(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
16 November 2006 (16.11.2006)

PCT

(10) International Publication Number
WO 2006/120433 A1

(51) International Patent Classification:
C12N 9/12 (2006.01)

(21) International Application Number:
PCT/GB2006/001700

(22) International Filing Date: 10 May 2006 (10.05.2006)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

0509508.8	10 May 2005 (10.05.2005)	GB
60/757,997	11 January 2006 (11.01.2006)	US

(71) Applicant *(for all designated States except US)*: **SOLEXA LIMITED** [GB/GB]; Chesterford Research Park, Little Chesterford, Saffron Walden, Essex CB10 1XL (GB).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **OST, Tobias, William, Barr** [GB/GB]; c/o Solexa Limited, Chesterford Research Park, Little Chesterford, Saffron Walden, Essex CB10 1XL (GB). **SMITH, Geoffrey, Paul** [GB/GB]; c/o Solexa Limited, Chesterford Research Park, Little Chesterford, Saffron Walden, Essex CB10 1XL (GB). **BALASUBRAMANIAN, Shankar** [GB/GB]; c/o Solexa Limited, Chesterford Research Park, Little Chesterford, Saffron Walden, Essex CB10 1XL (GB). **RIGATTI, Roberto** [IT/GB]; c/o Solexa Limited, Chesterford Research Park, Little Chesterford, Saffron Walden, Essex CB10 1XL (GB). **SANCHES, Raquel, Maria** [PT/GB]; c/o Solexa Limited, Chesterford Research Park, Little Chesterford, Saffron Walden, Essex CB10 1XL (GB).

(74) Agents: **BALDOCK, Sharon, Claire** et al.; Boult Wade Tennant, Verulam Gardens, 70 Gray's Inn Road, London WC1X 8BT (GB).

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

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: IMPROVED POLYMERASES

(57) Abstract: Modified DNA polymerases have an affinity for DNA such that the polymerase has an ability to incorporate one or more nucleotides into a plurality of separate DNA templates in each reaction cycle. The polymerases are capable of forming an increased number of productive polymerase-DNA complexes in each reaction cycle. The modified polymerases may be used in a number of DNA sequencing applications, especially in the context of clustered arrays.

WO 2006/120433 A1

IMPROVED POLYMERASES

Field of the invention

The present invention relates to polymerase enzymes and more particularly to modified DNA polymerases having an affinity for DNA such that the polymerase has an ability to incorporate a nucleotide or nucleotides into a plurality of separate DNA templates in each reaction cycle and is capable of forming an increased number of productive polymerase-DNA complexes in each reaction cycle. Also included in the scope of the present invention are methods of using the modified polymerases for DNA sequencing, especially in the context of clustered arrays.

Background

The three-dimensional crystal structure of certain DNA polymerases has revealed three separate subdomains, named palm, fingers and thumb (Joyce, C.M. and Steiz, T. A. (1994) Function and structure relationships in DNA polymerases, Annu. Rev. Biochem., 63, 777-822), each having key roles during DNA polymerisation.

The C terminal thumb subdomain of DNA polymerases has been implicated in DNA binding and processivity (Doublie et al. 1998. Nature 391, 251; Truniger et al. 2004. Nucleic Acids Research 32, 371). Residues in this region of DNA polymerases interact with the primer:template duplex.

Disruption of the structure of this region either by the introduction of site-directed mutations or truncation by the deletion of a small number of amino acids, has provided evidence for variants with reduced DNA affinity and processivity without gross changes in other physical properties such as dNTP affinity and nucleotide insertion

fidelity (Truniger et al. 2004. Nucleic Acids Research 32, 371; Minnick et al. 1996. J. Biol. Chem., 271, 24954; Polesky et al. 1990. J. Biol. Chem., 265, 14579).

Polymerases may be separated into two structurally distinct families called family A and family B.

The C-terminal subdomain of family B polymerases has been poorly studied, but is believed to be involved in DNA binding based primarily on the inspection of the x-ray crystal structure of the closed form (DNA-bound) of polymerase RB69. Mutagenesis studies have been conducted within this thumb domain for two examples of the family B class, namely Phi29 and T4. However, these studies were limited to amino acid deletions of large portions of the domain. The same type of deletion has been carried out for Klenow (a family A polymerase). The performance of the variants in these studies was evaluated in terms of their ability to bind and incorporate dNTPs, the effect the deletion had on fidelity, their affinity for DNA and also their interaction with accessory proteins.

No studies of the thumb domain of the polymerase from a thermophilic archaeon have previously been carried out.

SUMMARY OF THE INVENTION

The present invention is based upon the realisation that the tight binding of a polymerase to the DNA template is not always an advantageous property. This is particularly the case in the context of sequencing reactions in which only a single nucleotide incorporation event is required in each reaction cycle. Thus, for a polymerase that binds tightly to DNA, the ability of the polymerase to take part in incorporation of nucleotides on multiple DNA

strands is restricted compared to a variant polymerase that has a lower affinity for DNA.

The present inventors have devised a method for sequencing DNA that uses nucleotide analogues bearing modifications at the 3′ sugar hydroxyl group which block incorporation of further nucleotides (see WO03/048387 for example and the citations described therein). The use of nucleotides bearing a 3' block allows successive nucleotides to be incorporated into a polynucleotide chain in a controlled manner. After each nucleotide addition the presence of the 3' block prevents incorporation of a further nucleotide into the chain. Once the nature of the incorporated nucleotide has been determined, the block may be removed, leaving a free 3' hydroxyl group for addition of the next nucleotide.

In addition, in the context of reactions such as sequencing reactions involving modified nucleotides (as discussed supra and in more detail herein below), tight binding of a polymerase may in fact present certain disadvantages in terms of reaction completion. For example, if an inactive polymerase molecule that has a tight DNA binding affinity forms a stable complex with a template DNA molecule no extension is possible from that particular template DNA molecule.

With this realisation, the present invention provides altered polymerases which have a weaker interaction with template DNA. Thus, the polymerase of the invention has an improved ability to move from one template DNA molecule to another during a reaction cycle. This ability to form an increased number of productive polymerase-DNA complexes has the benefit that levels or reaction completion in reactions

involving addition of a single nucleotide in each reaction cycle may be much improved.

Unmodified polymerases tend to bind DNA with high affinity such that the equation:

$$Pol + DNA \underset{k_d}{\overset{k_a}{\rightleftharpoons}} [Pol{:}DNA]$$

is heavily shifted to favour the [Pol:DNA] complex.

In contrast, in the present invention, the altered polymerases bind to DNA less well, meaning that the equilibrium position is shifted towards the left hand side.

Therefore, the invention provides an altered polymerase having reduced affinity for DNA such that the polymerase has an ability to incorporate a nucleotide or nucleotides into a plurality of separate DNA templates in each reaction cycle.

By "DNA template" is meant any DNA molecule which may be bound by the polymerase and utilised as a template for nucleic acid synthesis.

"Nucleotide" is defined herein to include both nucleotides and nucleosides. Nucleosides, as for nucleotides, comprise a purine or pyrimidine base linked glycosidically to ribose or deoxyribose, but they lack the phosphate residues which would make them a nucleotide. Synthetic and naturally occurring nucleotides are included within the definition. Labelled nucleotides are included within the definition. The advantageous properties of the polymerases are due to their reduced affinity for the DNA template in combination with a retained affinity and fidelity for the nucleotides which they incorporate.

In one preferred aspect, an altered polymerase is provided having a reduced affinity for DNA such that the polymerase has an ability to incorporate at least one synthetic nucleotide into a plurality of DNA templates in

each reaction cycle. Prior to the present invention, the problem of modifying a polymerase adapted to incorporate non-natural nucleotides, to reduce its DNA affinity whilst retaining its advantageous properties has neither been realised nor addressed.

In one embodiment, the nucleotides comprise dideoxy nucleotide triphosphates (ddNTPs) as used in the well known Sanger sequencing reaction. These nucleotides may be labelled, for example with any of a mass label, radiolabel or a fluorescent label.

In a further embodiment, the nucleotides comprise nucleotides which have been modified at the 3' sugar hydroxyl such that the substituent is larger in size than the naturally occurring 3' hydroxyl group, compared to a control polymerase.

In a preferred embodiment, the nucleotides comprise those having a purine or pyrimidine base and a ribose or deoxyribose sugar moiety having a removable 3'-OH blocking group covalently attached thereto, such that the 3' carbon atom has attached a group of the structure

-O-Z

wherein Z is any of $-C(R')_2-O-R''$, $-C(R')_2-N(R'')_2$, $-C(R')_2-N(H)R''$, $-C(R')_2-S-R''$ and $-C(R')_2-F$,

wherein each R" is or is part of a removable protecting group;

each R' is independently a hydrogen atom, an alkyl, substituted alkyl, arylalkyl, alkenyl, alkynyl, aryl, heteroaryl, heterocyclic, acyl, cyano, alkoxy, aryloxy, heteroaryloxy or amido group, or a detectable label attached through a linking group; or $(R')_2$ represents an alkylidene group of formula $=C(R''')_2$ wherein each R''' may be the same

or different and is selected from the group comprising hydrogen and halogen atoms and alkyl groups; and

wherein said molecule may be reacted to yield an intermediate in which each R" is exchanged for H or, where Z is $-C(R')_2-F$, the F is exchanged for OH, SH or NH_2, preferably OH, which intermediate dissociates under aqueous conditions to afford a molecule with a free 3'OH;

with the proviso that where Z is $-C(R')_2-S-R''$, both R' groups are not H.

The nucleosides or nucleotides which are incorporated by the polymerases of the present invention according to one embodiment, comprise a purine or pyrimidine base and a ribose or deoxyribose sugar moiety which has a blocking group covalently attached thereto, preferably at the 3'O position, which renders the molecules useful in techniques requiring blocking of the 3'-OH group to prevent incorporation of additional nucleotides, such as for example in sequencing reactions, polynucleotide synthesis, nucleic acid amplification, nucleic acid hybridisation assays, single nucleotide polymorphism studies, and other such techniques.

Once the blocking group has been removed, it is possible to incorporate another nucleotide to the free 3'-OH group.

Preferred modified nucleotides are exemplified in International Patent Application publication number WO 2004/018497 in the name of Solexa Limited, which reference is incorporated herein in its entirety.

In a preferred embodiment the R' group of the modified nucleotide or nucleoside is an alkyl or substituted alkyl. In a further embodiment the -Z group of the modified nucleotide or nucleoside is of formula $-C(R')_2-N_3$. In a most

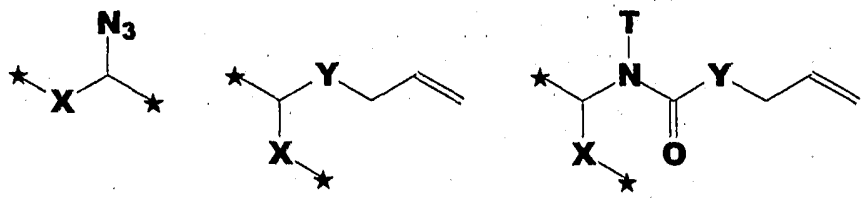
preferred embodiment the modified nucleotide or nucleoside includes a Z group which is an azido methyl group.

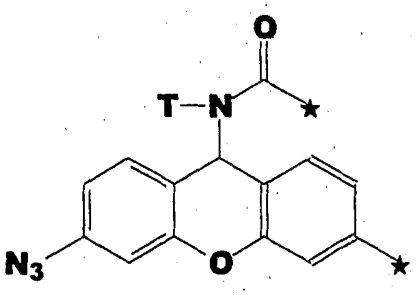
The preferred polymerases of the invention, as discussed in detail below, are particularly preferred for incorporation of nucleotide analogues wherein Z is an azido methyl group.

The modified nucleotide can be linked via the base to a detectable label by a desirable linker, which label may be a fluorophore, for example. The detectable label may instead, if desirable, be incorporated into the blocking groups of formula "Z". The linker can be acid labile, photolabile or contain a disulfide linkage. Other linkages, in particular phosphine-cleavable azide-containing linkers, may be employed in the invention as described in greater detail in WO 2004/018497, the contents of which are incorporated herein in their entirety.

Preferred labels and linkages include those disclosed in WO 03/048387. This reference is incorporated herein in its entirety.

In one embodiment the modified nucleotide or nucleoside has a base attached to a detectable label via a cleavable linker, characterised in that the cleavable linker contains a moiety selected from the group comprising:





(wherein X is selected from the group comprising O, S, NH and NQ wherein Q is a C_{1-10} substituted or unsubstituted alkyl group, Y is selected from the group comprising O, S, NH and N(allyl), T is hydrogen or a C_{1-10} substituted or unsubstituted alkyl group and * indicates where the moiety is connected to the remainder of the nucleotide or nucleoside).

In one embodiment the detectable label comprises a fluorescent label. Suitable fluorophores are well known in the art. In a preferred embodiment each different nucleotide type will carry a different fluorescent label. This facilitates the identification and incorporation of a particular nucleotide. Thus, for example modified Adenine, Guanine, Cytosine and Thymine would all have attached a separate fluorophore to allow them to be discriminated from one another readily. Surprisingly, it has been found that the altered polymerases are capable of incorporating modified nucleotide analogues carrying a number of different fluorescent labels. Moreover, the polymerases are capable of incorporating all four bases. These properties provide substantial advantages with regard to the use of the polymerases of the present invention in nucleic acid sequencing protocols.

As aforesaid, preferred nucleotide analogues include those containing O-azido methyl functionality at the 3' position. It will be appreciated that for other nucleotide

analogues the preferred amino acid sequence of the polymerase in the C terminal thumb sub-domain region, which contributes significantly to DNA binding, for optimum incorporation may vary. For any given nucleotide analogue, optimum sequence preferences in the C terminal thumb sub domain region (such as at residues Lys 790, 800, 844, 874, 878 and Arg 806 in RB69 and at residues Arg 743, Arg 713 and Lys 705 in 9°N polymerase, as discussed in greater detail below) may be determined by experiment, for example by construction of a library or discrete number of mutants followed by testing of individual variants in an incorporation assay system.

As aforementioned, the altered polymerases of the invention are capable of improved incorporation of all nucleotides, including a wide range of modified nucleotides having large 3' substituent groups of differing sizes and of varied chemical nature. The advantageous properties of the polymerases are due to their reduced affinity for the DNA template leading to increased dissociation of the polymerase from the DNA without adverse effects on affinity and fidelity for the nucleotides which they incorporate.

By virtue of the decreased DNA binding affinity of the polymerase of the invention, it is able to incorporate one or more nucleotides into several different DNA molecules in a single reaction cycle. Thus, the overall efficiency of reaction is improved, leading to greater levels of completion.

By "a reaction cycle" is meant a suitable reaction period to allow the incorporation of nucleotides into the template. Exemplary conditions for a single reaction cycle are one 30 minute, 45°C incubation period.

Many polymerisation reactions occur in the presence of an excess of DNA compared to polymerase. The polymerase of the present invention allows such a polymerisation reaction to proceed more effectively since the polymerase can catalyse numerous rounds of incorporation of a nucleotide or nucleotides on separate template DNA molecules. An unaltered polymerase on the other hand, particularly one which binds DNA much more tightly, will not have this ability since it is more likely to only participate in nucleotide incorporation on a single template in each reaction cycle. The polymerase according to the present invention allows high levels of reaction completion under conditions where the concentration of polymerase is limiting with respect to the concentration of DNA. In particular, the polymerase presents improved ability to incorporate one or more nucleotides into separate DNA molecules under conditions wherein the DNA:polymerase ratio is at least about 2:1, 3:1 or 5:1. However, at high concentrations of polymerase, the improvement may be masked.

Thus, an altered polymerase is provided having an affinity for DNA such that the polymerase is capable of forming an increased number of productive polymerase-DNA complexes in each reaction cycle.

The improved properties of the polymerases of the invention may be compared to a suitable control. "Control polymerase" is defined herein as the polymerase against which the activity of the altered polymerase is compared. The control polymerase is of the same type as the altered polymerase but does not carry the alteration which reduces the affinity of the polymerase for DNA. Thus, in a most preferred embodiment, the control polymerase is a 9°N polymerase and the modified polymerase is the same 9°N

polymerase except for the presence of one or more modifications which reduce the affinity of the 9°N polymerase for DNA.

