance of 3'-O modified nucleotides, such as 3'-O-reversibly terminated nucleotides including 3'-aminooxy nucleotides and 3'-O-azidomethyl nucleotides. Polymerases may have been engineered to

have improved tolerance to temperature or salt. Polymerases may have been engineered for improved solubility parameters.

Therefore disclosed is a method of enzymatic nucleic acid synthesis which includes:

- 5 a) providing a mixture comprising a 3'-reversibly terminated nucleotide and a template independent polymerase;
- b) providing an immobilized nucleic acid initiator sequence;
- c) adding a blocked nucleoside triphosphate to said initiator sequence through contacting the components of (a) and (b);
- 10 d) capping any initiator sequences to which addition did not occur in (c), wherein the capping process involves contacting the initiator sequences with a capping species and a template independent polymerase.

Disclosed is a method of enzymatic nucleic acid synthesis which includes:

- 15 a) providing a mixture comprising a 3'-reversibly terminated nucleotide and a template independent polymerase;
- b) providing an immobilized nucleic acid initiator sequence;
- c) adding a blocked nucleoside triphosphate to said initiator sequence through contacting the components of (a) and (b);
- 20 d) capping any initiator sequences to which addition did not occur in (c), wherein the capping process involves adding a capping species to the composition of step (c).

Disclosed is a method of enzymatic nucleic acid synthesis which includes:

- 25 a) providing a mixture comprising a 3'-reversibly terminated nucleotide and a template independent polymerase;
- b) providing an immobilized nucleic acid initiator sequence;
- c) adding a blocked nucleoside triphosphate to said initiator sequence through contacting the components of (a) and (b);
- d) capping any initiator sequences to which addition did not occur in (c), wherein the capping process involves contacting the initiator sequences with a capping species and a template independent polymerase;
- 30 e) cleaving the blocking group from the 3'-reversibly terminated nucleoside in the presence of a cleaving agent.

Steps (c) – (e) can be repeated using nucleotides in a chosen order in order to synthesise strands of a desired sequence. Optionally, a final step may be performed to purify full length nucleic acid products. Thus, a method of enzymatic nucleic acid synthesis is disclosed which includes:

- 5 a) providing a mixture comprising a 3'-reversibly terminated nucleotide and a template independent polymerase;
- b) providing an immobilized nucleic acid initiator sequence;
- c) adding a blocked nucleoside triphosphate to said initiator sequence through contacting the components of (a) and (b);
- 10 d) capping any initiator sequences to which addition did not occur in (c), wherein the capping process involves contacting the initiator sequences with a capping species bearing a chemical handle and a template independent polymerase, which may be the same template independent polymerase provided in step (a) and used in step (c), or may be a different polymerase;
- 15 e) cleaving the blocking group from the 3'-reversibly terminated nucleotide in the presence of a cleaving agent;
- f) repeating steps (c) to (e) in order to synthesise strands of a desired sequence;
- g) purifying full length synthesised strands, wherein purification involves contacting the synthesised strands with an appropriate immobilized capture partner that will interact
- 20 with the capping chemical handle.

In step (g), a subtractive affinity purification is performed. Full length products will not bear a chemical handle and thus not bind to the immobilised capture partner. Error-containing products will bear at least one chemical handle and thus bind to the immobilised capture partner. This affinity

25 purification may take place in a continuous flow system, such as in a column or a microfluidic device. The purification may take place in on a digital microfluidic device, which may employ electrowetting on dielectric phenomena to move droplets. The solid support may itself be held in place by a frit or by a magnetic field. For example, if the capture partner is immobilised on superparamagnetic particles, a magnet or magnet array may be used to isolate the capture partner from solution, thus

30 enabling recovery of the solution phase full-length synthesis product.

WO 2019/070593 refers to a method of nucleic acid synthesis using capping species. However the method disclosed therein is the opposite of the one described herein:

[0085] Oligo synthesis yields can be enhanced by the addition of a capping step after the addition of each reversible terminator but before the cleavage step for the terminator. One method to achieve this may be conducting the extension using a dideoxy nucleotide and TdT and/or Duplase. Under the above conditions, oligos that are not extended with a reversible terminator may be truncated and may not be able to be extended in subsequent steps.

[0086] If the final base added during synthesis is modified to select for only full-length sequences or if a 3' OH is required for subsequent utilization, this capping step may provide a simple method of purification. For example, one may modify the 3' end with a biotinylated base in the final synthesis step; then full-length oligos could be captured on a streptavidin-coated surface and truncated oligos may be left behind. Another purification method may involve synthesis using a nucleotide with a 3' -group that cannot be recognized by 3' exonucleases, thus allowing the truncated oligos to be enzymatically removed.

Thus the publication discloses a method where the modification is used to select the full length sequences, not to remove the unwanted capped sequences. The applicants wish the full length sequences to remain in solution, and thus the method of adding a moiety to the full length sequences such that they are the species removed from solution is not viable.

The plurality of nucleic acid strands may be immobilised on a solid support, for example the strands may be immobilised on super paramagnetic particles; plastic, glass, or polymer surfaces; biopolymers; synthetic polymers; hydrogels; or an inorganic support. The utility of immobilising the nucleic acid strands is to facilitate the iterative process of controlled de novo enzymatic nucleic acid synthesis. Reagents and buffers may be sequentially introduced to and removed from the solid-supported nucleic acid strands.

The method can be performed on a microfluidic device such as a digital microfluidic device.

Digital microfluidics (DMF) refers to a two-dimensional planar surface platform for lab-on-a-chip systems that is based upon the manipulation of microdroplets. Droplets can be dispensed, moved, stored, mixed, reacted, or analyzed on a platform with a set of insulated electrodes. Digital microfluidics can be used together with analytical analysis procedures such as mass spectrometry, colorimetry, electrochemical, and electrochemiluminescence.

The droplet can be moved using any means of electrokinesis. The aqueous droplet can be moved using electrowetting-on-dielectric (EWoD). Electrowetting on a dielectric (EWOD) is a variant of the electrowetting phenomenon that is based on dielectric materials. During EWoD, a droplet of a conducting liquid is placed on a dielectric layer with insulating and hydrophobic properties. Upon
5 activation of the electrodes the dielectric layer becomes less hydrophobic, thus causing the droplet to spread onto the surface.

The electrical signal on the EWoD or optically-activated amorphous silicon (a-Si) EWoD device can be delivered through segmented electrodes, active-matrix thin-film transistors or digital micromirrors.
10 Optically-activated s-Si EWoD devices are well known in the art for actuating droplets (*J. Adhes. Sci. Technol.*, 2012, **26**, 1747-1771).

The oil in the device can be any water immiscible or hydrophobic liquid. The oil can be mineral oil, silicone oil, an alkyl-based solvent such as decane or dodecane, or a fluorinated oil. The air in the
15 device can be any humidified gas.

The droplets can be actuated on a hydrophobic surface on the digital microfluidic device (ACS Nano 2018, 12, 6, 6050-6058). The hydrophobic surface can be a hydrophobic surface such as polytetrafluoroethylene (PTFE), Teflon AF (DuPont Inc), CYTOP (AGC Chemicals Inc), or FluoroPel
20 (Cytonix LLC). The hydrophobic surface may be modified in such a way to reduce biofouling, especially biofouling resulting from exposure to CFPS reagents or nucleic acid reagents. The hydrophobic surface may also be superhydrophobic, such as NeverWet (NeverWet LLC) or Ultra-Ever Dry (Flotech Performance Systems Ltd). Superhydrophobic surfaces prevent biofouling compared with typical fluorocarbon-based hydrophobic surfaces. Superhydrophobic surfaces thus prolong the
25 capability of digital microfluidic devices to move CFPS droplets and general solutions containing biopolymers (RSC Adv., 2017, 7, 49633-49648). The hydrophobic surface can also be a slippery liquid infused porous surface (SLIPS), which can be formed by infusing Krtox-103 oil (DuPont) with porous PTFE film (Lab Chip, 2019, 19, 2275).

30 In one embodiment, the polymerase used for 3'-reversibly terminated nucleotide incorporation is a terminal deoxynucleotidyl transferase (TdT). In one embodiment, the template independent polymerase is a modified TdT. The capping moiety can be added by the same polymerase, or a further polymerase can be used.