In one embodiment, the control polymerase is a wild type polymerase which is altered to provide an altered polymerase which can be directly compared with the unaltered polymerase.

In one embodiment, the control polymerase comprises substitution mutations at positions which are functionally equivalent to Leu408 and Tyr409 and Pro410 in the 9°N DNA polymerase amino acid sequence. Thus, in this embodiment the control polymerase has a substitution mutation at position 408 from leucine to a different amino acid, at position 409 from tyrosine to a different amino acid and at position 410 from proline to a different amino acid or at positions which are functionally equivalent if the polymerase is not a 9°N DNA polymerase. In a preferred embodiment, the control polymerase is a 9°N DNA polymerase comprising the said substitution mutations.

In another embodiment, the control polymerase comprises substitution mutations which are functionally equivalent to Leu408Tyr and Tyr409Ala and Pro410Val in the 9°N DNA polymerase amino acid sequence. Thus, in this embodiment the control polymerase has a substitution mutation at position 408 from leucine to tyrosine, at position 409 from tyrosine to alanine and at position 410 from proline to valine or at positions which are functionally equivalent if the polymerase is not a 9°N DNA polymerase. In a preferred embodiment, the control polymerase is a 9°N DNA polymerase comprising the said substitution mutations.

The control polymerase may further comprise a substitution mutation at the position functionally

equivalent to Cys223 in the 9°N DNA polymerase amino acid sequence. Thus, in this embodiment the control polymerase has a substitution mutation at position 223 from cysteine to a different amino acid, or at a position which is functionally equivalent if the polymerase is not a 9°N DNA polymerase. In a preferred embodiment, the control polymerase is a 9°N DNA polymerase comprising the said substitution mutation. In another embodiment, the control polymerase comprises the substitution mutation functionally equivalent to Cys223Ser in the 9°N DNA polymerase amino acid sequence. Thus, in this embodiment the control polymerase has a substitution mutation at position 223 from cysteine to serine, or at a position which is functionally equivalent if the polymerase is not a 9°N DNA polymerase. In a preferred embodiment, the control polymerase is a 9°N DNA polymerase comprising the said substitution mutation.

Preferably, the control polymerase is a 9°N DNA polymerase comprising a combination of the above mentioned mutations.

The polymerase will generally have a reduced affinity for DNA. This may be defined in terms of dissociation constant. Thus, wild type polymerases tend to have dissociation constants in the nano-picomolar range. For the purposes of the present invention, a polymerase having an affinity for DNA which is reduced compared to the control unaltered polymerase is suitable. Preferably, due to the alteration, the polymerase has at least a, or approximately a, 2-fold, 3-fold, 4-fold or 5-fold etc increase in its dissociation constant when compared to the control unaltered polymerase.

By "functionally equivalent" is meant the amino acid substitution that is considered to occur at the amino acid

position in another polymerase that has the same functional role in the enzyme. As an example, the mutation at position 412 from Tyrosine to Valine (Y412V) in the Vent DNA polymerase would be functionally equivalent to a substitution at position 409 from Tyrosine to Valine (Y409V) in the 9°N polymerase. The bulk of this amino acid residue is thought to act as a "steric gate" to block access of the 2'-hydroxyl of the nucleotide sugar to the binding site. Also, residue 488 in Vent polymerase is deemed equivalent to amino acid 485 in 9°N polymerase, such that the Alanine to Leucine mutation at 488 in Vent (A488L) is deemed equivalent to the A485L mutation in 9°N polymerase.

Generally, functionally equivalent substitution mutations in two or more different polymerases occur at homologous amino acid positions in the amino acid sequences of the polymerases. Hence, use herein of the term "functionally equivalent" also encompasses mutations that are "positionally equivalent" or "homologous" to a given mutation, regardless of whether or not the particular function of the mutated amino acid is known. It is possible to identify positionally equivalent or homologous amino acid residues in the amino acid sequences of two or more different polymerases on the basis of sequence alignment and/or molecular modelling.

The altered polymerase will generally be an "isolated" or "purified" polypeptide. By "isolated polypeptide" is meant a polypeptide that is essentially free from contaminating cellular components, such as carbohydrates, lipids, nucleic acids or other proteinaceous impurities which may be associated with the polypeptide in nature. Typically, a preparation of the isolated polymerase contains the polymerase in a highly purified form, i.e. at least

about 80% pure, preferably at least about 90% pure, more preferably at least about 95% pure, more preferably at least about 98% pure and most preferably at least about 99% pure. Purity of a preparation of the enzyme may be assessed, for example, by appearance of a single band on a standard SDS-polyacrylamide electrophoresis gel.

The altered polymerase may be a "recombinant" polypeptide.

The altered polymerase according to the invention may be any DNA polymerase. More particularly, the altered polymerase may be a family B type DNA polymerase, or a mutant or variant thereof. Family B DNA polymerases include numerous archael DNA polymerase, human DNA polymerase α and T4, RB69 and ϕ29 phage DNA polymerases. These polymerases are less well studied than the family A polymerases, which include polymerases such as *Taq*, and T7 DNA polymerase. In one embodiment the polymerase is selected from any family B archael DNA polymerase, human DNA polymerase α or T4, RB69 and ϕ29 phage DNA polymerases.

The archael DNA polymerases are in many cases from hyperthermophilic archea, which means that the polymerases are often thermostable. Accordingly, in a further preferred embodiment the polymerase is a thermophilic archaeon polymerase and is preferably selected from Vent, Deep Vent, 9°N and Pfu polymerase. Vent and Deep Vent are commercial names used for family B DNA polymerases isolated from the hyperthermophilic archaeon *Thermococcus litoralis* and *Pyrococcus furiosus* respectively. 9°N polymerase was also identified from *Thermococcus sp.* Pfu polymerase was isolated from *Pyrococcus furiosus*. As mentioned above, prior to the present invention the thumb domain from a thermophilic polymerase had not been studied. The most preferred

polymerase in the present invention is 9°N polymerase, including mutants and variants thereof. 9°N polymerase has no requirement for accessory proteins. This can be contrasted with previously studied polymerases in which deletions in the thumb domain were shown to adversely affect the interaction with accessory proteins whilst not altering other properties of the polymerase. In contrast, as is shown in the Experimental Section below, a deletion of a large number of residues of 9°N has a significant adverse effect on the important properties of 9°N such that catalytic activity is severely compromised.

It is to be understood that the invention is not intended to be limited to mutants or variants of the family B polymerases. The altered polymerase may also be a family A polymerase, or a mutant or variant thereof, for example a mutant or variant *Taq* or T7 DNA polymerase enzyme, or a polymerase not belonging to either family A or family B, such as for example reverse transcriptases. However, for reasons described herein family B polymerases are especially preferred.

A number of different types of alteration are contemplated by the invention, which result in a polymerase displaying the desired properties as a result of a reduced affinity for DNA. Particularly preferred are substitution mutations in the primary amino acid sequence of the polymerase, although addition and deletion mutations may also produce useful polymerases. Suitable alteration techniques, such as site directed mutagenesis for example, are well known in the art.

Thus, by "altered polymerase" it is meant that the polymerase has at least one amino acid change compared to the control polymerase enzyme. In general this change will

comprise the substitution of at least one amino acid for another. In preferred embodiments, these changes are non-conservative changes, although conservative changes to maintain the overall charge distribution of the protein are also envisaged in the present invention. Moreover, it is within the contemplation of the present invention that the modification in the polymerase sequence may be a deletion or addition of one or more amino acids from or to the protein, provided that the resultant polymerase has reduced DNA affinity and an ability to incorporate a nucleotide or nucleotides into a plurality of separate DNA templates in each reaction cycle compared to a control polymerase.

In one embodiment, the alteration to form the polymerase of the invention comprises at least one mutation, and preferably at least one substitution mutation, at a residue in the polymerase which destabilises the interaction of the polymerase with DNA. Thus, the resultant polymerase interacts in a less stable manner with DNA. As aforementioned, a decrease in affinity of the polymerase for DNA allows it to incorporate one or more nucleotides into several different DNA molecules in a single reaction cycle. Thus, the overall efficiency of reaction is improved, leading to greater levels of reaction completion.

In a further embodiment, the alteration comprises at least one mutation, and preferably at least one substitution mutation, at a residue in the polymerase which binds to DNA. Suitable target residues for mutation can be selected according to available crystal structures for suitable polymerases, particularly when crystallised in the closed state (bound to DNA). By reducing the number of binding contacts with the DNA, an overall reduction in DNA binding affinity may be achieved. Thus, the resultant polymerase

displays improved characteristics in the context of nucleotide incorporation reactions in which tight binding to DNA is disadvantageous.

In similar fashion, the polymerase may also carry an alteration which comprises at least one mutation, and preferably at least one substitution mutation, at a residue found in the DNA binding domain of the polymerase. Again, such a mutation is predicted to decrease the DNA binding affinity of the altered polymerase such that it is able to more readily bind to and dissociate from separate template DNA molecules during a reaction.

In one embodiment, the polymerase includes an alteration which comprises at least one mutation, and preferably at least one substitution mutation, at a basic amino acid residue in the polymerase. As is well known in the art, many positively charged amino acid residues in polymerases interact with the overall negatively charged DNA double helix, in particular with specific phosphate groups of nucleotides in the DNA.

As aforementioned, the preferred type of alteration resulting in a polymerase according to the invention comprises at least one substitution mutation. As is shown in the experimental section below, deletion of residues from the polymerase amino acid sequence can lead to a polymerase which, whilst having a reduced affinity for DNA, does not have overall advantageous properties since catalytic ability is impaired. In one particularly preferred embodiment, the polymerase comprises two substitution mutations, but may contain four, five, six or seven etc. mutations provided that the resultant polymerase has the desired properties.

Preferably, the affinity of the polymerase for nucleotides is substantially unaffected by the alteration.

As is shown in the experimental section (in particular example 6), it is possible to mutate a polymerase such that its affinity for DNA is reduced, whilst the affinity of the polymerase for a nucleotide, which may be a dNTP or ddNTP or a modified version thereof for example (see the definition of nucleotide supra) is not adversely affected. By "substantially unaffected" in this context is meant that the affinity for the nucleotide remains of the same order as for the unaltered polymerase. Preferably, the affinity for nucleotides is unaffected by the alteration.

Preferably, the fidelity of the polymerase is substantially unaffected by the alteration. As is shown in the experimental section (in particular example 6), it is possible to mutate a polymerase such that its affinity for DNA is reduced, whilst the fidelity of the polymerase is substantially unaffected by the alteration. By "substantially unaffected" in this context is meant that the misincorporation frequency for each nucleotide remains of the same order as for the unaltered polymerase. Preferably, the fidelity of the polymerase is unaffected by the alteration.

In terms of specific and preferred structural mutants, these may be based upon the most preferred polymerase, namely 9°N DNA polymerase. As discussed in example 1 below, an energy minimised overlaid alignment (contracted by Cresset) of the crystal structures of the open form of 9°N-7 DNA polymerase (PDB = 1qht), the open structure of a closely related DNA polymerase RB69 (PDB = 1ih7) and the closed form of RB69 (PDB = 1ig9) was used as a structural model for the identification of key residues involved in DNA binding. Accordingly, an altered polymerase is provided which comprises or incorporates one, two or three amino acid

substitution mutations to a different amino acid at the position or positions functionally equivalent to Lys705, Arg713 and/or Arg743 in the 9°N DNA polymerase amino acid sequence. Preferably, the polymerase is a 9°N DNA polymerase comprising these mutations. All combinations and permutations of one, two or three mutations are contemplated within the scope of the invention.

Mutations may also be made at other specific residues based upon alignment of the "open" 9°N DNA polymerase structure (i.e. not bound to DNA) with the known crystal structure of the RB69 polymerase complexed with DNA. Thus, an altered polymerase is provided which comprises or incorporates one or two amino acid substitution mutations to a different amino acid at the position or positions functionally equivalent to Arg606 and/or His679 in the 9°N DNA polymerase amino acid sequence. Preferably, the polymerase is a 9°N DNA polymerase comprising these mutations. All combinations and permutations of different mutations are contemplated within the scope of the invention. Thus, these mutations may be made in combination with the other mutations discussed supra.

In one preferred embodiment, the polymerase comprises at least a substitution mutation to a different amino acid at the position functionally equivalent to either Arg713 or Arg743 in the 9°N DNA polymerase amino acid sequence. These two positions represent particularly preferred sites for mutation, as discussed in more detail in the experimental section below. Both residues may be mutated in the same polymerase to a different amino acid.

In terms of the nature of the different amino acid, the substitution mutation or mutations preferably convert the substituted amino acid to a non-basic amino acid (i.e. not

lysine or arginine). Any non-basic amino acid may be chosen.

Preferred substitution mutation or mutations convert the substituted amino acid to an amino acid selected from:

(i) acidic amino acids,

(ii) aromatic amino acids, particularly tyrosine (Y) or phenylalanine (F); and

(iii) non-polar amino acids, particularly, alanine (A), glycine (G) or methionine (M).

In one embodiment, the substitution mutation or mutations convert the substituted amino acid to alanine.

In a more specific embodiment, an altered polymerase is provided comprising the substitution mutation or mutations which are functionally equivalent to Lys705Ala and/or Arg713Ala and/or Arg743Ala in the 9°N DNA polymerase amino acid sequence. Thus, in this embodiment the polymerase has a substitution mutation at position 705 from lysine to alanine and/or at position 713 from arginine to alanine and/or at position 743 from arginine to alanine or at positions which are functionally equivalent if the polymerase is not a 9°N DNA polymerase. In a preferred embodiment, the polymerase is a 9°N DNA polymerase comprising the said substitution mutations.

In one embodiment, the altered polymerase comprises the amino acid substitution functionally equivalent to Arg713Ala and in a further embodiment, the altered polymerase comprises the amino acid substitution functionally equivalent to Arg743Ala. Preferably, the altered polymerase is a 9°N DNA polymerase.

Specific structural mutants may also be based upon other types of polymerase, such as the RB69 polymerase for which the "open" and "closed" structures are known.

Accordingly, an altered polymerase is provided which comprises or incorporates one, two, three, four, five or six amino acid substitution mutations to a different amino acid at the position or positions functionally equivalent to Lys790, Lys800, Arg806, Lys844, Lys874 and/or Lys878 in the RB69 DNA polymerase amino acid sequence. Preferably, the polymerase is a 9°N DNA polymerase comprising these analogous or functionally equivalent mutations. All combinations and permutations of one, two, three, four, five or six mutations are contemplated within the scope of the invention.

In terms of the nature of the different amino acid, the substitution mutation or mutations preferably convert the substituted amino acid to a non-basic amino acid (i.e. not lysine or arginine). Any non-basic amino acid may be chosen.

Preferred substitution mutation or mutations convert the substituted amino acid to an amino acid selected from:

(i) acidic amino acids,

(ii) aromatic amino acids, particularly tyrosine (Y) or phenylalanine (F); and

(iii) non-polar amino acids, particularly, alanine (A), glycine (G) or methionine (M).

In one embodiment, the substitution mutation or mutations convert the substituted amino acid to alanine.

It should be noted that the present invention is not limited to polymerases which have only been altered in the above mentioned manner. Polymerases of the invention may include a number of additional mutations, such as for example the preferred mutant polymerases disclosed in detail in WO 2005/024010. In particular, a polymerase comprising substitution mutations at positions which are functionally

equivalent to Leu408 and Tyr409 and Pro410 in the 9°N DNA polymerase amino acid sequence is contemplated. In a preferred embodiment, the polymerase is a 9°N DNA polymerase comprising the said substitution mutations.

In a specific embodiment, the polymerase comprises the substitution mutations which are functionally equivalent to at least one or two but preferably all of Leu408Tyr and Tyr409Ala and Pro410Val in the 9°N DNA polymerase amino acid sequence. In a preferred embodiment, the polymerase is a 9°N DNA polymerase comprising all the said substitution mutations.

The polymerase may further comprise a substitution mutation at the position functionally equivalent to Cys223 in the 9°N DNA polymerase amino acid sequence. In a preferred embodiment, the polymerase is a 9°N DNA polymerase comprising the said substitution mutation. In one embodiment, the polymerase comprises the substitution mutation functionally equivalent to Cys223Ser in the 9°N DNA polymerase amino acid sequence. In a preferred embodiment, the polymerase is a 9°N DNA polymerase comprising the said substitution mutation.