References herein to an 'initiator sequence' refer to a short oligonucleotide with a free 3'-end which the 3'-reversibly terminated nucleotide monomers can attach. In one embodiment, the initiator sequence is a DNA initiator sequence. In an alternative embodiment, the initiator sequence is an RNA initiator sequence.

5

References herein to a 'DNA initiator sequence' include a small sequence of DNA which a 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 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.

10

In one embodiment, the initiator sequence is immobilised on a solid support. 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.

15

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.

20

In one embodiment, the initiator sequence contains a base or base sequence recognisable by an enzyme. A base recognised by an enzyme, such as a glycosylase, may be removed to generate an abasic site which may be cleaved by chemical or enzymatic means. A base sequence may be recognised and cleaved by a restriction enzyme.

25

In the methods described herein, the synthesised strands may be cleaved prior to the incomplete terminated (capped) strands being removed. Thus the immobilised initiator strands, having undergone extension to form the oligonucleotide product pool, are released from their solid support. The undesired products having the capping species may then be removed from the product pool, for example using capturing onto a further support.

30

Thus in one embodiment, the resultant contiguous oligonucleotide sequence is released from being immobilised. In one embodiment, this release occurs by removing a non-canonical base from one or more of the immobilised oligonucleotides and cleaving the strands at the resultant abasic site. In

35

one embodiment, the non-canonical base is uracil, which is removed by uracil DNA glycosylase. In an alternate embodiment, the non-canonical base is 8-oxoguanine, which is removed by formamidopyrimidine DNA glycosylase.

- 5 In an alternate 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 contiguous oligonucleotide sequence by cleaving the chemical linker through the addition of tris(2-carboxyethyl)phosphine (TCEP) or dithiothreitol (DTT) for a disulfide linker; palladium complexes or
10 an allyl linker; or TCEP for an azide-masked hemiaminal ether linker.

- In one embodiment, the template independent polymerase is a terminal deoxynucleotidyl transferase (TdT). In one embodiment, the template independent polymerase is a modified TdT. In one embodiment, the TdT is added in the presence of an extension solution comprising one or more
15 buffers (e.g., Tris or cacodylate), one or more salts (e.g., Na^+ , K^+ , Mg^{2+} , Mn^{2+} , Cu^{2+} , Zn^{2+} , Co^{2+} , 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 nucleoside triphosphate hydrolysis by TdT. Therefore, the use of
20 an inorganic pyrophosphatase has the advantage of reducing the rate of (1) backwards reaction and (2) TdT strand dismutation.

- References herein to 'cleaving agent' refer either to a substance which is able to cleave the 3'-blocking group from the 3'-blocked nucleoside or to a substance which is able to cleave immobilised
25 oligonucleotides from the solid support. In one embodiment, the cleaving agent is a chemical cleaving agent. In an alternative 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'-nucleoside blocking group used. For instance, tris(2-carboxyethyl)phosphine (TCEP) or tris(hydroxypropyl)phosphine (THPP) can be used to cleave a 3'-O-
30 azidomethyl group, palladium 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 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.

The inventors have previously developed a selection of engineered terminal transferase enzymes, any of which may be used in the current process.

The term modified terminal transferase enzyme refers to a modified terminal deoxynucleotidyl transferase (TdT) enzymes or the homologous amino acid sequence of Pol μ , Pol β , Pol λ , and Pol θ of any species or the homologous amino acid sequence of X family polymerases of any species. Terminal transferase enzymes are ubiquitous in nature and are present in many species. Many known TdT sequences have been reported in the NCBI database <http://www.ncbi.nlm.nih.gov/>.

The inventors have modified the terminal transferase from *Lepisosteus oculatus* TdT (spotted gar) (shown below). However the corresponding modifications can be introduced into the analogous terminal transferase sequences from any other species, including the sequences listed above in the various NCBI entries, including truncated versions thereof.

The amino acid sequence of the spotted gar (*Lepisosteus oculatus*) is shown below (SEQ ID 1)

MLHIPIFPPIKKRQKLPEsrNSCKYEVKFSEVAIFLVERKMGSSRRKFLTNLARSKGFRIEDVLSDAVTHVVAEDNSAD
ELWQWLQNSSLGDLskIEVLDisWFTECMGAGKPVQVEARHCLVKSCPVIDQYLEPSTVETVSQYACQRRTTMEN
HNQIFTDAFAILAENAEFNESEGPClAFMRAASLLKSLPHAISSSKDLEGLPCLGDQTKAVIEDILEYGQCSKVQDVLC
DDRYQTIKLFTSVFGVGLKTAEKWYRKGfHSLEEVQADNAIHFTKMqKAGFLYYDDISAAVCKAEaQAIGQIVEET
VRLIAPDAIVTLTGGFRRGKECGHDVDFLITTPeMGKEVWLLNRLINRLQNQGILLYDIVESTFDKTRLPCRKFAM
DHFQKCFaIIKLKkELAAGRvQKDWKAIRVDFVAPPVDNFAfallGWTGSRQFERDLRRfARHERKMllLDNHAlY
DKTKKIFLPAKTEEDIFaHLGLDYIDPWQRNA

The inventors have identified various amino acids modifications in the amino acid sequence having improved properties. Certain regions improve the solubility and handling of the enzyme. Certain

other regions improve the ability to incorporate nucleotides with modifications; these modifications include modifications at the 3'-position of the sugar and modifications to the base.

Described herein are modified terminal deoxynucleotidyl transferase (TdT) enzymes comprising amino acid modifications when compared to a wild type sequence SEQ ID NO 1 or a truncated version thereof or the homologous amino acid sequence of a terminal deoxynucleotidyl transferase (TdT) enzyme in other species or the homologous amino acid sequence of Pol μ , Pol β , Pol λ , and Pol θ of any species or the homologous amino acid sequence of X family polymerases of any species, wherein the amino acid is modified at one or more of the amino acids:

V32, A33, I34, F35, A53, V68, V71, E97, I101, M108, G109, A110, Q115, V116, S125, T137, Q143, M152, E153, N154, H155, N156, Q157, I158, I165, N169, N173, S175, E176, G177, P178, C179, L180, A181, F182, M183, R184, A185, L188, H194, A195, I196, S197, S198, S199, K200, E203, G204, D210, Q211, T212, K213, A214, I216, E217, D218, L220, Y222, V228, D230, Q238, T239, L242, L251, K260, G261, F262, H263, S264, L265, E267, Q269, A270, D271, N272, A273, H275, F276, T277, K278, M279, Q280, K281, S291, A292, A293, V294, C295, K296, E298, A299, Q300, A301, Q304, I305, T309, V310, R311, L312, I313, A314, I318, V319, T320, G328, K329, E330, C331, L338, T341, P342, E343, M344, G345, K346, W349, L350, L351, N352, R353, L354, I355, N356, R357, L358, Q359, N360, Q361, G362, I363, L364, L365, Y366, Y367, D368, I369, V370, K376, T377, C381, K383, D388, H389, F390, Q391, K392, F394, I397, K398, K400, K401, E402, L403, A404, A405, G406, R407, D411, A421, P422, P423, V424, D425, N426, F427, A430, R438, F447, A448, R449, H450, E451, R452, K453, M454, L455, L456, D457, N458, H459, A460, L461, Y462, D463, K464, T465, K466, K467, T474, D477, D485, Y486, I487, D488, P489.

Modifications which improve the incorporation of modified nucleotides can be at one or more of selected regions shown below. Regions were selected according to mutation data, sequence alignment, and structural data obtained from spotted gar TdT co-crystallized with DNA and a 3'-modified dNTP. The second modification can be selected from one or more of the amino acid regions VAIF, MGA, MENHNQI, SEGPCLAFMRA, HAISSS, DQTKA, KGFHS, QADNA, HFTKMQK, SAAVCK, EAQA, TVRLI, GKEC, TPEMGK, YYDIV, DHFQK, LAAG, APPVDNF, FARHERKMLLDNHALYDKTKK, and DYIDP shown highlighted in the sequence below.

MLHIPIFPPIKKRQKLPEsrNSCKYEVKFSEVAIFLVERKMGSSRRKFLTNLARSKGFRIEDVLSDAV
THVVAEDNSADELWQWLQNSSLGDLISKIEVLDISWFTECMGAGKPVQVEARHCLVKSCPVIDQYLEPS
TVETVSQYACQRRTTMENHNQIFTDAFAILAENAEFNESEGPCLAFMRAASLLKSLPHAISSSKDLEG

LPCLGDDTKAVIEDILEYGQCSKVQDVLCDTRYQTIKLFVSFVGLKTAEKWYRKGFHSLLEEVQADN
 AIFTKMQKAGFLYYDDISAAVCKAEQAIGQIVEETVRLIAPDAIVTLTGGFRRGKECGHDVDFLIT
 TPENMGKEVWLLNRLINRLQNQGILLYYDIVESTFDKTRLPCRKFAMDHQKCFAIKLLKELAAAGRV
 QKDWKAIRVDFVAPPVDNFAFALLGWTGSRQFERDLRRFARHERKMLLDNHALYDKTKKIFLPAKTEE
 5 DIFAHLGLDYIDPWQRNA

References to particular sequences include truncations thereof. Included herein are modified
 terminal deoxynucleotidyl transferase (TdT) enzyme comprising at least one amino acid modification
 when compared to a wild type sequence SEQ ID NO 1 or a truncated version thereof, or the
 10 homologous amino acid sequence of a terminal deoxynucleotidyl transferase (TdT) enzyme in other
 species, wherein the modification is selected from one or more of the amino acid regions
 WLLNRLINRLQNQGILLYYDIV, VAIF, MGA, MENHNQI, SEGPCLAFMRA, HAISSS, DQTKA, KGFHS,
 QADNA, HFTKMQK, SAAVCK, EAQA, TVRLI, GKEC, TPENMGK, DHFQK, LAAG, APPVDNF,
 FARHERKMLLDNHALYDKTKK, and DYIDP of the sequence of SEQ ID NO 1 or the homologous regions
 15 in other species.

Truncated proteins may include at least the region shown below

TVSQYACQRRRTTMENHNQIFTDFAFILAENAEFNESEGPCLAFMRAASLLKSLPHAISSSKDLEGLPC
 20 LGDQTKAVIEDILEYGQCSKVQDVLCDTRYQTIKLFVSFVGLKTAEKWYRKGFHSLLEEVQADNAIH
 FTKMQKAGFLYYDDISAAVCKAEQAIGQIVEETVRLIAPDAIVTLTGGFRRGKECGHDVDFLITTP
 MGKEVWLLNRLINRLQNQGILLYYDIVESTFDKTRLPCRKFAMDHQKCFAIKLLKELAAAGRVQKD
 WKAIRVDFVAPPVDNFAFALLGWTGSRQFERDLRRFARHERKMLLDNHALYDKTKKIFLPAKTEEDIF
 AHLGLDYIDPWQRNA

Described herein is a modified terminal deoxynucleotidyl transferase (TdT) enzyme comprising at
 least the sequence:

TVSQYACQRRRTTMENHNQIFTDFAFILAENAEFNESEGPCLAFMRAASLLKSLPHAISSSKDLEGLPC
 30 LGDQTKAVIEDILEYGQCSKVQDVLCDTRYQTIKLFVSFVGLKTAEKWYRKGFHSLLEEVQADNAIH
 FTKMQKAGFLYYDDISAAVCKAEQAIGQIVEETVRLIAPDAIVTLTGGFRRGKECGHDVDFLITTP
 MGKEVWLLNRLINRLQNQGILLYYDIVESTFDKTRLPCRKFAMDHQKCFAIKLLKELAAAGRVQKD
 WKAIRVDFVAPPVDNFAFALLGWTGSRQFERDLRRFARHERKMLLDNHALYDKTKKIFLPAKTEEDIF
 AHLGLDYIDPWQRNA

or the homologous regions in other species, wherein the sequence has one or more amino acid modifications in one or more of the amino acid regions WLLNRLINRLQNQGILLYDI, MENHNQI, SEGPCLAFMRA, HAISSS, DQTKA, KGFHS, QADNA, HFTKMQK, SAAVCK, EAQA, TVRLI, GKEC, TPMEGK, DHFQK, LAAG, APPVDNF, FARHERKMILLDNHALYDKTKK, and DYIDP of the sequence.

5

TVSQYACQRRTTMENHNQIFTDFAFAILAENAEFNESEGPCLAFMRAASLLKSLPHAISSSKDLEGLPC
 LGDQTKAVIEDILEYGQCSKVQDVLCDTRYQTIKLFSTVFGVGLKTAEKWYRKGFHSLLEEVQADNAIH
 FTKMOKAGFLYYDDISAAVCKAEQAIGQIVEETVRLIAPDAIVTLTGGRGKFCGHDVDFLITTE
 MGKEVWLLNRLINRLQNQGILLYDIVESTFDKTRLPCRKFEAMDHFQKCFAIKLLKKELAAGRVQKD
 10 WKAIRVDFVAPPVDNFAFALLGWTGSRQFERDLRRFARHERKMILLDNHALYDKTKKIFLPAKTEEDIF
 AHLGLDYIDPWQRNA

Sequence homology extends to all modified or wild-type members of family X polymerases, such as DNA Pol μ (also known as DNA polymerase mu or POLM), DNA Pol β (also known as DNA polymerase
 15 beta or POLB), and DNA Pol λ (also known as DNA polymerase lambda or POLL). It is well known in the art that all family X member polymerases, of which TdT is a member, either have terminal transferase activity or can be engineered to gain terminal transferase activity akin to terminal deoxynucleotidyl transferase (Biochim Biophys Acta. 2010 May; 1804(5): 1136–1150). For example, when the following human TdT loop1 amino acid sequence

20

...ESTFEKLRRLPSRKVDALDHF...

was engineered to replace the following human Pol μ amino acid residues

25

...HSCCESPTRLAQQSHMDAF...,

the chimeric human Pol μ containing human TdT loop1 gained robust terminal transferase activity (Nucleic Acids Res. 2006 Sep; 34(16): 4572–4582).

30

Furthermore, it was generally demonstrated in US patent application no. 2019/0078065 that family X polymerases when engineered to contain TdT loop1 chimeras could gain robust terminal transferase activity. Additionally, it was demonstrated that TdT could be converted into a template-dependent polymerase through specific mutations in the loop1 motif (Nucleic Acids Research, Jun 2009, 37(14):4642-4656). As it has been shown in the art, family X polymerases can be trivially modified to either display template-dependent or template-independent nucleotidyl transferase activities.

35

Therefore, all motifs, regions, and mutations demonstrated in this patent can be trivially extended to

modified X family polymerases to enable modified X family polymerases to incorporate 3'-modified nucleotides, reversibly terminated nucleotides, and modified nucleotides in general to effect methods of nucleic acid synthesis.

- 5 Disclosed is a kit for performing the method. The kit may comprise comprising 3'-reversibly terminated nucleotides, a modified terminal transferase enzyme and a capping nucleotide species having a moiety allowing for capture of the capped strands. The kit may further comprise a solid support for capturing the capped strands.

10 EXAMPLES

Blocking capping process

- An example of a blocking capping process is shown in Figure 2. In step (A), an initiator sequence is contacted with reversibly terminated nucleoside triphosphates and a template independent polymerase. The template independent polymerase effects a transformation of the length N initiator
- 15 to length N+1 initiator. Due to one or more reaction conditions (e.g. time, temperature, ionic strength, divalent metal, buffer identity, pH) or enzyme properties (e.g. intolerance of reversible terminator or substrate sequence context) some initiator sequences may not convert to length N+1, and thus will remain length N. In step (B), all initiator sequences are exposed to capping nucleotides and a template independent polymerase. The template independent polymerase facilitates addition
- 20 of the capping species to any available (i.e. not reversibly terminated) 3' termini. Step (C) shows an example initiator pool following (A) addition and (B) capping. Step (D) describes treating the initiator sequences with a deblocking solution that transforms the reversible terminator moiety (e.g. an azidomethyl or aminooxy moiety with TCEP or nitrite respectively) into a 3'-hydroxyl moiety. Steps (A) to (D) can be repeated in a de novo enzymatic nucleic acid synthesis process. Step (E) shows an
- 25 example product pool following six cycles of de novo enzymatic nucleic acid synthesis. One strand was capped in the first cycle, one strand was capped in the fourth cycle, and one strand underwent six successful cycles and is thus full length. Each capped strand has a synthesised portion as long as the cycle in which it was capped. Exposure of the pool in (E) to a capture structure, which may be immobilised, leads to a chemical or affinity interaction of the chemical handle present in the
- 30 blocking capping moiety. Unbound, full length product can then be removed and is shown purified in step (F).