Preferably, the polymerase is a 9°N DNA polymerase comprising a combination of the above mentioned mutations.

The invention also relates to a 9°N polymerase molecule comprising, consisting essentially of or consisting of the amino acid sequence shown as any one of SEQ ID NO: 1, 3, 5 or 21. The invention also encompasses polymerases having amino acid sequences which differ from those shown as SEQ ID NOs: 1, 3, 5 and 21 only in amino acid changes which do not affect the function of the polymerase to a material extent. In this case the relevant function of the polymerase is defined as a reduced affinity for DNA such that the

polymerase has an ability to incorporate a nucleotide or nucleotides into a plurality of separate DNA templates in each reaction cycle (compared to a control polymerase) and/or that the polymerase is capable of forming an increased number of productive polymerase-DNA complexes in each reaction cycle (compared to a control polymerase).

Thus, conservative substitutions at residues which are not important for this activity of the polymerase variants having reduced DNA affinity are included within the scope of the invention. The effect of further mutations on the function of the enzyme may be readily tested, for example using well known nucleotide incorporation assays (such as those described in the examples of WO 2005/024010 and in examples 3 and 4 below).

The altered polymerase of the invention may also be defined directly with reference to its reduced affinity for DNA, which together with a substantially unaltered fidelity and affinity for nucleotides produce the advantages associated with the polymerases of the invention. Thus, an altered polymerase is provided which has a dissociation constant (K_D) for DNA of at least, or in the region of between, approximately 2-fold greater, 3-fold greater, 4-fold greater or 5-fold greater than the unaltered control polymerase.

In one embodiment, an altered polymerase is provided which will dissociate from DNA in the presence of a salt solution, preferably a NaCl solution, having a concentration of less than or equal to about 500mM, preferably less than 500 mM. The salt solution may be of a suitable concentration such that the reduced affinity polymerase of the invention can be distinguished from an unaltered polymerase which binds DNA more tightly. Suitable salt

solution concentrations (preferably NaCl) are in the region of approximately 150mM, 200mM, 250mM, 300mM or 350mM preferably 200mM. Any suitable double stranded DNA molecule may be utilised to determine whether the alteration has the desired effect in terms of reducing DNA affinity. Preferably, the DNA molecule from which the polymerase dissociates comprises the sequence set forth as SEQ ID No.: 18. Preferably, at least approximately 40%, 50%, 60%, 70%, etc., of the polymerase will dissociate from the DNA at the relevant NaCl concentration in the wash solution.

Dissociation experiments may be carried out by any known means, such as, for example, by utilising the washing assay detailed in example 5 of the Experimental Section below (see also Figures 6 and 7).

As aforementioned, the reduction in DNA affinity is (preferably) achieved without a notable or significant decrease in the affinity of the polymerase for nucleotides. Surprisingly, the altered polymerase of the invention may also display comparable activity, for example, in terms of Vmax, to the unmodified polymerase even though the DNA binding affinity has been decreased. This surprising property displayed by the polymerases of the present invention is shown in the kinetic analysis of certain enzymes of the invention in particular in example 6 of the Experimental Section below and with reference to Figure 8.

The altered polymerase of the invention may also be defined directly with reference to its improved ability to be purified from host cells in which the polymerase is expressed. Thus, thanks to the reduced affinity of the altered polymerase for DNA (which together with a substantially unaltered affinity for nucleotides and fidelity produce the advantages associated with the

polymerases of the invention) the polymerase can more readily be purified. Less endogenous DNA from the host cell is carried over during purification of the enzyme. Thus, a more pure product results since less endogenous DNA remains bound to the polymerase following the purification process. An additional advantage of the reduced affinity for DNA of the altered polymerases is that less severe purification procedures need to be utilised in order to provide a substantially pure polymerase preparation. Accordingly, less polymerase will be adversely affected by the purification process itself leading to a polymerase preparation with higher levels of overall activity. In addition, more uniform purification should be possible leading to less variability between batches of polymerase. Representative data regarding the improvement in carry over of endogenous DNA during the purification procedure is provided in Example 7 of the experimental section below.

Preferably, less than about 60 ng/ml, 50 ng/ml, 40 ng/ml, 30 ng/ml, 20 ng/ml, 10 ng/ml and more preferably less than about 5 ng/ml of host DNA is carried over following purification of the polymerase. Standard purification protocols may be utilised, such as for example see Colley *et al.*, *J. Biol. Chem.* 264:17619-17622 (1989); *Guide to Protein Purification*, in *Methods in Enzymology*, vol. 182 (Deutscher, ed., 1990).

Thus, the invention provides an altered polymerase having an affinity for DNA such that;

(i) the polymerase has a dissociation constant for DNA of at least about, or approximately, 2-fold, 3-fold, 4-fold or 5-fold greater than the unaltered/control polymerase and/or

(ii) at least 50%, 60%, 70% or 80% of the polymerase dissociates from DNA to which the polymerase is bound when a sodium chloride solution having a concentration of between about 200nM and 500nM, preferably between about 200 nM and 300 nM is applied thereto, and/or

(iii) less than about 60, 50, 45, 40, 35, 30, 25, 20, 15, 10, 5, 3, 1 or 0.5 ng/ml of endogenous DNA remains bound to the polymerase following a purification process from the cell in which the polymerase is expressed;

the alteration not significantly adversely affecting nucleotide binding ability or fidelity such that the polymerase is capable of;

(a) forming an increased number of productive polymerase-DNA complexes over a reaction cycle (giving improved levels of reaction completion), and/or

(b) catalysing an improved(/increased/elevated) overall level of nucleotide incorporation;

especially under conditions where the concentration of polymerase is limiting with respect to the concentration of DNA.

The invention further relates to nucleic acid molecules encoding the altered polymerase enzymes of the invention.

For any given altered polymerase which is a mutant version of a polymerase for which the amino acid sequence and preferably also the wild type nucleotide sequence encoding the polymerase is known, it is possible to obtain a nucleotide sequence encoding the mutant according to the basic principles of molecular biology. For example, given that the wild type nucleotide sequence encoding 9°N polymerase is known, it is possible to deduce a nucleotide sequence encoding any given mutant version of 9°N having one or more amino acid substitutions using the standard genetic

code. Similarly, nucleotide sequences can readily be derived for mutant versions other polymerases from both family A and family B polymerases such as, for example, Vent™, Pfu, Tsp JDF-3, Taq, etc. Nucleic acid molecules having the required nucleotide sequence may then be constructed using standard molecular biology techniques known in the art.

In one particular embodiment the invention relates to nucleic acid molecules encoding mutant versions of the 9°N polymerase.

Therefore, the invention provides a nucleic acid molecule which encodes an altered 9°N polymerase, the nucleic acid molecule comprising, consisting essentially of or consisting of the nucleotide sequence of any of SEQ ID NO: 2, 4, 6, 19 or 20.

In accordance with the present invention, a defined nucleic acid includes not only the identical nucleic acid but also any minor base variations including, in particular, substitutions in cases which result in a synonymous codon (a different codon specifying the same amino acid residue) due to the degenerate code in conservative amino acid substitutions. The term "nucleic acid sequence" also includes the complementary sequence to any single stranded sequence given regarding base variations.

The nucleic acid molecules described herein may also, advantageously, be included in a suitable expression vector to express the polymerase proteins encoded therefrom in a suitable host. Thus, there is provided an expression vector comprising, consisting essentially of or consisting of the nucleotide sequence of any of SEQ ID NO: 2, 4, 6, 19 or 20. Incorporation of cloned DNA into a suitable expression vector for subsequent transformation of said cell and

subsequent selection of the transformed cells is well known to those skilled in the art as provided in Sambrook et al. (1989), Molecular cloning: A Laboratory Manual, Cold Spring Harbour Laboratory.

Such an expression vector includes a vector having a nucleic acid according to the invention operably linked to regulatory sequences, such as promoter regions, that are capable of effecting expression of said DNA fragments. The term "operably linked" refers to a juxtaposition wherein the components described are in a relationship permitting them to function in their intended manner. Such vectors may be transformed into a suitable host cell to provide for the expression of a protein according to the invention.

The nucleic acid molecule may encode a mature protein or a protein having a prosequence, including that encoding a leader sequence on the preprotein which is then cleaved by the host cell to form a mature protein.

The vectors may be, for example, plasmid, virus or phage vectors provided with an origin of replication, and optionally a promoter for the expression of said nucleotide and optionally a regulator of the promoter. The vectors may contain one or more selectable markers, such as, for example, an antibiotic resistance gene.

Regulatory elements required for expression include promoter sequences to bind RNA polymerase and to direct an appropriate level of transcription initiation and also translation initiation sequences for ribosome binding. For example, a bacterial expression vector may include a promoter such as the lac promoter and for translation initiation the Shine-Delgarno sequence and the start codon AUG. Similarly, a eukaryotic expression vector may include a heterologous or homologous promoter for RNA polymerase II,

a downstream polyadenylation signal, the start codon AUG, and a termination codon for detachment of the ribosome. Such vectors may be obtained commercially or be assembled from the sequences described by methods well known in the art.

Transcription of DNA encoding the polymerase of the invention by higher eukaryotes may be optimised by including an enhancer sequence in the vector. Enhancers are cis-acting elements of DNA that act on a promoter to increase the level of transcription. Vectors will also generally include origins of replication in addition to the selectable markers.

Preferred uses of the altered polymerases

In a further aspect the invention relates to use of an altered polymerase having reduced affinity for DNA according to the invention for the incorporation of a nucleotide into a polynucleotide. As mentioned above, the nature of the nucleotide is not limiting since the altered polymerases of the invention retain affinity for the relevant nucleotides.

As aforementioned, the invention is based upon the realization that, the tight binding of a polymerase to the DNA template is not always an advantageous property. This is particularly the case in the context of sequencing reactions in which only a single nucleotide incorporation event is required in each reaction cycle for each template DNA molecule. In many of these sequencing reactions a labelled nucleotide is utilised.

Thus, the invention provides for use of a polymerase which has been altered such that it displays a reduced affinity for DNA and an ability to incorporate a labelled

nucleotide into a plurality of separate DNA templates in each reaction cycle for incorporation of a labelled nucleotide into a polynucleotide, the label being utilised to determine the nature of the nucleotide added.

In one embodiment, the nucleotide comprises a ddNTP. Thus, the polymerase of the invention may be utilised in a conventional Sanger sequencing reaction, the details of which are well known in the art.

In a preferred embodiment, the nucleotide is a modified nucleotide which has been modified at the 3' sugar hydroxyl such that the substituent is larger in size than the naturally occurring 3' hydroxyl group.

The polymerases of the invention may be used in any area of technology where it is required/desirable to be able to incorporate nucleotides, for example modified nucleotides having a substituent at the 3' sugar hydroxyl position which is larger in size than the naturally occurring hydroxyl group, into a polynucleotide chain. They may be used in any area of technology where any of the desirable properties of the enzyme, for example improved rate of incorporation of nucleotides even under conditions where the DNA is present in excess and increased levels of reaction completion under these conditions, are required. This may be a practical, technical or economic advantage.

Although the altered polymerases exhibit desirable properties in relation to incorporation of modified nucleotides having a large 3' substituent due to their decreased affinity for DNA, the utility of the enzymes is not confined to incorporation of such nucleotide analogues. The desirable properties of the altered polymerase due to its reduced affinity for DNA may provide advantages in relation to incorporation of any other nucleotide, including

unmodified nucleotides, relative to enzymes known in the art. In essence, the altered polymerases of the invention may be used to incorporate any type of nucleotide that they have the ability to incorporate.

The polymerases of the present invention are useful in a variety of techniques requiring incorporation of a nucleotide into a polynucleotide, which include sequencing reactions, polynucleotide synthesis, nucleic acid amplification, nucleic acid hybridisation assays, single nucleotide polymorphism studies, and other such techniques. Use in sequencing reactions represents a most preferred embodiment. All such uses and methods utilizing the modified polymerases of the invention are included within the scope of the present invention.

The invention also relates to a method for incorporating nucleotides into DNA comprising allowing the following components to interact:

(i) A polymerase according to the invention;

(ii) a DNA template; and

(iii) a nucleotide solution.

As discussed above, the polymerase of the invention has particular applicability in reactions where incorporation of only a single or relatively few nucleotides are required in each reaction cycle. Often in these reactions one or more of the nucleotides will be labelled. Accordingly, the invention provides a method for incorporating labelled nucleotides into DNA comprising allowing the following components to interact:

(i) A polymerase which has been altered such that it displays a reduced affinity for DNA and an ability to incorporate a labelled nucleotide

into a plurality of separate DNA templates in each reaction cycle,

(ii) a DNA template; and

(iii) a nucleotide solution.

In one specific embodiment, the invention provides a method for incorporating nucleotides which have been modified at the 3' sugar hydroxyl such that the substituent is larger in size than the naturally occurring 3' hydroxyl group into DNA comprising allowing the following components to interact:

- A polymerase according to the present invention (as described above)
- a DNA template; and
- a nucleotide solution containing the nucleotides which have been modified at the 3' sugar hydroxyl such that the substituent is larger in size than the naturally occurring 3' hydroxyl group.

Particularly preferred are uses and methods carried out on a clustered array. Clustered arrays of nucleic acid molecules may be produced using techniques generally known in the art. By way of example, WO 98/44151 and WO 00/18957 (both of which are incorporated by reference herein) both describe methods of nucleic acid amplification which allow amplification products to be immobilised on a solid support in order to form arrays comprised of clusters or "colonies" of immobilised nucleic acid molecules. Reference is also made to WO 2005/078130 including the citations referred to therein, the contents of all of which are hereby incorporated by reference. Incorporation on clusters, in particular sequencing on clustered arrays, provides specific advantages because the polymerase is able to incorporate

nucleotides into multiple DNA templates located in close proximity, thus providing a highly efficient reaction.

The above components are allowed to interact under conditions which permit the formation of a phosphodiester linkage between the 5' phosphate group of a nucleotide and a free 3' hydroxyl group on the DNA template, whereby the nucleotide is incorporated into a polynucleotide. Preferred nucleotides, including modified nucleotides, are described in detail above.

The incorporation reactions may occur in free solution or the DNA templates may be fixed to a solid support.

The rate of incorporation of the nucleotide exhibited by a mutant enzyme may be similar to the rate of incorporation of nucleotides exhibited by the unaltered enzyme. Due to the improved activity of the modified enzyme, thanks to its reduced affinity for DNA, the same rate of incorporation combined with the ability to incorporate nucleotides into a plurality of templates in a single reaction cycle improves the overall rates of completion. However, it is not necessary for the rate of incorporation of nucleotides to be precisely the same to that of the unaltered enzyme for a mutant enzyme to be of practical use. The rate of incorporation may be less than, equal to or greater than the rate of incorporation of nucleotides by the unaltered enzyme, provided the overall reaction efficiency in terms of reaction completion is improved.

In one particular embodiment of the invention, the altered polymerases of the invention may be used to incorporate modified nucleotides into a polynucleotide chain in the context of a sequencing-by-synthesis protocol. In this particular aspect of the method the nucleotides may

have been modified at the 3' sugar hydroxyl such that the substituent is larger in size than the naturally occurring 3' hydroxyl group. These nucleotides are detected in order to determine the sequence of a DNA template.

Thus, in a still further aspect, the invention provides a method of sequencing DNA comprising allowing the following components to interact:

- A polymerase according to the present invention (as described above)
- a DNA template; and
- a nucleotide solution containing the nucleotides which have been modified at the 3' sugar hydroxyl such that the substituent is larger in size than the naturally occurring 3' hydroxyl group

followed by detection of the incorporated modified nucleotides thus allowing sequencing of the DNA template.

The DNA template for a sequencing reaction will typically comprise a double-stranded region having a free 3' hydroxyl group which serves as a primer or initiation point for the addition of further nucleotides in the sequencing reaction. The region of the DNA template to be sequenced will overhang this free 3' hydroxyl group on the complementary strand. The primer bearing the free 3' hydroxyl group may be added as a separate component (e.g. a short oligonucleotide) which hybridises to a region of the template to be sequenced. Alternatively, the primer and the template strand to be sequenced may each form part of a partially self-complementary nucleic acid strand capable of forming an intramolecular duplex, such as for example a hairpin loop structure. Nucleotides are added successively to the free 3' hydroxyl group, resulting in synthesis of a

polynucleotide chain in the 5' to 3' direction. After each nucleotide addition the nature of the base which has been added will be determined, thus providing sequence information for the DNA template.