Non-blocking capping process

An example of a non-blocking capping process is shown in Figure 3. In step (A), an initiator sequence is contacted with reversibly terminated nucleoside triphosphates and a template independent polymerase. The template independent polymerase effects a transformation of the length N initiator to length N+1 initiator. Due to one or more reaction conditions (e.g. time, temperature, ionic strength, divalent metal, buffer identity, pH) or enzyme properties (e.g. intolerance of reversible terminator or substrate sequence context) some initiator sequences may not convert to length N+1, and thus will remain length N. In step (B), all initiator sequences are exposed to capping nucleotides and a template independent polymerase. The template independent polymerase facilitates addition of the capping species to any available (i.e. not reversibly terminated) 3' termini. Step (C) shows an example initiator pool following (A) addition and (B) capping. Step (D) describes treating the initiator sequences with a deblocking solution that transforms the reversible terminator moiety (e.g. an azidomethyl or aminooxy moiety with TCEP or nitrite respectively) into a 3'-hydroxyl moiety. Steps (A) to (D) can be repeated in a de novo enzymatic nucleic acid synthesis process. Step (E) shows an example product pool following six cycles of de novo enzymatic nucleic acid synthesis. One strand was capped in the first cycle, one strand was capped in the fourth cycle, and one strand underwent six successful cycles and is thus full length. As a non-blocking capping process was employed, the length of the strands may be equal at the end of the process. Additionally, one strand may bear multiple chemical handles, for example two chemical handles could be installed in one step or one chemical handle could be installed in each of two distinct steps. Exposure of the pool in (E) to a capture structure, which may be immobilised, leads to a chemical or affinity interaction of the chemical handle present in the blocking capping moiety. Unbound, full length product can then be removed and is shown purified in step (F).

Non-canonical nucleotide capping process

An example of a non-canonical nucleotide capping process is shown in Figure 4. In step (A), an initiator sequence is contacted with reversibly terminated nucleoside triphosphates and a template independent polymerase. The template independent polymerase effects a transformation of the length N initiator to length N+1 initiator. Due to one or more reaction conditions (e.g. time, temperature, ionic strength, divalent metal, buffer identity, pH) or enzyme properties (e.g. intolerance of reversible terminator or substrate sequence context) some initiator sequences may not convert to length N+1, and thus will remain length N. In step (B), all initiator sequences are exposed to non-canonical nucleotides and a template independent polymerase. The template independent polymerase facilitates addition of the non-canonical nucleotide species to any available (i.e. not reversibly terminated) 3' termini. Step (C) shows an example initiator pool following (A)

addition and (B) capping. Some initiators that failed to undergo addition have one non-canonical nucleotide on their 3' terminus while others may have greater than one non-canonical nucleotide on their 3' terminus. Step (D) describes treating the initiator sequences with a deblocking solution that transforms the reversible terminator moiety (e.g. an azidomethyl or aminooxy moiety with TCEP or nitrite respectively) into a 3'-hydroxyl moiety. Steps (A) to (D) can be repeated in a de novo enzymatic nucleic acid synthesis process. Step (E) shows an example product pool following six cycles of de novo enzymatic nucleic acid synthesis. Two strands were capped in the first cycle (one with a single non-canonical species and one with multiple), one strand was capped in the fourth cycle, and one strand underwent six successful cycles and is thus full length. As a non-blocking capping process was employed, the length of the capped strands may be equal to or greater than correct strands at the end of the process. Additionally, one strand may bear multiple non-canonical handles, for example two chemical handles could be installed in one step or one chemical handle could be installed in each of two distinct steps. Exposure of the pool in (E) to a condition that digests the non-canonical nucleotides (e.g. a glycosylase enzyme) converts any strand containing a non-canonical nucleotide to a cleaved species with a blocked 3' terminus (e.g. a 3'-phosphate). Such digested species are not functional in downstream applications such as further extension or ligation. As such, the correct full-length species are purified from those which experienced a failed addition. Example pairs of non-canonical nucleotides and digestion conditions are (i) 2'-deoxyuridine 5'-triphosphate (dUTP) and uracil DNA glycosylase (UDG), and (ii) 8-Oxo-2'-deoxyguanosine 5'-triphosphate (8oxo-dGTP) and formamidopyrimidine DNA Glycosylase (FPG).

Experimental method and data demonstrating capping with a non-canonical nucleotide followed by digestion

This example demonstrates that dUTP tailing can be used as an enzymatic DNA synthesis capping strategy to remove unreacted 3'-OH termini of extending eDNA molecules during enzymatic synthesis, as illustrated in the schematic diagram shown in Figure 1.

Method

All subsequent steps were automated on a ThermoFisher Kingfisher Flex platform using custom scripts unless otherwise stated.

Enzymatic N5mer synthesis (Step 1 and 2, Figure 1)

Super paramagnetic particles (oligo-magbeads) bearing oligonucleotide initiators were prepared as previously described, using oligo D649 as initiator (Table 1). This oligo contains a 5' portion

compatible with Illumina TruSeq library preparation methods (labelled as P5, Figure 1) and allows the amplification product to be sequenced by Illumina SBS.

Table 1. Oligos used in this experiment

Note that * denotes an internal phosphorothioate bond

Oligo ID	Sequence (5' to 3')
D649	T*T*T*TTTTTTTUGAGTCCTCAGCACACGACGCTCTTCCGATCTAGAAGACTT
D412	AATGATACGGCGACCACCGAGATCTACACCACAGACTACACTCTTCCCTACACGACGCTCTTCCGATC*T
D413	AATGATACGGCGACCACCGAGATCTACACCGACACTTACACTCTTCCCTACACGACGCTCTTCCGATC*T
D470	CAAGCAGAAGACGGCATACGAGATAGACATGCGTGACTGGAGTTCAGACGTG*T
D475	CAAGCAGAAGACGGCATACGAGATGCCTTCTTGTGACTGGAGTTCAGACGTG*T

- 5 60 micrograms of oligo-magbeads were used per reaction and this formed the initiator substrate upon which the N5mer oligos were written. Five cycles of automated enzymatic DNA synthesis reactions were conducted, cycling through the steps detailed in Table 2, to write the oligo sequence TAGCG onto the end of the oligo-magbeads.

Table 2. Enzymatic synthesis reaction cycle conditions

Step	Details	Time (mins)
1	Addition (250 μ M dNTP-ONH ₂ , 2 μ M TdT, reaction buffer pH 7.2)	3
2	Wash Buffer 1	2
3	Wash Buffer 2	2
4	Deblock (sodium nitrite, acetate buffer pH 5.5)	5
5	Wash Buffer 1	2
6	Wash 2 buffer	2

7	Repeat steps 1-4 four more times
---	----------------------------------

The 3' terminus of the N5mer at the end of this process was 3'-OH (i.e. recovered after Step 2, Figure 1).

5 PolyU capping (Step 3, Figure 1)

Following synthesis, the N5mer extended oligo-magbead was tailed with dUTP by incubating in a PolyU-capping reaction, Table 3:

Table 3. PolyU-capping reaction conditions		
Step	Details	Time (mins)
1	PolyU-capping mix (250 μ M dUTP, 2 μ M TdT, reaction buffer pH 7.2)	1
2	Wash Buffer 1	2
3	Wash Buffer 2	2

- 10 The tailing reaction was stopped by removing the oligo-magbeads from the PolyU-capping reaction and washing in Wash Buffer 1.

dCTP-ONH₂ addition (Step 4, Figure 1)

- 15 PolyU-tailed N5mer oligo magbeads were split in half (30 micrograms each) and each half subjected to a subsequent round of enzymatic DNA synthesis, of either i) mock conditions (Table 4, dCTP-ONH₂ omitted from the addition mix) or ii) full reaction conditions (as shown in Table 4).

Table 4. dCTP-ONH ₂ enzymatic synthesis reaction conditions		
Step	Details	Time (mins)
1	Addition (250 μ M dCTP-ONH ₂ , 2 μ M TdT, reaction buffer pH 7.2)	3
2	Wash Buffer 1	2

3	Wash Buffer 2	2
4	Deblock (sodium nitrite, acetate buffer pH 5.5)	5
5	Wash Buffer 1	2
6	Wash 2 buffer	2

The product of this reaction was either N5mer (mock) or N6mer (full) polyU-capped oligo-magbeads.

PolyU-capped tail digestion (Step 5, Figure 1).

- 5 Each 30 microgram oligo-magbead sample from the previous step was split again into half (15 micrograms) with each half subjected to either i) mock UDG digestion (Table 5, no UDG) or ii) full UDG digestion (as shown in Table 5).

Table 5. PolyU-capping tail digestion reaction conditions		
Step	Details	Time (mins)
1	Digestion mix (1U/20 μ L USER reagent in reaction buffer pH 7.9)	90
2	Wash Buffer 1	2

- 10 The final sample processing conditions are shown in Table 6.

Table 6. Sample processing summary				
Sample	Treatment			
	N5 synth	PolyU-cap	N6 dCTP synth	UDG digest
SR138_C10	+	+	-	+
SR138_D10	+	+	-	-
SR138_C05	+	+	+	+

SR138_D05	+	+	+	-
-----------	---	---	---	---

NGS library prep and index PCR (Step 6, Figure 1).

Samples SR138_D10, C10, D05 and C05 were converted into Illumina SBS-compatible sequencing libraries by appending TruSeq style adapters, and then dual indexing each sample uniquely by index

5 PCR (Tables 7A and B).

Table 7A. Indexing PCR reactions

Sample	PCR_F	PCR_R	PCR mix	Cycles
SR138_D10	D413	D475	25 pmol PCR_F 25 pmol PCR_R 1X Thermo Phire HS 2X PCR mastermix 5 µL sample 2U Q5U DNA polymerase	20 per reaction
SR138_C10	D412	D475		
SR138_D05	D413	D470		
SR138_C05	D412	D470		

Table 7B: Indexing PCR thermocycling conditions

Step	Description	Temperature (°C)	Time (sec)
1	Initial denaturation	95	120
2	Denaturation	95	20
3	Annealing	65	20
4	Extension	72	20
5	Repeat steps 2-5	19 cycles	
6	Final extension	72	60
7	Hold	10	infinite

Results

Dual indexed samples were pooled and sequenced on an Illumina MiniSeq (Mid Output kit, 101 cycle read 1, 8 cycle i5 and i7 index reads) according to the manufacturer's specification. Basecalling and demultiplexing was performed with bcl2fastq. Fastq files for each of the samples were analysed using regular expressions to yield the results shown in Table 8.

5

Table 8. NGS read count summary				
Metric	Sample			
	C10	D10	C05	D05
PolyU-cap	+	+	+	+
N6 dCTP synth	-	-	+	+
UDG digest	+	-	+	-
Total number of reads containing initiator sequence	1315	34898	808	36902
Percent of N5 reads with dU tail	0.08	0.53	0.05	0.05
Percent of N5 reads with dU tail and terminal dC	0	0	0.03	0.48
Percent of correct sequence N6 reads	0	0	0.29	0.25

Conclusion

Figure 6 and Table 8 demonstrate the benefit of polyU-capping as a method for removing 3'OH unreacted oligo termini during enzymatic DNA synthesis. Comparing samples C10 and D10 it is evident that the UDG digest treatment has dramatically reduced the presence of oligonucleotides with a dU tail in the pool, from 53% of the product pool to 8% of the product pool. Comparing samples C05 and D05 it is evident that the UDG digest treatment has dramatically reduced the presence of oligonucleotides with a dU tail in the pool, from 48% of the product pool to 3% of the product pool. It can also be seen that the inclusion of the polyU-capping method increases the percent of correct sequences (e.g. TAGCGC) by 16% (from 0.25 without treatment to 0.29 with treatment) in the C05/D05 pair.

10

15

Claims

1. A method of template independent nucleic acid synthesis comprising,
 - a. providing an immobilized initiator oligonucleotide of length N bases;
 - 5 b. extending the immobilized initiator oligonucleotides with a 3'-O reversibly terminated nucleotide species and a template independent polymerase;
 - c. extending any remaining unreacted immobilized initiator oligonucleotide of length N bases with a capping species having a moiety allowing for capture or digestion of the capped strands, wherein the capping species is a nucleotide;
 - 10 d. cleaving the 3'-O terminating group in the presence of a cleaving agent; and repeating steps b. to d. to add greater than one nucleotide to the above nucleic acid initiator sequence whilst capping the unextended oligonucleotide products to produce an oligonucleotide product pool; and
 - e. capturing or digesting the capped strands to remove them from the product pool.
- 15 2. The method according to claim 1 wherein the 3'-reversibly terminated nucleotide is one of 3'-O-azidomethyl, 3'-aminooxy, 3'-O-allyl group, 3'-phosphate, 3'-phosphomonoester, 3'-phosphodiester, 3'-O-cyanoethyl, 3'-O-acetyl, 3'-O-nitrate, 3'-O-acetyl levulinic ester, 3'-O-thiocarbamate, 3'-O-methylthiocarbamate, 3'-O-dimethylthiocarbamate, 3'-O-tert butyl dimethyl silane, 3'-O-trimethyl(silyl)ethoxymethyl, 3'-O-ortho-nitrobenzyl, and 3'-O-para-nitrobenzyl.
- 20 3. The method according to claim 1 wherein the 3'-reversibly terminated nucleotide is 3'-O-azidomethyl, 3'-phosphate, or 3'-aminooxy nucleotides.
4. The method according to any one of claims 1 to 3 wherein the template independent polymerase is a terminal deoxynucleotidyl transferase.
- 25 5. The method according to claim 4 wherein the terminal deoxynucleotidyl transferase has been modified from a naturally occurring sequence.
6. The method according to claim 5 wherein the capping nucleotide has a free 3'-OH or is a 3'-deoxynucleotide or has a non-reversible 3'-terminator.
7. The method of claim 6 wherein the capping nucleotide has a free 3'-OH.
- 30 8. The method according to any one of claims 1 to 7 wherein the capping nucleotide is a non-canonical nucleotide.
9. The method according to claim 8 wherein capping nucleotide base is uracil or 8-oxo-guanine.
10. The method according to claim 8 or claim 9 wherein the product pool is treated with an enzyme to digest the capped strands.

11. The method according to claim 10 wherein the digestion also releases the immobilised oligonucleotide product pool.
12. The method according to any one of claims 1 to 9 wherein the immobilised oligonucleotide product pool is released prior to capturing the capped strands.
- 5 13. The method according to one of claims 1 to 12 wherein the moiety for capture is a small molecule such as biotin for protein capture or a fluorous tag for fluorous capture.
14. The method according to claim 13 wherein the moiety for capture is biotin.
15. The method according to any one of claims 1 to 12 wherein the moiety for capture is a moiety for chemical capture such as a thiol, an azide or an alkyne.
- 10 16. The method according to claim 15 wherein the moiety for capture is an azide or an alkyne.
17. The method according to any one of claims 1-16 wherein the method is performed on a microfluidic device.
18. The method of claim 17 wherein the device is a digital microfluidic device.
19. The method according to claim 1 comprising
 - 15 a. providing an immobilized initiator oligonucleotide of length N bases;
 - b. extending the immobilized initiator oligonucleotides with a 3'-O reversibly terminated nucleotide species and a template independent polymerase;
 - c. extending any remaining unreacted immobilized initiator oligonucleotide of length N bases with a capping species having a moiety allowing for capture or digestion of the
 - 20 capped strands, wherein the capping species is a dUTP or UTP;
 - d. cleaving the 3'-O terminating group in the presence of a cleaving agent; and repeating steps b. to d. to add greater than one nucleotide to the above nucleic acid initiator sequence whilst capping the unextended oligonucleotide products to produce an oligonucleotide product pool; and
 - 25 e. digesting the capped strands to remove strands containing internal uracil bases from the product pool.
20. A kit for use in a method as defined in any one of claims 1-19, the kit comprising 3'-reversibly terminated nucleotides, a modified terminal transferase enzyme and a capping nucleotide species having a free 3'-OH or 3'-H and a moiety allowing for capture of the capped strands.
- 30 21. The kit according to claim 20 further comprising a solid support for capturing the capped strands.
22. The kit according to claim 20 or 21 wherein the capping nucleotide is dUTP.