Such DNA sequencing may be possible if the modified nucleotides can act as chain terminators. Once the modified nucleotide has been incorporated into the growing polynucleotide chain complementary to the region of the template being sequenced there is no free 3'-OH group available to direct further sequence extension and therefore the polymerase can not add further nucleotides. Once the nature of the base incorporated into the growing chain has been determined, the 3' block may be removed to allow addition of the next successive nucleotide. By ordering the products derived using these modified nucleotides it is possible to deduce the DNA sequence of the DNA template. Such reactions can be done in a single experiment if each of the modified nucleotides has attached a different label, known to correspond to the particular base, to facilitate discrimination between the bases added at each incorporation step. Alternatively, a separate reaction may be carried out containing each of the modified nucleotides separately.

In a preferred embodiment the modified nucleotides carry a label to facilitate their detection. Preferably this is a fluorescent label. Each nucleotide type may carry a different fluorescent label. However the detectable label need not be a fluorescent label. Any label can be used which allows the detection of the incorporation of the nucleotide into the DNA sequence.

One method for detecting the fluorescently labelled nucleotides, suitable for use in the methods of the invention, comprises using laser light of a wavelength

specific for the labelled nucleotides, or the use of other suitable sources of illumination.

In one embodiment, the fluorescence from the label on the nucleotide may be detected by a CCD camera.

If the DNA templates are immobilised on a surface they may preferably be immobilised on a surface to form a high density array, which is preferably a clustered or "colonial" array as discussed supra. In one embodiment, and in accordance with the technology developed by the applicants for the present invention, the high density array comprises a single molecule array, wherein there is a single DNA molecule at each discrete site that is detectable on the array. Single-molecule arrays comprised of nucleic acid molecules that are individually resolvable by optical means and the use of such arrays in sequencing are described, for example, in WO 00/06770, the contents of which are incorporated herein by reference. Single molecule arrays comprised of individually resolvable nucleic acid molecules including a hairpin loop structure are described in WO 01/57248, the contents of which are also incorporated herein by reference. The polymerases of the invention are suitable for use in conjunction with single molecule arrays prepared according to the disclosures of WO 00/06770 of WO 01/57248. However, it is to be understood that the scope of the invention is not intended to be limited to the use of the polymerases in connection with single molecule arrays.

Single molecule array-based sequencing methods may work by adding fluorescently labelled modified nucleotides and an altered polymerase to the single molecule array. Complementary nucleotides base-pair to the first base of each nucleotide fragment and are then added to the primer in

a reaction catalysed by the improved polymerase enzyme. Remaining free nucleotides are removed.

Then, laser light of a specific wavelength for each modified nucleotide excites the appropriate label on the incorporated modified nucleotides, leading to the fluorescence of the label. This fluorescence may be detected by a suitable CCD camera that can scan the entire array to identify the incorporated modified nucleotides on each fragment. Thus millions of sites may potentially be detected in parallel. Fluorescence may then be removed.

The identity of the incorporated modified nucleotide reveals the identity of the base in the sample sequence to which it is paired. The cycle of incorporation, detection and identification may then be repeated approximately 25 times to determine the first 25 bases in each oligonucleotide fragment attached to the array, which is detectable.

Thus, by simultaneously sequencing all molecules on the array, which are detectable, the first 25 bases for the hundreds of millions of oligonucleotide fragments attached in single copy to the array may be determined. Obviously the invention is not limited to sequencing 25 bases. Many more or less bases may be sequenced depending on the level of detail of sequence information required and the complexity of the array.

Using a suitable bioinformatics program the generated sequences may be aligned and compared to specific reference sequences. This allows determination of any number of known and unknown genetic variations such as single nucleotide polymorphisms (SNPs) for example.

The utility of the altered polymerases of the invention is not limited to sequencing applications using single-

molecule arrays. The polymerases may be used in conjunction with any type of array-based (and particularly any high density array-based) sequencing technology requiring the use of a polymerase to incorporate nucleotides into a polynucleotide chain, and in particular any array-based sequencing technology which relies on the incorporation of modified nucleotides having large 3' substituents (larger than natural hydroxyl group), such as 3' blocking groups.

The polymerases of the invention may be used for nucleic acid sequencing on essentially any type of array formed by immobilisation of nucleic acid molecules on a solid support. In addition to single molecule arrays suitable arrays may include, for example, multi-polynucleotide or clustered arrays in which distinct regions on the array comprise multiple copies of one individual polynucleotide molecule or even multiple copies of a small number of different polynucleotide molecules (e.g. multiple copies of two complementary nucleic acid strands).

In particular, the polymerases of the invention may be utilised in the nucleic acid sequencing method described in WO 98/44152, the contents of which are incorporated herein by reference. This International application describes a method of parallel sequencing of multiple templates located at distinct locations on a solid support. The method relies on incorporation of labelled nucleotides into a polynucleotide chain.

The polymerases of the invention may be used in the method described in International Application WO 00/18957, the contents of which are incorporated herein by reference. This application describes a method of solid-phase nucleic acid amplification and sequencing in which a large number of distinct nucleic acid molecules are arrayed and amplified

simultaneously at high density via formation of nucleic acid colonies and the nucleic acid colonies are subsequently sequenced. The altered polymerases of the invention may be utilised in the sequencing step of this method.

Multi-polynucleotide or clustered arrays of nucleic acid molecules may be produced using techniques generally known in the art. By way of example, WO 98/44151 and WO 00/18957 both describe methods of nucleic acid amplification which allow amplification products to be immobilised on a solid support in order to form arrays comprised of clusters or "colonies" of immobilised nucleic acid molecules. The contents of WO 98/44151 and WO 00/18957 relating to the preparation of clustered arrays and use of such arrays as templates for nucleic acid sequencing are incorporated herein by reference. The nucleic acid molecules present on the clustered arrays prepared according to these methods are suitable templates for sequencing using the polymerases of the invention. However, the invention is not intended to be limited to use of the polymerases in sequencing reactions carried out on clustered arrays prepared according to these specific methods.

The polymerases of the invention may further be used in methods of fluorescent *in situ* sequencing, such as that described by Mitra *et al.* Analytical Biochemistry 320, 55-65, 2003.

The present invention also contemplates kits which include the polymerase of the invention, possibly packaged together with suitable instructions for use. The polymerase will be provided in a form suitable for use, for example provided in a suitable buffer or may be in a form which can be reconstituted for use (e.g. in a lyophilized form).

Thus, a kit is provided for use in a nucleotide incorporation reaction or assay comprising a polymerase of the invention as described herein and a solution of nucleotides, the nucleotides being such that the polymerase can incorporate them into a growing DNA strand. Preferred nucleotides include suitably labelled nucleotides which can thus be used in sequencing reactions for example. Labels may include fluorescent labels, radiolabels and/or mass labels as are well known in the art.

In one preferred embodiment, the nucleotide solution comprises, consists essentially of or consists of synthetic (i.e. non-natural) nucleotides such as ddNTPs for example. The kit may thus be utilised in a Sanger sequencing reaction for example.

In a further embodiment, the nucleotide solution comprises, consists essentially of or consists of modified nucleotides. Preferred modified nucleotides are defined above with respect to the polymerases of the invention and this description applies *mutatis mutandis* here.

The kit may, in a further embodiment, also incorporate suitable primer and/or DNA template molecules which allow a nucleotide incorporation reaction to be carried out.

In a still further aspect, the invention provides a method for producing a polymerase according to the invention comprising:

(i) selecting residues for mutagensis in the polymerase;

(ii) producing a mutant polymerase in accordance with the selection made in (i);

(iii) determining the affinity of the mutant polymerase for DNA; and

(iv) if the affinity for DNA is reduced, testing the polymerase for an ability to form an increased number of productive polymerase-DNA complexes in each reaction cycle.

Preferably affinity for nucleotides is unaffected, but may be considered satisfactory if it remains of the same order as for the unmodified polymerase.

In one embodiment, the method further comprises ensuring that the fidelity of the polymerase remains of the same order following mutagenesis.

Preferably fidelity is unaffected, but may be considered acceptable if it remains of the same order as for the modified polymerase.

Reaction cycle is as defined above.

In a preferred embodiment, the test of the polymerase includes the use of synthetic nucleotides to determine whether an increased number of productive polymerase-DNA complexes are being formed. Suitable nucleotide incorporation assays in which the polymerase may be tested are known in the art (e.g. see WO2005/024010) and are described in more detail in the experimental section below.

In one embodiment, residues are selected on the basis of the 9°N primary amino acid sequence. In one embodiment, the selection is made by predicting which amino acids will contact the DNA. Alternatively, residues may be selected which are predicted to stabilise the interaction of the polymerase with DNA and/or which are found in the DNA binding domain of the polymerase and/or which are basic. Predictions may be based on crystal structures of a suitable polymerase, as discussed supra and in the experimental section (example 1).

Methods of mutagenesis, in particular site-directed mutagenesis, are well characterised in the art and kits are

commercially available. Accordingly, these techniques are not discussed in detail. Any suitable technique may be utilized in the method of the invention.

The reduction in affinity for DNA may be measured by any suitable method. Preferably, the affinity is reduced at least, or approximately, 1.5-fold, 2-fold, 3-fold, 4-fold or 5-fold etc. compared to the original unaltered polymerase. This affinity may be measured with reference to the dissociation constant for example.

In a preferred embodiment, the polymerase is a family B polymerase, preferably derived from a thermophilic archaeon and most preferably is 9°N polymerase.

BRIEF DESCRIPTION OF THE DRAWINGS

The invention will be further understood with reference to the following experimental section and figures in which:

Figure 1 shows overexpression of mutant enzymes of the invention.

Figure 2 shows results of a NUNC tube assay using crude preparations of the mutant enzymes.

Figure 3 shows results of a single base incorporation assay utilising the mutant enzymes.

Figure 4 is a further presentation of the activity of the mutant enzymes.

Figure 5 presents results of timecourses for a single base incorporation assay at [DNA] > [pol] (ratio 5:1) for the control polymerase (YAV) and for each of the three mutant enzymes (K705A, R713A and R743A).

Figure 6 shows results of the washing assay, with the fluorescence image of the NUNC wells shown.

Figure 7 also presents results of the washing assay, showing the affinity of the respective polymerases for a DNA template (the data for K705A has been omitted for clarity).

Figure 8 represents Michaelis plots showing the kinetic characterization of the polymerase enzymes, which are shown overlaid.

Figure 9 presents the nucleotide and amino acid sequences encoded by the codon-modified gene of clone 9

Figure 10 presents the results of SDS-PAGE experiments comparing the expression of 3 clones (1,2 and 4) of Pol 52 (codon-modified gene of clone 9 when expressed in pET11-a expression vector) with Pol 19 (clone 9 gene expressed from the pNEB917 expression vector) and Pol 43 (clone 9 gene expressed from the pET11-a expression vector) in the crude lysates of uninduced (gel I) and induced (Gel I,II) cultures.

Abbreviations: MW-Molecular Weight; PM-protein marker; cl 9-clone 9

DETAILED DESCRIPTION OF THE INVENTION

Experimental Section

Example 1 - Preparation of altered polymerases

Rationale

Site- directed mutations were introduced in the C-terminal region of 9°N-7 YAV C223S polymerase in an attempt to reduce the affinity of the enzyme for DNA (wild-type 9°N-7 polymerase has a very high affinity for DNA, Kd = 50 pM; Southworth et al. 1996. PNAS. 93, 5281).

An energy minimised overlaid alignment (contracted by Cresset) of the crystal structures of the open form of 9°N-7

DNA polymerase (PDB = 1qht), the open structure of a closely related DNA polymerase RB69 (PDB = 1ih7) and the closed form of RB69 (PDB = 1ig9) was used as a structural model for the identification of key residues involved in DNA binding. The crystal structure of the closed form of RB69 polymerase (Franklin et al. 2001. Cell 105, 657) identified a number of residues that formed H-bond or electrostatic interactions with the complexed DNA, either directly to the nucleotide bases or the phosphate backbone. A high proportion of these residues were basic (Lys790, 800, 844, 874, 878 and Arg806), consistent with their likely interaction with acidic phosphate groups. Inspection of the closed RB69 structure showed that the majority of these residues adopted orientations toward the bound duplex. No analogous structure for the closed form of 9°N-7 pol exists and so we used our structural alignment to identify basic residues in the open form of 9°N-7 pol which adopted analogous conformations to the basic residues (of those above) from the RB69 open structure. Of the 6 basic residues from RB69, 3 were found to have a corresponding basic residue in 9°N-7, these were: Arg743 (RB69 Lys878), Arg713 (Lys800) and Lys705 (Lys844). It was decided to engineer 4 mutant enzymes, the alanine variants of the residues shown (R743A, R713A and K705A) and a 71 amino acid deletion (Δ71), which removed an α-helix from the thumb subdomain (residues disordered in the 9°N-7 pol structure) within which the three residues above were located.

Mutagenesis and cloning

Mutations were introduced into pSV19 (plasmid encoding 9°N-7 YAV C223S exo- polymerase) via a PCR method using

Stratagene Quikchange XL kit and the protocol thereof (also see WO 2005/024010)

Mutagenic primers used:

- R743A.

fwd 5'-CCCGGCGGTGGAGGCGATTCTAAAAGCC-3' (SEQ ID NO: 9)

rev 3'-GGGCCGCCACCTCCGCTAAGATTTTCGG-5' (SEQ ID NO: 10)

- R713A

fwd 5'-GAAGGATAGGCGACGCGGCGATTCCAGCTG-3' (SEQ ID NO: 11)

rev 3'-CTTCCTATCCGCTGCGCCGCTAAGGTCGAC-5' (SEQ ID NO: 12)

- K705A

fwd 5'-GCTACATCGTCCTAGCGGGCTCTGGAAGG-3' (SEQ ID NO: 13)

rev 3'-CGATGTAGCAGGATCGCCCGAGACCTTCC-5' (SEQ ID NO: 14)

- Δ71 (C-terminus 704)

fwd 5'-GCTACATCGTCCTATGAGGCTCTGGAAGG-3' (SEQ ID NO: 15)

rev 3'-CGATGTAGCAGGATACTCCGAGACCTTCC-5' (SEQ ID NO: 16)

Potential clones were selected and PCR fragments of the gene sequenced to confirm the presence of the mutation.
Positive clones were produced for all mutants.

Overexpression and growth:

- Transformed into expression strain Novagen RosettaBlue DE3 pLysS
- Growth and induction carried out as described in Experimental section of WO 2005/024010.
- Harvest and lysis carried out as described in Experimental section of WO 2005/024010.
- Purification carried out as described in Experimental section of WO 2005/024010.

Results:

- Successful overexpression of mutant enzymes was acheieved. All mutant enzymes were overexpressed. SDS-PAGE

gels were run to check overexpression of the constructs (- = uninduced; + = IPTG induced). The resulting gels are shown in Figure 1.

Example 2 - NUNC tube assay using crude protein preparation.

Small 5 ml cultures of the mutant enzymes (along with a culture of YAV C223S exo- for direct comparison) were taken through a quick purification as outlined in WO 2005/024010 up until the heat treatment step. At this point, the samples were considered to be sufficiently pure to test their activity.

The buffers for each of the crude preparations were exchanged into enzymology buffer (50 mM Tris pH 8.0, 6 mM MgSO4, 1 mM EDTA, 0.05% Tween20) using an S300 gel filtration spin-column. The samples were not normalised for concentration. The test employed was a simple incorporation of ffTTP into surface-coupled A-template hairpin. 2 pmoles of 5'-amino oligo 815

(5'-CGATCACGATCACGATCACGATCACGATCACGATCACGCTGATGTGCATGCTGTTG TTTTTTTACAACAGCATGCACATCAGCG-3') (SEQ ID NO: 17)
was coupled to a NUNC-nucleolink strip according to the manufacturers protocol.