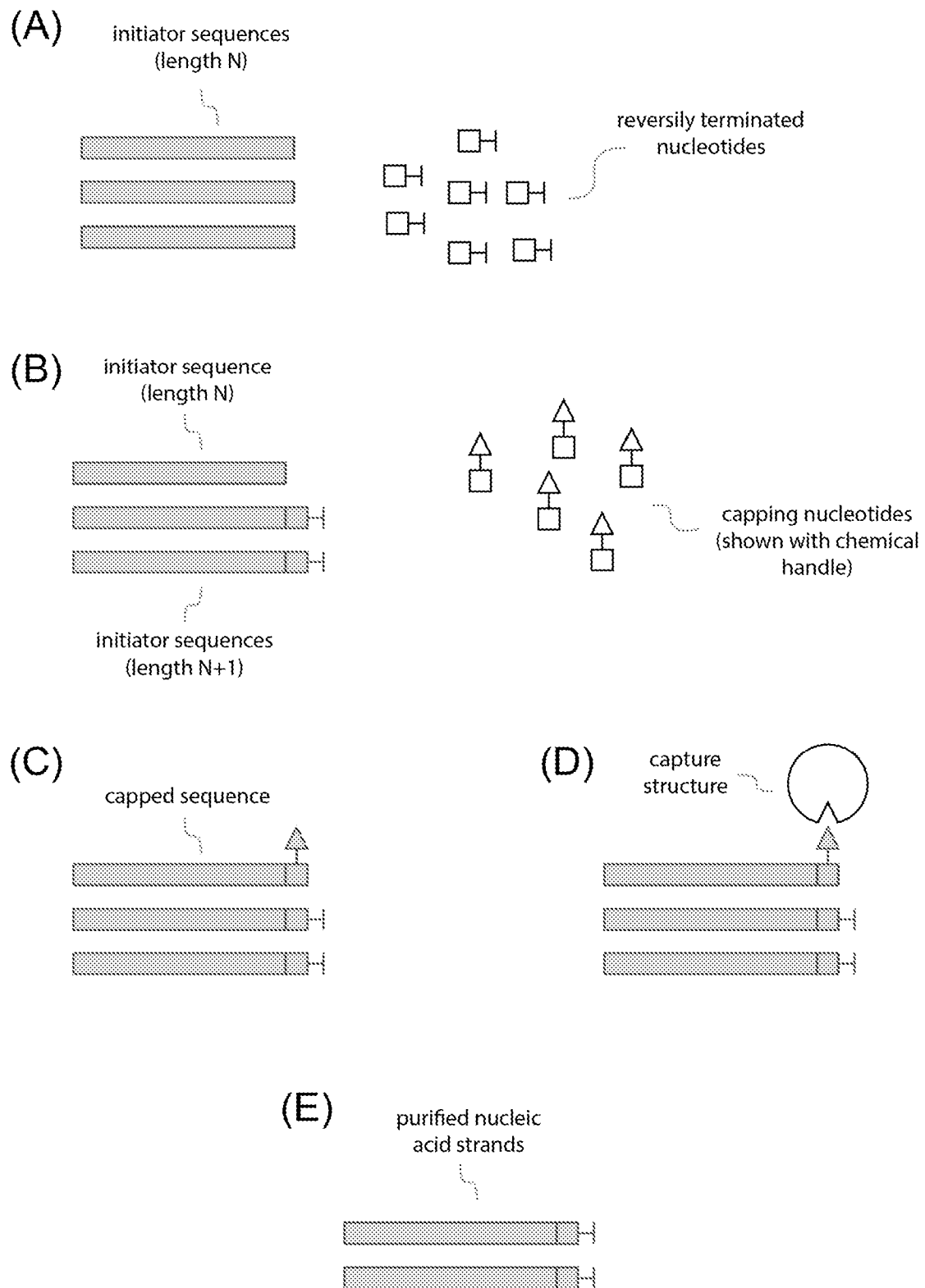


FIGURE 1

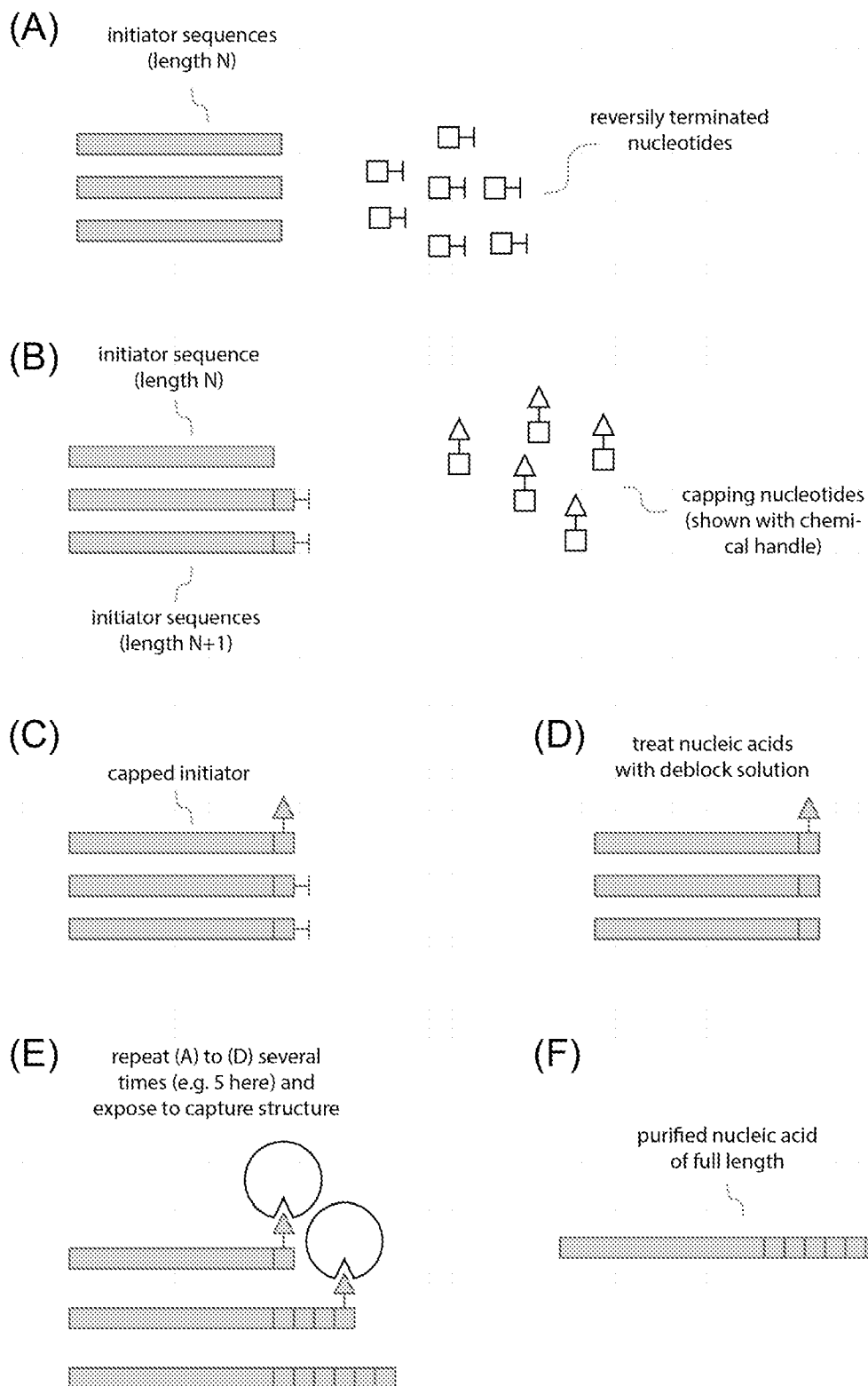


FIGURE 2

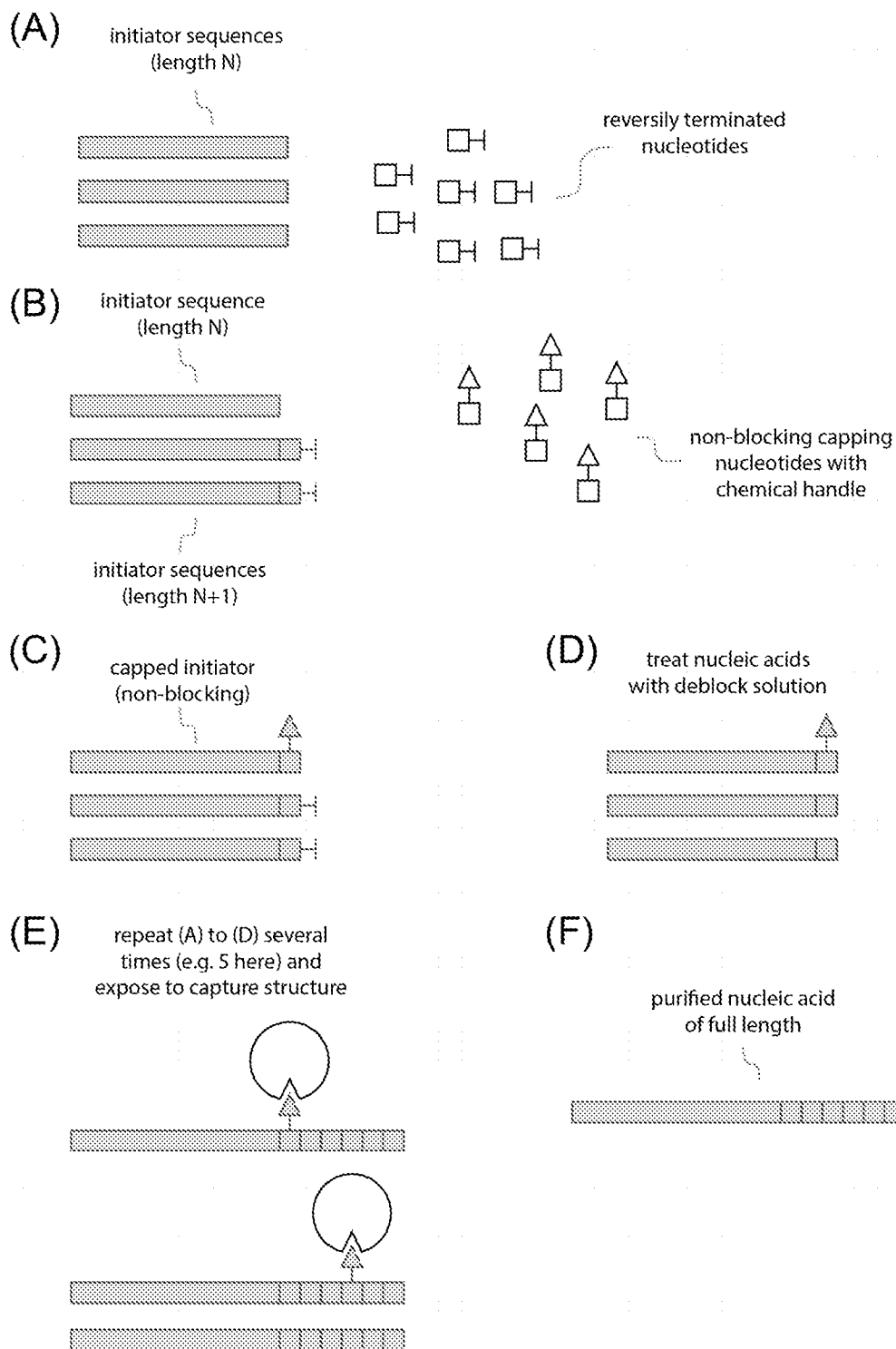


FIGURE 3

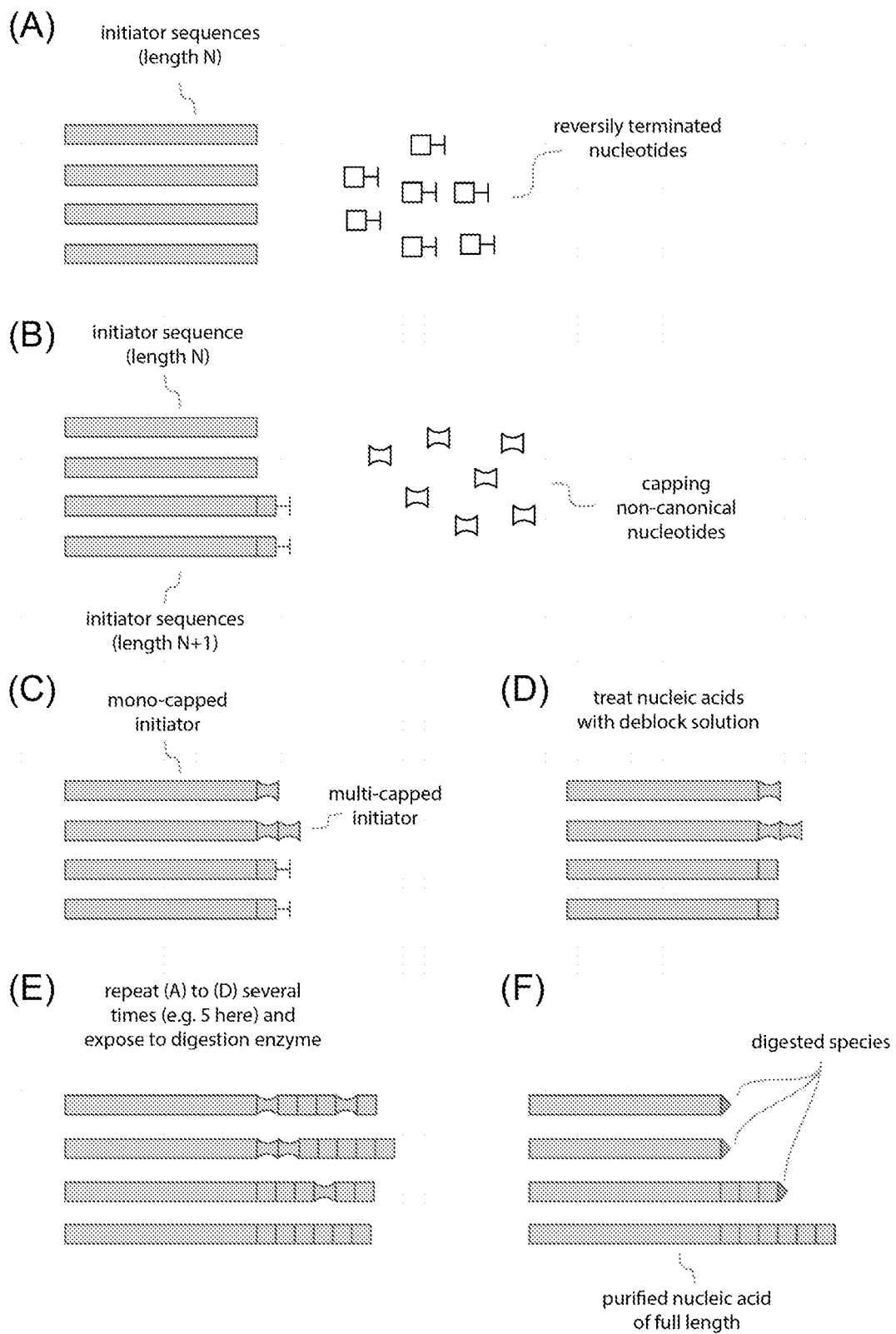


FIGURE 4

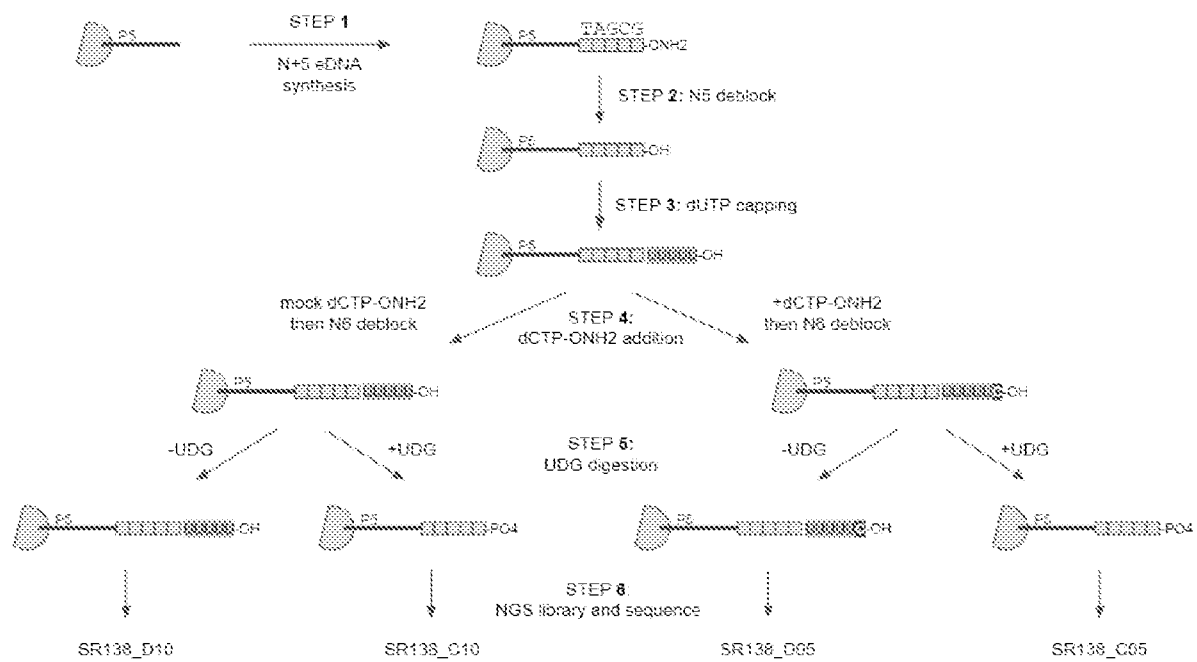


FIGURE 5

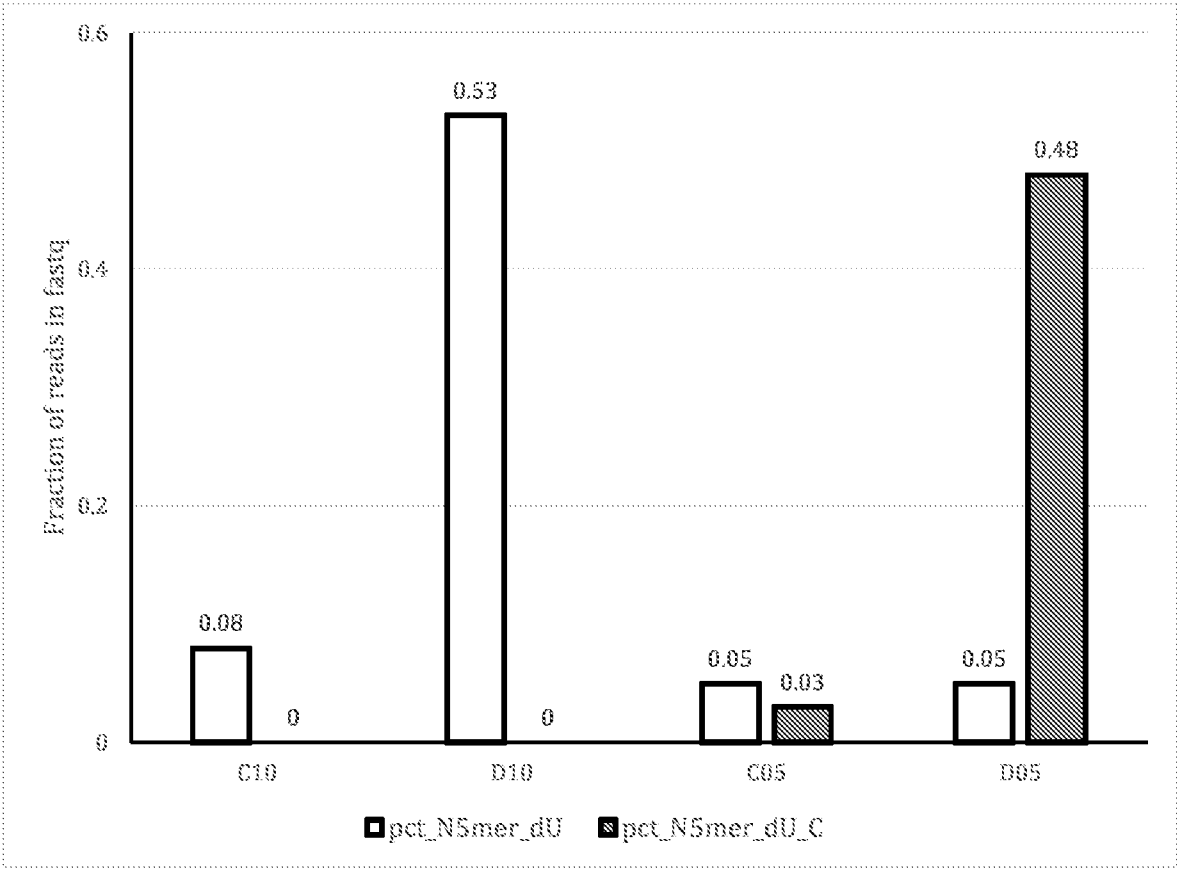


FIGURE 6

INTERNATIONAL SEARCH REPORT

International application No

PCT/GB2021/052095

A. CLASSIFICATION OF SUBJECT MATTER

INV. C12Q1/6806
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)

C12Q

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, CHEM ABS Data, EMBASE, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X Y	US 2016/046974 A1 (EFCAVITCH J WILLIAM [US] ET AL) 18 February 2016 (2016-02-18) abstract; claims 1,9-16,22-27,32,40; figures 5, 11, 21 paragraphs [0007], [0020], [0021], [0031], [0032], [0054], [0055], [0093], [0094] -----	1-18, 20-22 19