Once washed, each well was incubated with a 20 µl aliquot of a crude enzyme preparation (identity of enzyme listed below) and 5 µM ffT-N3-647. The strip was then incubated at 45 °C for 30 minutes. The experiment was performed in duplicate. Upon completion of the 30 minute incubation, wells were washed with 3 x 100 µl of high salt wash buffer (10 mM Tris pH 8.0, 1M NaCl, 10 mM EDTA) and

then 3 x 100 µl of MilliQ water. Strips were scanned on a typhoon fluorescence imager CY5 filter, PMT = 450 V).

The results are presented in Figure 2, in which the wells are as follows:

1 = 20 µl enzymology buffer only + 1 µl 100 µM ffT-N3-647

2 = 20 µl crude YAV C223S exo- + 1 µl 100 µM ffT-N3-647

3 = 20 µl crude YAV C223S R743A exo- (clone 12) + 1 µl 100 µM ffT-N3-647

4 = 20 µl crude YAV C223S K705A exo- (clone 15) + 1 µl 100 µM ffT-N3-647

5 = 20 µl crude YAV C223S R743A exo- (clone 16) + 1 µl 100 µM ffT-N3-647

6 = 20 µl crude YAV C223S R713A exo- (clone 24) + 1 µl 100 µM ffT-N3-647

7 = 20 µl crude YAV C223S Δ71 exo- (clone 38) + 1 µl 100 µM ffT-N3-647

8 = 20 µl crude YAV C223S R713A exo- (clone 39) + 1 µl 100 µM ffT-N3-647

Results

Enzymology was observed in all wells except the background wells (MilliQ only) and well 1 (no enzyme control). The fluorescence density is proportional to the amount of ffTTP incorporation - the darker the well, the greater the level of incorporation. Performance of the mutant enzymes will be discussed relative to YAV (clone 9)(YAV C223S exo-). Deletion of the tip of the thumb subdomain (Δ71 mutant) results in an enzyme that is severely catalytically compromised, and only incorporates to 35 % of the level seen for clone 9. Mutant K705A was equivalent to clone 9. The two arginine mutants R743A and R713A showed

elevated levels of incorporation, showing ~ 45% improvements over clone 9.

Conclusion

Mutant enzymes K705A, R713A and R743A display improved levels of incorporation and decreased affinity of the enzyme for DNA. Removal of all three of these basic residues, in combination with deletion of additional residues, abolishes activity (Δ71 mutant). It may be that substitution of all three residues would not lead to a decrease in activity, in the absence of further mutations/deletions.

Example 3 - Single base incorporation assay

The activity of the crude enzyme preparations (normalised concentrations) was measured using the single base incorporation assay as described in WO 2005/024010. 10 minute incubations were run with either 30 or 3 µg/ml crude enzyme preparation in the presence of 2 µM ffT-N3-cy3 and 20 nM 10A hairpin DNA (^{32}P-labelled), aliquots of the reaction mixture were withdrawn at 0, 30, 60, 180 and 600s and run on a 12 % acrylamide gel.

Results

The Gel images are shown in Figure 3.

The band intensities were quantified using Imagequant and the fluorescence intensity plotted versus incubation time to generate the time-courses shown in Figure 4.

These data give an estimate of the performance of the mutant enzymes for the first base incorporation of ffTTP relative to YAV. Due to the concentration normalisation, the activities are directly comparable. The Δ71 mutant is

essentially inactive (kobs is 21 % of that observed for YAV), R743A and K705A have comparable activities to YAV, but R713A shows a significant enhancement in both kobs (2x that observed for YAV) and the level of cycle completion.

Example 4 - Single base incorporation assay for purified polymerases under conditions where [DNA] is greater than [pol].

The activity of the purified enzyme preparations of Clone 9 polymerase (YAV C223S exo-) and the thumb sub-domain mutants K705A, R713A and R743A was measured using the single base incorporation assay as described in WO 2005/024010. The experiment was carried out such that the respective concentrations of DNA and polymerase were at a ratio of approximately 5:1. Thus, the ability of the enzyme to incorporate nucleotides into multiple DNA template molecules in a single reaction cycle was investigated. 30 minute incubations were run with 4 nM purified enzyme in the presence of 20 nM 10A hairpin DNA (^{32}P-labelled) and 2 μM ffT-N3-cy3, aliquots of the reaction mixture were withdrawn at 0, 15, 30, 60, 180, 480, 900 and 1800s intervals and run on a 12 % acrylamide gel.

Results

The band intensities were quantified using Imagequant and the fluorescence intensity, converted into percentage completion (based on the relative intensities of the starting material and final product bands on the gel) plotted versus incubation time to generate the timecourses shown in Figure 5.

Timecourse plots for clone 9 and K705A are biphasic in nature, displaying an initial exponential "burst" phase (black line) followed by a linear dependence of product conversion with time (grey line). The amplitude of the burst phase is greater for K705A than for clone 9 (~ 28 % and 19 % respectively) and the gradient of the linear phase is steeper (hence faster) for K705A than clone 9. The significance of this observation is discussed below.

In contrast to this, both R713A and R743A mutant enzymes do not show this biphasic nature, instead, only the fast exponential phase is observed. In both cases, the amplitude of the exponential phase is ~ 90 % indicating a higher degree of product conversion within this exponential phase than either clone 9 or K705A. The burst phase equates to the rate of incorporation of ffTTP of the population of DNA molecules associated with a polymerase prior to reaction initiation i.e. maximum rate at which the ternary pol:DNA:ffTTP complex can turnover. Any subsequent phase is attributed to a slower dissociation/re-association process required for the polymerase to sequester new substrate molecules (DNA and ffTTP). The biphasic nature observed for clone 9 and K705A suggests that the slow post-burst phase is caused by the difficulty of the enzyme to dissociate and re-associate with DNA, most likely due to their low Kd(DNA).

The mutation of basic residues that may contact duplex DNA when bound by the polymerase (namely R713 and R743) to remove this functionality results in mutant enzymes which only display burst kinetics (R713A and R743A). We interperet this in one of two ways, i) as having improved the enzymes ability to dissociate and re-associate with DNA by decreasing the affinity for DNA (increased Kd(DNA)) and/or ii) the decrease in affinity for DNA in these mutants

results in a larger "active enzyme" fraction in the polymerase preparation. It has been shown that impure DNA polymerase (contaminated with E. coli genomic DNA carried through from lysis) inhibits the enzyme by reducing the active enzyme fraction of the preparation.

The crude fitting of the timecourses suggests that the observed rate constants for the burst phase seen for clone 9 and K705A are comparable (kobs ~ 0.06 s-1) whereas this rate constant is smaller for R713A (kobs ~ 0.01 s-1) and R743A (kobs ~ 0.004 s-1). Under these experimental conditions, the burst is faster for clone 9 and K705A than for R713A or R743A, but the latter two enzymes reach completion in a shorter period of time due to the absence of the slow, linear dissociation/re-association phase inherent to clone 9 and K705A.

Example 5 - Washing assay.

Employing a washing assay qualitatively assesses the affinity of purified enzyme preparations for DNA. 4 (1x8) NUNC nucleolink strips were functionalized with 2 pmoles of 5'-amino A-template hairpin, oligo 815
(5' H2N-
CGATCACGATCACGATCACGATCACGATCACGATCACGCTGATGTGCATGCTGTTGTTTT
TTTACAACAGCATGCACATCAGCG-3') (SEQ ID NO: 18) according to the manufacturer's protocol.

Once washed, each well was incubated with a 20 µl aliquot of 500 nM enzyme (clone 9, K705A, R713A or R743A mutants) at 45 °C for 30 minutes. Post incubation, each well was washed with 3 x 100 ml of 10 mM Tris pH 8.0, 10 mM EDTA including varying concentrations of NaCl (0, 0.05, 0.1, 0.3, 0.4, 0.75, 1.0, 2.0 M) and then 3 x 100 ml MilliQ water.

Wells were subsequently pre-equilibrated with enzymology buffer prior to a further incubation of 20 µl of 2 µM ffT-N3-647 at 45 °C for 30 minutes. Wells were washed with 3 x 100 ml high salt wash buffer (10 mM Tris pH 8.0, 1M NaCl, 10 mM EDTA) and then 3 x 100 ml MilliQ water. Strips scanned on Typhoon fluorescence imager (y5 filter, PMT = 500 V).

Results

The fluorescence image of the NUNC wells is shown in Figure 6.

Any fluorescence in the wells is due to residual enzyme bound to the surface-coupled DNA post-wash. Increasing the ionic strength of the wash buffer between incubation should destabilise the interaction between the polymerase and the DNA by masking electrostatic interactions. Enzyme should be more effectively washed off the DNA at higher ionic strength.

When a low ionic strength wash is employed between incubations all enzymes tested displayed a high level of incorporation, therefore ineffective at dissociating enzyme from DNA. As the concentration of NaCl in the wash buffer increased, the behaviour of the enzymes relative to each other changed. Mutant enzymes R713A and R743A were more effectively removed from the DNA at [NaCl] < 200 mM, whereas K705A and clone 9 showed a similar response to each other, but required higher [NaCl] to remove them from the DNA. Even after a wash with 2 M NaCl, a significant (ca. 75 %) level of incorporation relative to a 0 M NaCl wash was observed for clone 9. This is clearly illustrated in the plot shown in Figure 7 (the data for K705A has been omitted for clarity). Interestingly, none of the enzymes tested appeared

to be completely removed from the DNA after experiencing a 2 M NaCl wash.

From this experiment, it is clear that mutating residues R713 and R743 result in enzymes that display lower affinity for DNA than clone 9, as evidenced by their ability to be washed from DNA by lower ionic strength washes.

Example 6 - Incorporation kinetics of ffT-N3-cy3 by Clone 9, R713A and R743A.

The kinetic characterization of the enzymes was conducted using NUNC tube assay and involved the measurement of rate constants for the first order incorporation of ffT N3 cy3 where [DNA] << [pol] or [ffNTP], at a variety of [ffTTP]. Below is described the methodology used for each of the three polymerases tested.

Six (1x8) NUNC nucleolink strips were functionalized with 2 pmoles of 5'-amino A template hairpin oligo 815 (5' H2N-CGATCACGATCACGATCACGATCACGATCACGATCACGCTGATGTGCATGCTGTTGTTTTTTTACAACAGCATGCACATCAGCG-3') (SEQ ID NO: 18), according to the manufacturers protocol.

Each strip was employed for a time-course experiment at a particular [ffT-N3-cy3]. 20 µl of enzymology buffer (50 mM Tris pH 8.0, 6 mM MgSO4, 1 mM EDTA, 0.05 % Tween20) was incubated in each NUNC well at 45 °C for 2 minutes.

Time-courses were initiated by addition of a 20 µl aliquot of 2x enzymology mix (X µM ffT-N3-cy3, 1.1 µM polymerase in enzymology buffer) pre-equilibrated at 45 °C for 2 minutes using an 8-channel multipipette in order to start reactions in individual wells at identical time-points. The action of adding the 2x enzymology mix to the

buffer in the well is sufficient to allow adequate mixing. The reactions were stopped at desired time-points by the addition of 125 µl of 250 mM EDTA. After reactions in all 8 wells stopped, strips were washed with 3 x 100 ml high salt wash (10 mM Tris pH 8.0, 1 M NaCl, 10 mM EDTA) and then 3 x 100 ml MilliQ water and then scanned on a Typhoon fluorescence imager (Cy3 filter, PMT = 500 V). Fluorescence intensities in each well were quantified using Imagequant. Plotting the variation in Cy3 fluorescence intensity vs. time generates time-course graphs. Under our experimental conditions, these time-course plots evaluate well to a single exponential decay process (fitted to equation: $y = yo + Aexp(x/t)$) from which the reaction half life, t, is determined, the inverse of which is termed the observed rate constant kobs ($kobs = 1/t$).

The magnitude of the observed rate constant is dependent on the concentration of ffT-N3-cy3, so by repeating this experiment at different ffT-N3-cy3 concentrations a range of kobs values can be determined for a particular enzyme. The variation of kobs with ffT-N3-cy3 concentration is hyperbolic and fits well to the Michealis-Menten equation: $Vmax = (kpolx[S])/(Kd+[S])$ here S = ffT N3-cy3, according to standard enzymological analysis.
From the Michaelis plot, key values characteristic of a particular enzyme catalyzing a particular reaction can be obtained, namely kpol (defined as the rate constant for the process at infinite substrate concentration) and Kd (defined as the dissociation constant, the concentration of substrate at kpol/2). This process was repeated for clone 9, R713A and R743A mutants.

Michaelis plots for all of the enzymes are shown overlaid in Figure 8.

Results

The kinetic characteristics of ffT-N3-cy3 incorporation for the enzymes tested are summarized below.

	Clone 9	R713A	R743A
k_{pol} / s^{-1}	0.061	0.10	0.068
K_d / μM	1.72	3.32	1.92

From this, it appears as though the mutations to the DNA-binding region of the polymerases have not adversely affected either the activity of the enzymes (at high substrate concentrations, kpol approximates to Vmax) or the affinity the enzymes have for fully functional nucleotide (in this case ffT-N3-cy3, but the trend is considered to be applicable to all bases). This is an ideal situation, as the mutations have had the desired effect of modifying the DNA-binding affinity of the enzymes without affecting other key catalytic properties.

Example 7 - Purification of the polymerases and measurement of levels of carry over DNA.

DNA contamination

Pico green assay

(Molecular Probes kit, cat # P11496).

Solutions required

TE buffer

10mM Tris.HCl pH 7.5

1mM EDTA

40 mL required, 2 mL of 20 X TE buffer added to 38 mL H_20

λ DNA

Solution 1 (2µg/mL λDNA) dilute 15 µL of λ DNA with 735 λL of 1 x TE buffer.

Solution 2 (50 ng/mL λ) dilute 25 µL of λ DNA with 975 µl of 1 x TE buffer.

Standard curve

In 2 mL eppendorfs the following samples were made:

Sample λ DNA (ng)	λ DNA @ 2 mg mL (µL)	λ DNA @ 50 ng mL (µL)	glycerol storage buffer (µL)	TE (µL)
100	160		400	1040
25	40		400	1160
10	16		400	1184
2.5		160	400	1040
1		64	400	1136
0.25		16	400	1184
0.025		1.6	400	1198.4
0			400	1200

3 x 500 µL from each sample was put into 3 eppendorfs.

Enzyme samples

In 5 mL bijou bottles the following samples were made:

sample		Amount (µL)	glycerol storage buffer (µL)	TE (µL)
1	enzyme stock	400		1800
2	sample 1	1100	200	900
3	sample 2	1100	200	900
4	sample 3	1100	200	900

2 x 500 µL from each sample was put into 2 eppendorfs.

A picogreen solution was prepared; 85 µL of picogreen stock added to 17 mL of 1 x TE buffer.

500 µL of this solution was added to each of the standard curve and enzyme samples, and was mixed well by pipetting and then all samples were transfered to 1.5 mL fluorimeter cuvettes.

Using the fluorimeter

The advanced reads program of the Cary Eclipse file was utilised. The λ excitation was set to 480 nm and the λ emission was set to 520 nm, and 1000 volts were used.

Analysis

Data for the standard curve was entered into Graph pad Prism a standard curve of the formula y=ax+c was fitted. The concentration values, x, was then determined.

Results

Polymerase sample	**Concentration of DNA associated with purified polymerase**
Clone 9 batch 5	62.9 ng ± 1.9 ng
Clone 9 batch 6	63.7 ng ± 2.1 ng
Clone 9 R743A	0.04 ng ± 6.4 ng
Clone 9 R713A	8.2 ng ± 4.2 ng

From this experiment, it is clear that the alterations in the polymerases enhance purification of the enzyme since less endogenous DNA is carried over during purification. As mentioned above, carry over of endogenous DNA can adversely

influence activity of the enzyme and so the mutations are clearly advantageous.

Example 8: Preparation of a modified optimised codon usage nucleic acid sequence which encodes the clone 9 polymerase.

The amino acid sequence shown in SEQ ID NO 1 was translated into a nucleic acid sequence using the optimal nucleic acid sequence at each codon to encode for the required/desired amino-acid.

The deduced nucleic acid sequence is shown in SEQ ID NO.19.

In a similar scenario, the nucleic acid sequence presented as SEQ ID NO:20 was deduced based upon the amino-acid sequence of the polymerase presented as SEQ ID NO: 21. The polymerase having the amino acid sequence presented as SEQ ID NO: 21 comprises the R743A mutation and also carries a substitution mutation to Serine at both residues 141 and 143. Nucleic acid molecules and proteins comprising the respective nucleotide and amino acid sequences form a part of the invention.

Cloning of a codon- modified gene of clone 9 into the expression vector pET11-a using NdeI - Nhe I sites (to preserve the internal Bam H I site).

Synthesis of a codon-optimised gene of clone 9

The nucleic acid sequence of SEQ ID NO 19 was synthesized and supplied in pPCR -Script by GENEART.

The DNA and protein sequences were confirmed (results not shown).

Cloning of pSV57 (codon-modified gene of clone 9 in the pPCRScript vector) into pET11-a (hereinafter named pSV 52)

Preparation of the pET11-a vector

The pET11-a vector (Novagen catalog No. 69436-3) was digested with Nde I and Nhe I, dephosphorylated, and any undigested vector ligated using standard techniques.

The digested vector was purified on a 0.8% agarose gel and using the MinElute® Gel extraction kit protocol from Qiagen® .

The purified digested pET11-a vector was quantified using a polyacrylamide TB 4-20% gel.

Preparation of the insert (codon-modified gene of clone 9)

The codon-modified gene of clone 9 synthesized by GENEART in the pPCRSCript vector (hereinafter pSV 57) was digested with Nde I and Nhe.

The digested insert was purified on a 0.8% agarose gel and using the MinElute® Gel extraction kit protocol from Qiagen® .

The purified digested insert was quantified using a polyacrylamide TB 4-20% gel.

Ligation

The pET11-a vector and the insert were ligated (ratio 1:3) at the Nde I and Nhe I restriction sites using the Quick ligation kit (NEB, M2200S).

Transformation

2µl of the ligation mixture was used to transform XL10-gold ultracompetent cells (Stratagene catalog No 200315). PCR screening of the colonies containing the insert.

Transformants were picked and DNA minipreps of 3 positive clones of XL10-gold transformed with the ligation product were prepared. The three purified plasmids (hereinafter pSV52, clones 1,2 and 4 were sequenced at the cloning sites and all three clones were found to have the correct sequence at the cloning sites.

The minipreps were also used to transform the expression *E. coli* host BL21-CodonPlus (DE3)-RIL (Stratagene catalog No. 230245) as described below.

Southern blotting

pVent (pNEB917 derived vector), pSV43 (clone 9 in pET11a), pSV54 (codon-optimised clone in pET11-a) and pSV57 (codon-modified gene in pPCR-Script supplied by GENEART) were restricted and Southern blotted to check for cross hybridisation between the genes (results not shown).

Expression Studies of Pol 52

Transformation of pSV52 (clones 1,2 and 4) into the expression host *E. coli* BL21-CodonPlus (DE3) RIL (Stratagene catalog No 230245).

21-25ng of purified pSV52 plasmid DNA (clones 1,2 and 4) was used to transform competent cells of the expression host *E.*

coli BL21-CodonPlus (DE3) RIL (hereinafter RIL) using the manufacturer's instructions.

50 µl of each transformation was plated onto fresh Luria-Bertani (LB) agar medium containing 100 µg/ml of carbenicillin and 34 µg/ml of chloramphenicol (LBCC agar medium) and incubated overnight at 37° C.

The following glycerol stocks were also plated onto LBCC agar plates to be used as controls for the expression studies and incubated overnight at 37° C.

SOL10204:RIL-pSV19 (clone 9 in pNEB 917 vector)
SOL10354:RIL-pSV43 (clone 9 in pET11-a vector)

Production of cell pellets expressing Pol 52 and the positive controls of clone 9

Single transformed *E. coli* colonies were used to inoculate starter cultures of 3ml LBCC media in culture tubes and incubated overnight at 37° C with shaking (225rpm).

The starter cultures were diluted 1/100 into 50ml LBCC media in sterile vented Erlenmeyer flasks and incubated at 37° C with vigorous shaking (300rpm) for approximately 4 hours until OD_{600nm} was approximately 1.0.

10ml of the uninduced cultures was removed and the cells harvested (as described below).

IPTG was added to a final concentration of 1mM and the cultures induced for 2 hours at

37^0 C with vigorous shaking (300rpm).

10ml of the induced cultures was removed and the cells harvested as follows:

Induced and uninduced cells were harvested by centrifugation at 5000 X *g* for 30min at 4^0 C .

The cell pellets were washed and resuspended in 1/10th of the culture volume of 1X Phosphate Buffered Saline (PBS) and centrifuged as above.

The supernatants were decanted and the pellets stored at -20^0 C until required for the cell lysis and purification steps.

Cell Lysis and crude purification of Pol 52 and clone 9

The cell pellets were thawed and resuspended in1/50th of culture volume of 1X Wash buffer (50mM Tris-HCl pH 7.9, 50 mM glucose, 1mM EDTA) containing 4mg/ml lysozyme freshly added to the 1X buffer and incubated at room temperature for 15min.

An equal volume of 1X Lysis buffer (10mM Tris-HCl pH 7.9, 50mM KCl, 1mM EDTA,0.5% (w/v) Tween 20) containing 0.5%(w/v) Tergitol NP-40 and 1X "complete EDTA-free" proteinase inhibitor cocktail (both added freshly to the 1X Lysis buffer) was added to the cells which were gently mixed and incubated at room temperature for 30min.

The cells were heated at 80° C for 1hr in a water bath then centrifuged at 38,800 X *g*
for 30min at 4° C to remove cell debris and denatured protein..

Preparation of samples normalised for volume and SDS-PAGE analysis

The expression of Pol 52 and clone 9 DNA polymerases was assessed by analysis of the crude lysates of the uninduced and induced control samples on SDS-PAGE followed by Coomassie blue staining.

Supernatants were carefully removed and the samples normalised to volume by the addition of 50:50 (v/v) 1X Wash buffer and 1X Lysis buffer to a final volume of 370µl.

Preparation of samples for Gel I

10 µl of the normalised crude lysates (from uninduced and induced samples) were mixed with 10 µl of loading buffer containing 143mM DTT.

Preparation of samples for Gel II

Normalised crude lysates from the induced samples only were dilute 1/10 in distilled water to a final volume of 10 µl and mixed with 10 µl of loading buffer containing 143mM DTT.

All samples were heated at 70° C for 10 minutes.

SDS-PAGE

A NuPage® 4-12% Bis-Tris gel (Invitrogen catalog No NP0321BOX) was prepared according to the manufacturer's instructions.

10 µl of SeeBlue® Plus2 pre-stained proteins standard (Invitrogen catalog No LC5925) and µl of each sample were loaded and the gels run at a constant 200V for 50minutes.

The gels were stained with Coomassie blue (SimplyBlue™ Safe stain, Invitrogen, catalog No. LC 6060).

Results

The results of the SDS-PAGE are shown in figure 10.
The estimated expression level in this experiment is 20mg/L of culture.
Similar levels of expression of the codon-modified gene of clone 9 in *E. coli* host BL21-CodonPlus (DE3)-RIL (Pol52) were obtained using the expression vector pET11-a when compared to the un-modified gene of clone 9 in the same cells using either the expression vector pNEB917 (Pol19) or pET11 (Pol 43).

No significant differences were observed in the levels of expression of the 3 different clones of Pol 52.

References

• Crystal structure of a bacteriophage T7 DNA replication complex at 2.2 Å resolution.
Doublie et al. 1998. Nature 391, 251.

• Function of the C-terminus of Phi29 DNA polymerase in DNA and terminal protein binding.
Truniger et al. 2004. Nucleic Acids Research 32, 371.

• A thumb subdomain mutant of the large fragment of Escherichia coli DNA polymerase I with reduced DNA binding affinity, processivity and frameshift fidelity.
Minnick et al. 1996. J. Biol. Chem., 271. 24954.

• Identification of residues critical for the polymerase activity of the Klenow fragment of DNA polymerase I from Escherichia coli.
Polesky et al. 1990. J. Biol. Chem., 265, 14579.

• Cloning of thermostable DNA polymerases from hyperthermophilic marine archaea with emphasis on Thermococcus sp. 9°N-7 and mutations affecting 3'-5' exonuclease activity.
Southworth et al. 1996. PNAS. 93, 5281

• Structure of the replicating complex of a pol alpha family DNA polymerase.
Franklin et al. 2001. Cell 105, 657.

• Crystal structure of a pol alpha family DNA polymerase from the hyperthermophilic archaeon Thermococcus sp. 9°N-7.
Rodriguez et al. 2000. J. Mol. Biol., 299, 471.

Claims

1. An altered polymerase having a reduced affinity for DNA such that the polymerase has an ability to incorporate a nucleotide or nucleotides into a plurality of separate DNA templates in each reaction cycle compared to a control polymerase.

2. An altered polymerase having a reduced affinity for DNA such that the polymerase is capable of forming an increased number of productive polymerase-DNA complexes in each reaction cycle compared to a control polymerase.

3. The polymerase of claim 1 or claim 2 wherein the control polymerase is of the same type as the polymerase.

4. The polymerase of any of claims 1 to 3 wherein the control polymerase comprises the substitution mutations which are functionally equivalent to Leu408Tyr and Tyr409Ala and Pro410Val in the 9°N DNA polymerase amino acid sequence.

5. The polymerase of claim 4 wherein the control polymerase further comprises the substitution mutation functionally equivalent to Cys223Ser in the 9°N DNA polymerase amino acid sequence.

6. The polymerase of claim 4 or 5 wherein the control polymerase is a 9°N DNA polymerase.

7. The polymerase of any preceding claim wherein the alteration comprises at least one mutation at a residue in

the polymerase which stabilises the interaction of the polymerase with DNA.

8. The polymerase of any preceding claim wherein the alteration comprises at least one mutation at a residue in the polymerase which binds to DNA.

9. The polymerase of any preceding claim wherein the alteration comprises at least one mutation at a residue found in the DNA binding domain of the polymerase.

10. The polymerase of any preceding claim wherein the alteration comprises at least one mutation at a basic amino acid residue in the polymerase.

11. The polymerase of any preceding claim wherein the alteration comprises at least one substitution mutation.

12. The polymerase of claim 11 comprising two substitution mutations.

13. The polymerase of any preceding claim wherein the polymerase is a DNA polymerase.

14. The polymerase according to claim 13 wherein the DNA polymerase is a family B type DNA polymerase.

15. The polymerase according to claim 14 which is selected from any family B archael DNA polymerase, human DNA polymerase α or T4, RB69 and ϕ29 phage DNA polymerases.

16. The family B archael DNA polymerase according to claim 15 which is selected from Vent, Deep Vent, 9°N and Pfu polymerase.

17. The polymerase according to claim 16 wherein the family B archael DNA polymerase is 9°N polymerase.

18. The polymerase according to any preceding claim wherein the affinity of the polymerase for nucleotides and/or the fidelity of the polymerase is substantially unaffected by the alteration.

19. The polymerase according to any preceding claim which comprises one, two or three amino acid substitution mutations to a different amino acid at the position or positions functionally equivalent to Lys705, Arg713 and/or Arg743 in the 9°N DNA polymerase amino acid sequence.

20. An altered polymerase which comprises one, two or three amino acid substitution mutations to a different amino acid at the position or positions functionally equivalent to Lys705, Arg713 and/or Arg743 in the 9°N DNA polymerase amino acid sequence.

21. The polymerase according to claim 20 which comprises at least a substitution mutation to a different amino acid at the position functionally equivalent to either Arg713 or Arg743 in the 9°N DNA polymerase amino acid sequence.

22. The altered polymerase of any preceding claim comprising one or two amino acid substitution mutations to a different amino acid at the position or positions

functionally equivalent to Arg606 and/or His 679 in the 9°N DNA polymerase amino acid sequence.

23. An altered polymerase comprising one or two amino acid substitution mutations to a different amino acid at the position or positions functionally equivalent to Arg606 and/or His 679 in the 9°N DNA polymerase amino acid sequence.

24. An altered polymerase comprising one, two, three, four, five or six amino acid substitution mutations to a different amino acid at the position or positions functionally equivalent to Lys790, Lys800, Arg806, Lys844, Lys874 and/or Lys878 in the RB69 DNA polymerase amino acid sequence.

25. The polymerase according to any one of claims 19 to 24 wherein the substitution mutation or mutations convert the substituted amino acid to a non-basic amino acid.

26. The polymerase according to claim 25 wherein the substitution mutation or mutations convert the substituted amino acid to an amino acid selected from:

(i) acidic amino acids,

(ii) aromatic amino acids, particularly tyrosine (Y) or phenylalanine (F); and

(iii) non-polar amino acids, particularly, alanine (A), glycine (G) or methionine (M).

27. The polymerase according to claim 26 wherein the substitution mutation or mutations convert the substituted amino acid to alanine.

28. An altered polymerase comprising the substitution mutation or mutations which are functionally equivalent to Lys705Ala and/or Arg713Ala and/or Arg743Ala in the 9°N DNA polymerase amino acid sequence.

29. The altered polymerase of claim 28 which comprises the amino acid substitution functionally equivalent to Arg713Ala.

30. The altered polymerase of claim 28 or 29 which comprises the amino acid substitution functionally equivalent to Arg743Ala.

31. The altered polymerase of any preceding claim further comprising one or more amino acid substitution mutations at the position or positions which are functionally equivalent to Leu408 and/or Tyr409 and/or Pro410 in the 9°N DNA polymerase amino acid sequence.

32. The altered polymerase of claim 31 comprising the substitution mutations functionally equivalent to Leu408Tyr and Tyr409Ala and Pro410Val in the 9°N DNA polymerase amino acid sequence.

33. The altered polymerase of any preceding claim further comprising a substitution mutation to a different amino acid at the position functionally equivalent to Cys223 in the 9°N DNA polymerase amino acid sequence.

34. The altered polymerase of claim 33 wherein the polymerase comprises the substitution mutation functionally

equivalent to Cys223Ser in the 9°N polymerase amino acid sequence.

35. An altered 9°N polymerase comprising the amino acid sequence of any one of SEQ ID NO: 1, 3, 5 or 21.

36. A nucleic acid molecule encoding an altered polymerase as defined in any one of the preceding claims.

37. A nucleic acid molecule according to claim 36 which encodes an altered 9°N polymerase, the nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO: 2, 4 or 6.

38. A nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO. 19 or 20.

39. An expression vector comprising the nucleic acid molecule of claim 36 to 38.

40. A host cell containing the vector of claim 39.

41. Use of a polymerase which has been altered such that it displays a reduced affinity for DNA such that the polymerase has an ability to incorporate a labelled nucleotide into a plurality of separate DNA templates in each reaction cycle for incorporation of a labelled nucleotide into a polynucleotide, the label being utilised to determine the nature of the nucleotide added.

42. Use of a polymerase as defined in any one of claims 1 to 35 for incorporation of a nucleotide into a polynucleotide.

43. The use according to claim 41 or 42 wherein the incorporation occurs on a clustered array.

44. A method for incorporating labelled nucleotides into DNA comprising allowing the following components to interact:

(i) A polymerase which has been altered such that it displays a reduced affinity for DNA such that the polymerase has an ability to incorporate a labelled nucleotide into a plurality of separate DNA templates in each reaction cycle,

(ii) a DNA template; and

(iii) a nucleotide solution.

45. A method for incorporating nucleotides into DNA comprising allowing the following components to interact:

(i) A polymerase according to any one of claims 1 to 35,

(ii) a DNA template; and

(iii) a nucleotide solution.

46. The method according to claim 44 or 45 wherein the DNA template comprises a clustered array.

47. A kit for use in carrying out a nucleotide incorporation reaction comprising:

a polymerase as defined in any of claims 1 to 36 and a nucleotide solution.

48. The kit of claim 47 wherein the nucleotide solution comprises labelled nucleotides.

49. The kit of claim 47 or 48 wherein the nucleotides comprise synthetic nucleotides.

50. The kit of any of claims 47 to 49 wherein the nucleotides comprise modified nucleotides.

51. The kit of claim 50 wherein the nucleotides have been modified at the 3' sugar hydroxyl such that the substituent is larger in size than the naturally occurring 3' hydroxyl group.