X Y	WO 2019/053443 A1 (NUCLERA NUCLEICS LTD [GB]) 21 March 2019 (2019-03-21) abstract; claims 7,8,12,15,18-21,23,27,28 page 18, line 22 - line 25 page 20, line 16 - line 27 -----	1-22 19



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

4 November 2021

Date of mailing of the international search report

15/11/2021

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

Celler, Jakub

INTERNATIONAL SEARCH REPORT

International application No.

PCT/GB2021/052095

Box No. I Nucleotide and/or amino acid sequence(s) (Continuation of item 1.c of the first sheet)

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of a sequence listing:
 - a. ☐ forming part of the international application as filed:
 - ☐ in the form of an Annex C/ST.25 text file.
 - ☐ on paper or in the form of an image file.
 - b. ☐ furnished together with the international application under PCT Rule 13~~ter~~.1(a) for the purposes of international search only in the form of an Annex C/ST.25 text file.
 - c. ☒ furnished subsequent to the international filing date for the purposes of international search only:
 - ☒ in the form of an Annex C/ST.25 text file (Rule 13~~ter~~.1(a)).
 - ☐ on paper or in the form of an image file (Rule 13~~ter~~.1(b) and Administrative Instructions, Section 713).
2. ☐ In addition, in the case that more than one version or copy of a sequence listing has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that forming part of the application as filed or does not go beyond the application as filed, as appropriate, were furnished.
3. Additional comments:

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/GB2021/052095

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2016046974	A1	18-02-2016	NONE

WO 2019053443	A1	21-03-2019	GB 2581044 A 05-08-2020
			US 2020270296 A1 27-08-2020
			WO 2019053443 A1 21-03-2019