52. The kit according to claim 51 wherein the nucleotides which have been modified at the 3' sugar hydroxyl such that the substituent is larger in size than the naturally occurring 3' hydroxyl group comprise a modified nucleotide or nucleoside molecule comprising a purine or pyrimidine base and a ribose or deoxyribose sugar moiety having a removable 3'-OH blocking group covalently attached thereto, such that the 3' carbon atom has attached a group of the structure

-O-Z

wherein Z is any of $-C(R')_2-O-R''$, $-C(R')_2-N(R'')_2$, $-C(R')_2-N(H)R''$, $-C(R')_2-S-R''$ and $-C(R')_2-F$,

wherein each R" is or is part of a removable protecting group;

each R' is independently a hydrogen atom, an alkyl, substituted alkyl, arylalkyl, alkenyl, alkynyl, aryl, heteroaryl, heterocyclic, acyl, cyano, alkoxy, aryloxy, heteroaryloxy or amido group, or a detectable label attached

through a linking group; or $(R')_2$ represents an alkylidene group of formula $=C(R''')_2$ wherein each R''' may be the same or different and is selected from the group comprising hydrogen and halogen atoms and alkyl groups; and

wherein said molecule may be reacted to yield an intermediate in which each R" is exchanged for H or, where Z is $-C(R')_2-F$, the F is exchanged for OH, SH or NH_2, preferably OH, which intermediate dissociates under aqueous conditions to afford a molecule with a free 3'OH;

with the proviso that where Z is $-C(R')_2-S-R"$, both R' groups are not H.

53. The kit according to claim 52 wherein R' of the modified nucleotide or nucleoside is an alkyl or substituted alkyl.

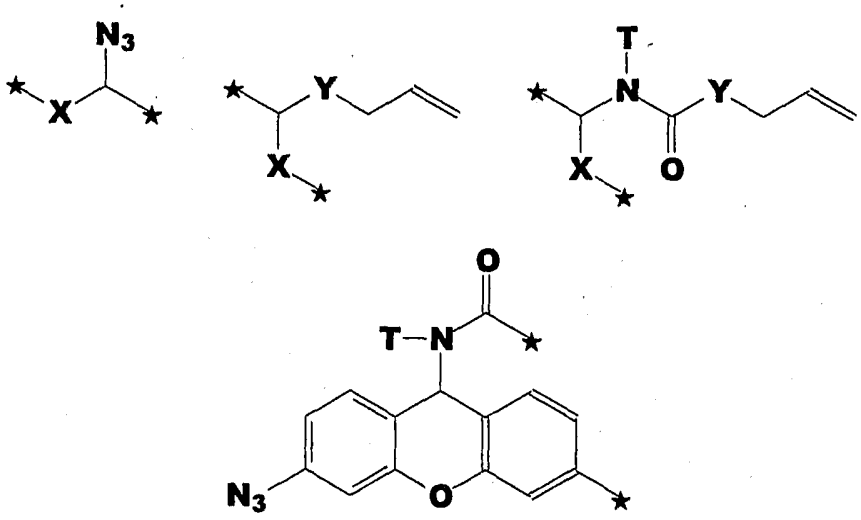
54. The kit according to claim 53 wherein -Z of the modified nucleotide or nucleoside is of formula $-C(R')_2-N_3$.

55. The kit according to claim 54 wherein Z is an azidomethyl group.

56. The kit according to claim 51 wherein the nucleotides which have been modified at the 3' sugar hydroxyl such that the substituent is larger in size than the naturally occurring 3' hydroxyl group are fluorescently labelled to allow their detection.

57. The kit according to claim 51 wherein the nucleotides which have been modified at the 3' sugar hydroxyl such that the substituent is larger in size than the naturally occurring 3' hydroxyl group comprises a nucleotide or

nucleoside having a base attached to a detectable label via a cleavable linker, characterised in that the cleavable linker contains a moiety selected from the group comprising:



(wherein X is selected from the group comprising O, S, NH and NQ wherein Q is a C_{1-10} substituted or unsubstituted alkyl group, Y is selected from the group comprising O, S, NH and N(allyl), T is hydrogen or a C_{1-10} substituted or unsubstituted alkyl group and * indicates where the moiety is connected to the remainder of the nucleotide or nucleoside).

58. The kit according to claim 57 wherein the detectable label comprises a fluorescent label

59. The kit of any of claims 47 to 58 further comprising one or more DNA template molecules and/or primers.

60. A method of producing a polymerase as defined in any of claims 1 to 35 comprising:

(i) selecting suitable residues for mutagenesis in the polymerase;

(ii) producing a mutant polymerase in accordance with the selection made in (i);

(iii) determining the affinity of the mutant polymerase for DNA; and

(iv) if the affinity for DNA is reduced, testing the polymerase for an ability to form an increased member of productive polymerase-DNA complexes in each reaction cycle.

61. An altered polymerase as hereinbefore described with reference to the accompanying figures.

62. A use as hereinbefore described with reference to the accompanying figures.

63. A method for incorporating nucleotides as hereinbefore described with reference to the accompanying figures.

64. A kit as hereinbefore described with reference to the accompanying figures.

65. A method of producing a polymerase as hereinbefore described with reference to the accompanying figures.

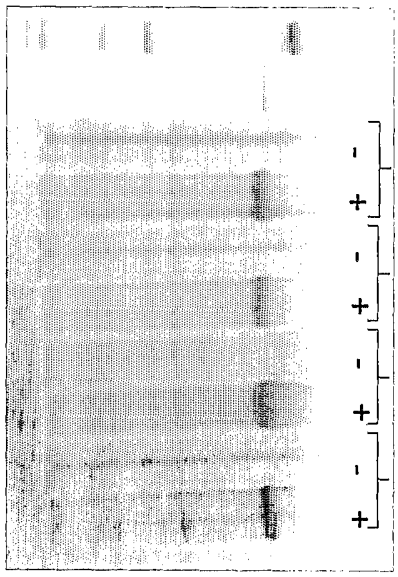


FIG. 1A

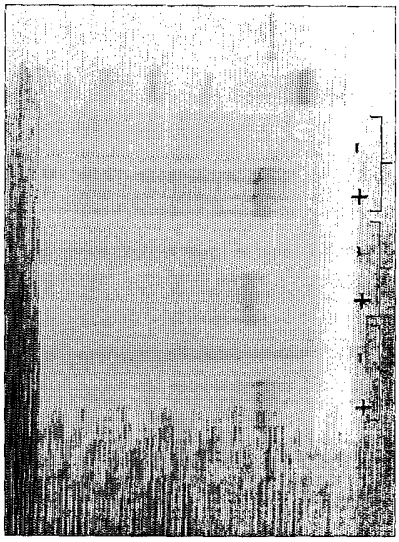


FIG. 1B

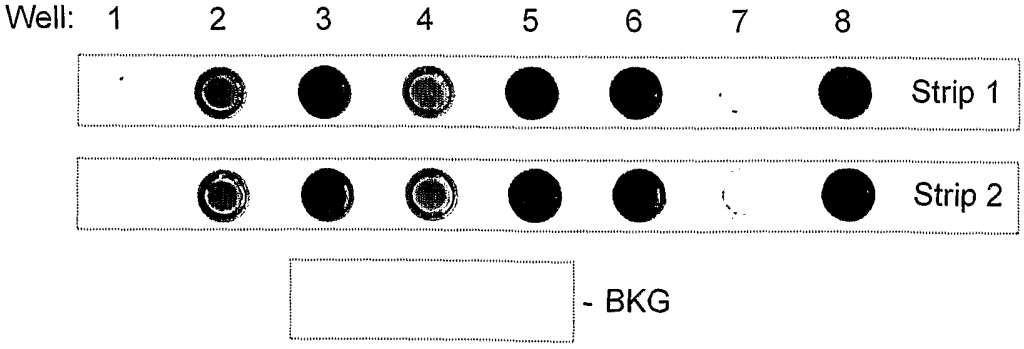


FIG. 2

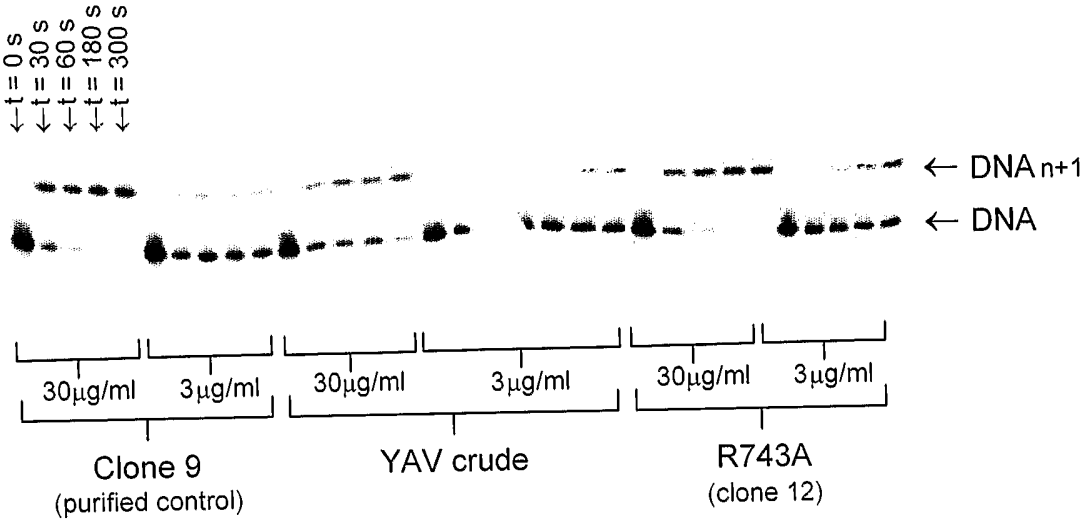

t = 0 s
t = 30 s
t = 60 s
t = 180 s
t = 300 s
DNA n+1
DNA
30µg/ml
3µg/ml
30µg/ml
3µg/ml
30µg/ml
3µg/ml
Clone 9
(purified control)
YAV crude
R743A
(clone 12)

FIG. 3A

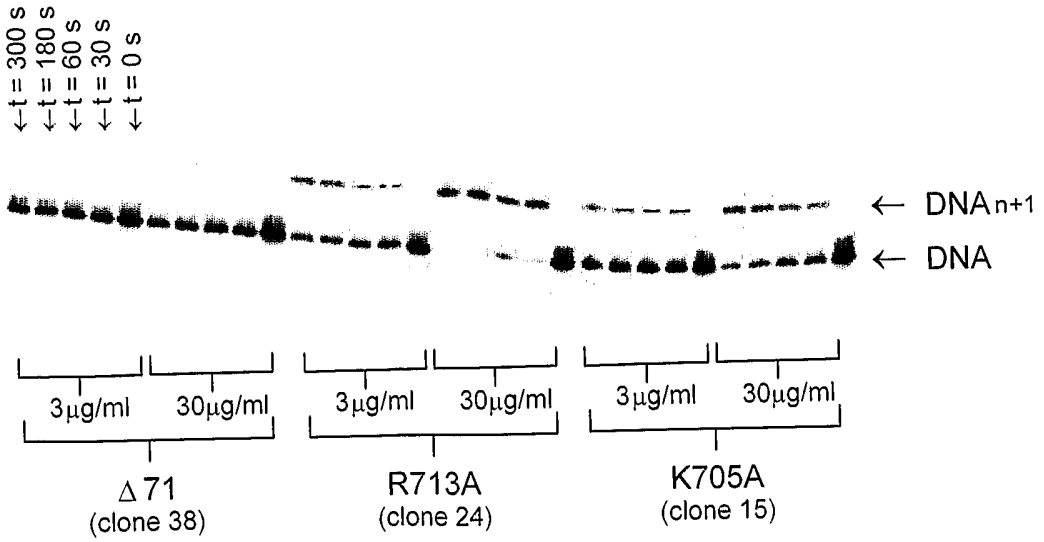

t = 300 s
t = 180 s
t = 60 s
t = 30 s
t = 0 s
DNA n+1
DNA
3µg/ml
30µg/ml
3µg/ml
30µg/ml
3µg/ml
30µg/ml
Δ71
(clone 38)
R713A
(clone 24)
K705A
(clone 15)

FIG. 3B

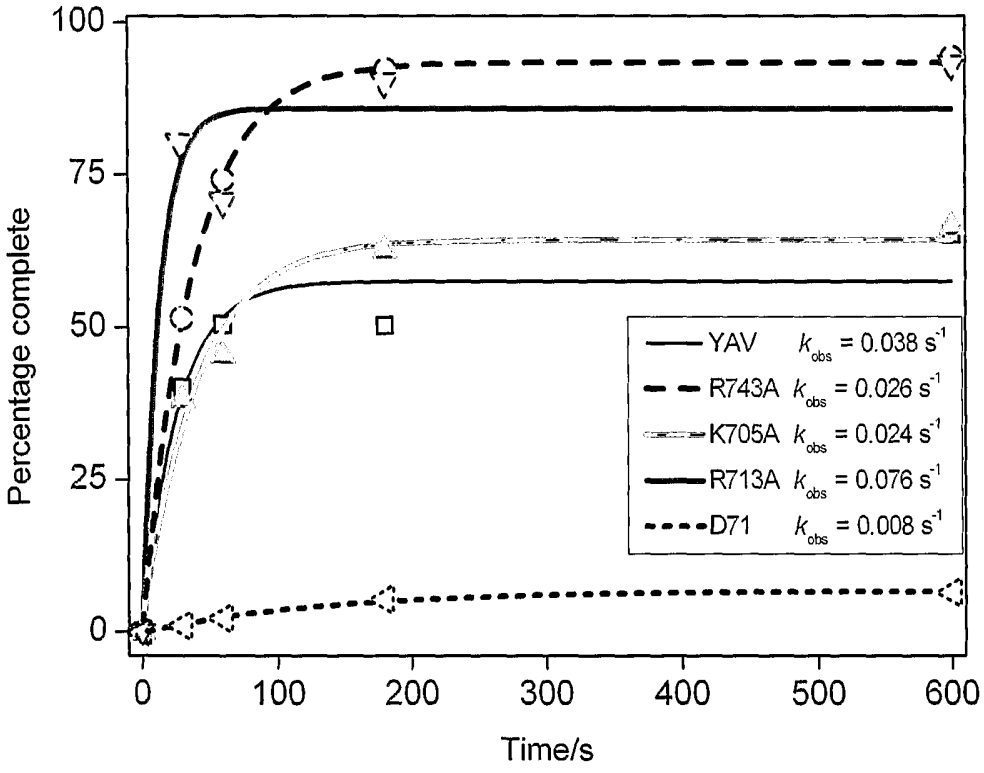

100
75
50
25
0
Percentage complete
0
100
200
300
400
500
600
Time/s
YAV k_{obs} = 0.038 s^{-1}
R743A k_{obs} = 0.026 s^{-1}
K705A k_{obs} = 0.024 s^{-1}
R713A k_{obs} = 0.076 s^{-1}
D71 k_{obs} = 0.008 s^{-1}

FIG. 4

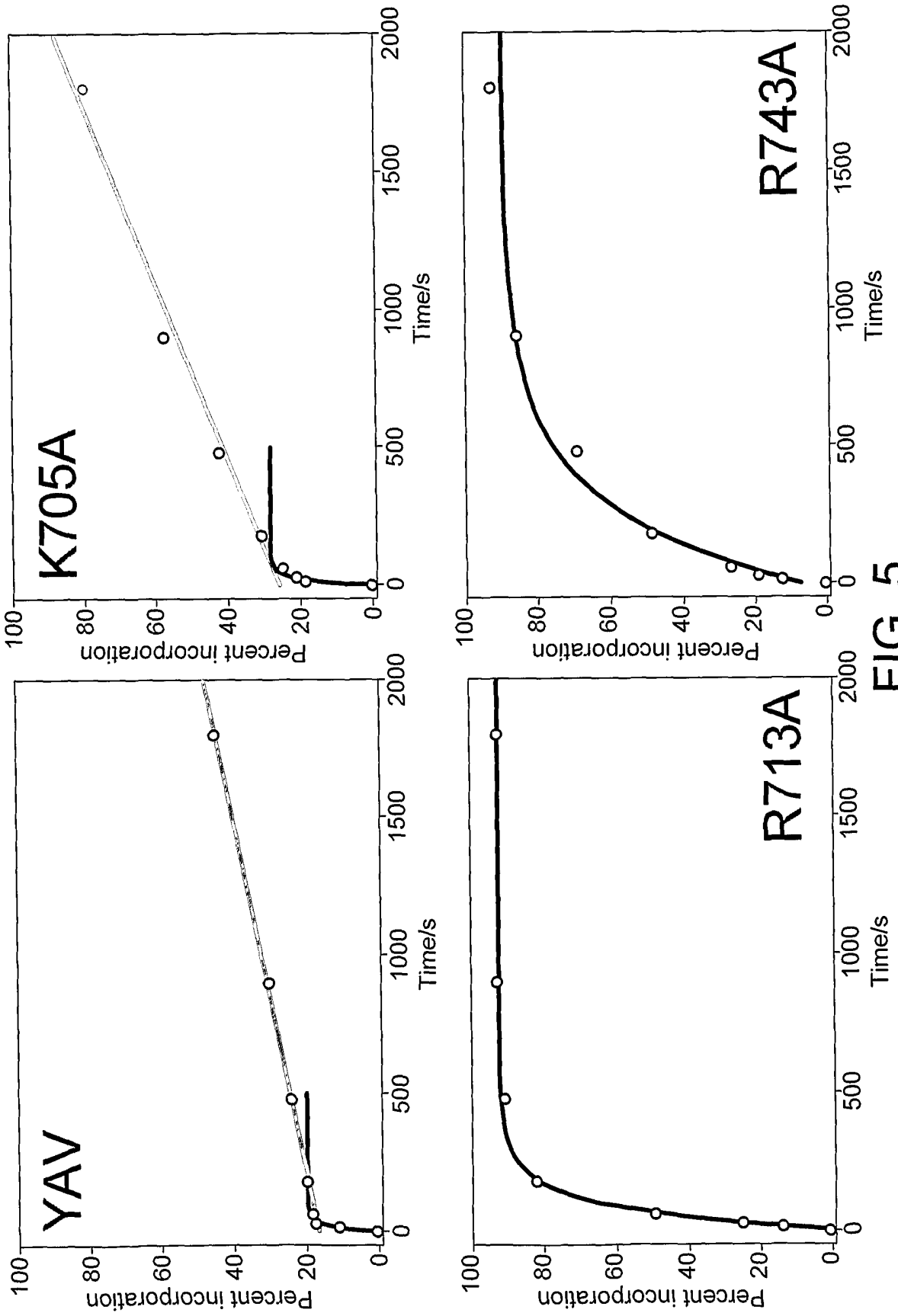

YAV
K705A
R713A
R743A
Percent incorporation
Time/s
0
20
40
60
80
100
500
1000
1500
2000

FIG. 5

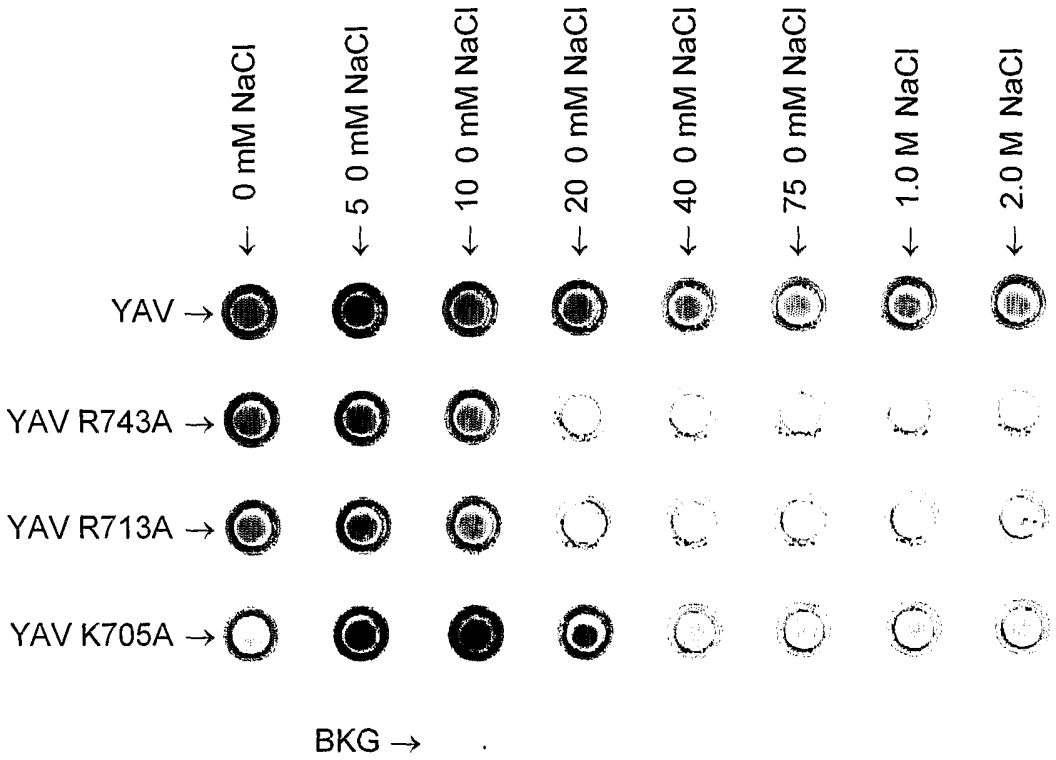

0 mM NaCl
5 0 mM NaCl
10 0 mM NaCl
20 0 mM NaCl
40 0 mM NaCl
75 0 mM NaCl
1.0 M NaCl
2.0 M NaCl
YAV →
YAV R743A →
YAV R713A →
YAV K705A →
BKG →

FIG. 6

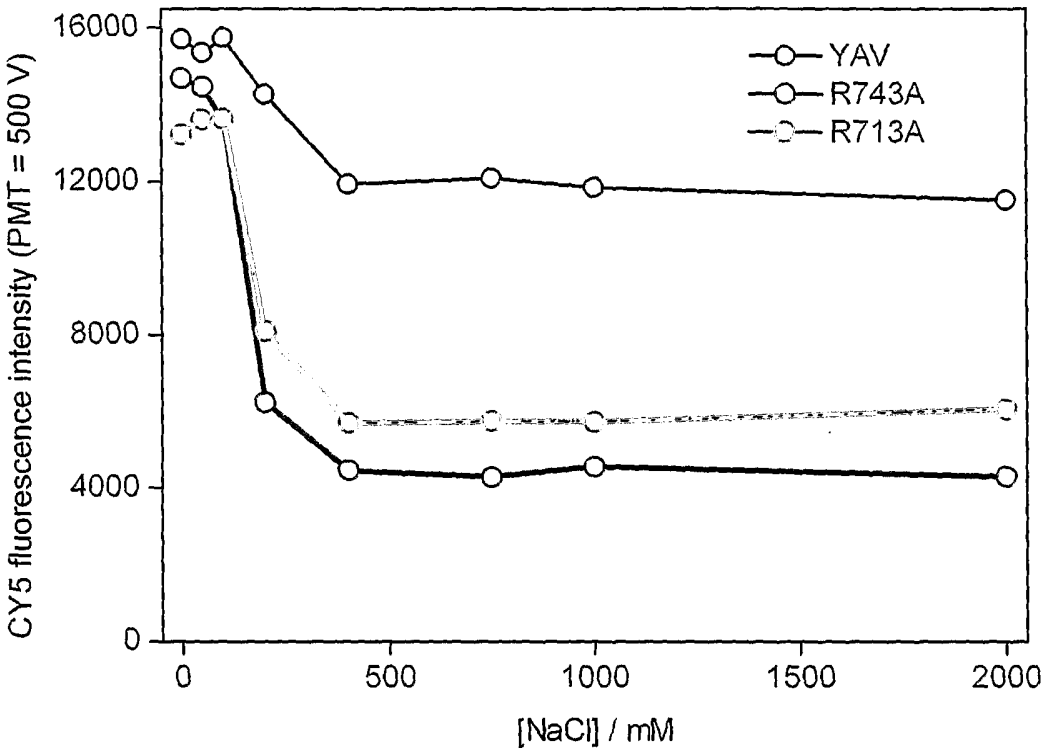

16000
12000
8000
4000
0
0
500
1000
1500
2000
CY5 fluorescence intensity (PMT = 500 V)
[NaCl] / mM
YAV
R743A
R713A

FIG. 7

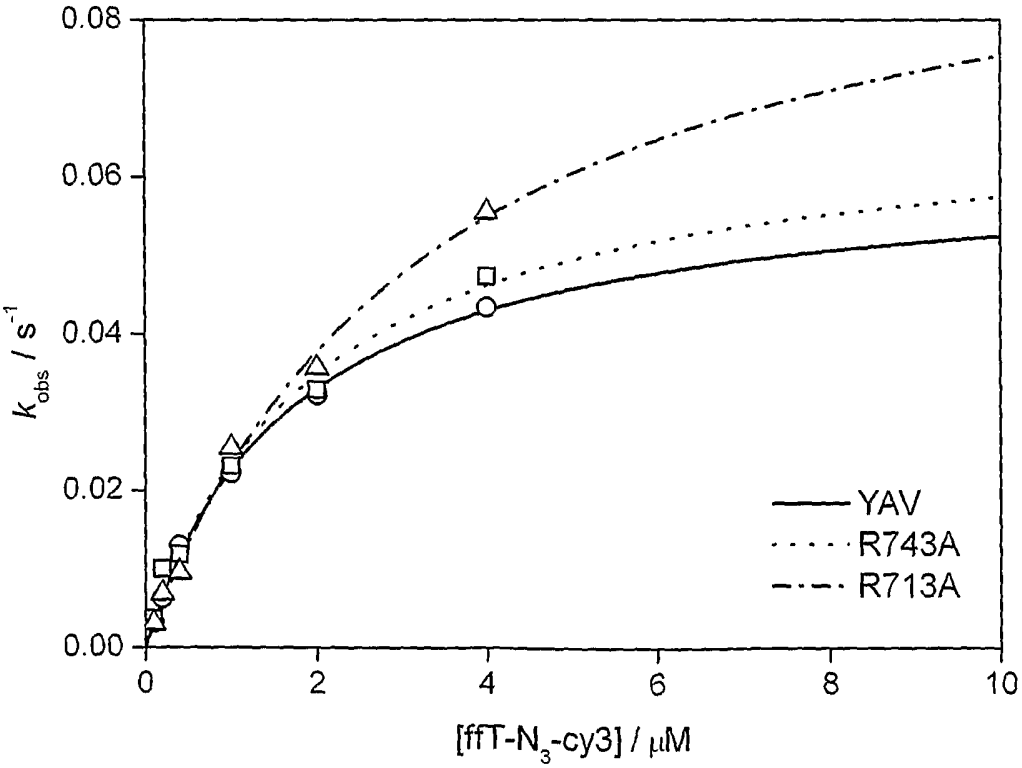

0.08
0.06
0.04
0.02
0.00
0
2
4
6
8
10
k_{obs} / s^{-1}
[ffT-N_3-cy3] / μM
YAV
R743A
R713A

FIG. 8

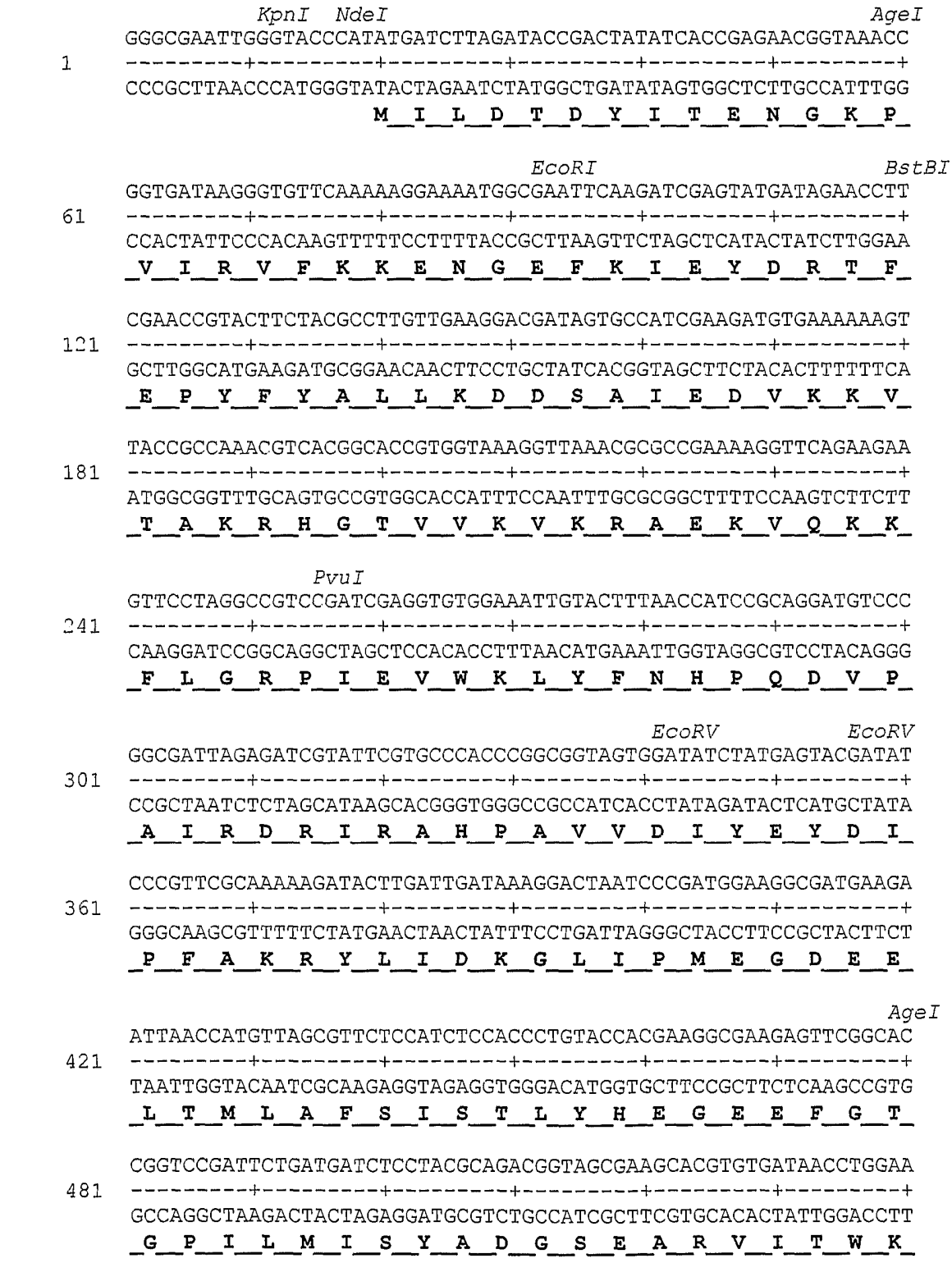

KpnI NdeI AgeI
GGGCGAATTGGGTACCCATATGATCTTAGATACCGACTATATCACCGAGAACGGTAAACC
1
CCCGCTTAACCCATGGGTATACTAGAATCTATGGCTGATATAGTGGCTCTTGCCATTTGG
M I L D T D Y I T E N G K P
EcoRI BstBI
GGTGATAAGGGTGTTCAAAAAGGAAAATGGCGAATTCAAGATCGAGTATGATAGAACCTT
61
CCACTATTCCCACAAGTTTTTCCTTTTACCGCTTAAGTTCTAGCTCATACTATCTTGGAA
V I R V F K K E N G E F K I E Y D R T F
CGAACCGTACTTCTACGCCTTGTTGAAGGACGATAGTGCCATCGAAGATGTGAAAAAAGT
121
GCTTGGCATGAAGATGCGGAACAACTTCCTGCTATCACGGTAGCTTCTACACTTTTTTCA
E P Y F Y A L L K D D S A I E D V K K V
TACCGCCAAACGTCACGGCACCGTGGTAAAGGTTAAACGCGCCGAAAAGGTTCAGAAGAA
181
ATGGCGGTTTGCAGTGCCGTGGCACCATTTCCAATTTGCGCGGCTTTTCCAAGTCTTCTT
T A K R H G T V V K V K R A E K V Q K K
PvuI
GTTCCTAGGCCGTCCGATCGAGGTGTGGAAATTGTACTTTAACCATCCGCAGGATGTCCC
241
CAAGGATCCGGCAGGCTAGCTCCACACCTTTAACATGAAATTGGTAGGCGTCCTACAGGG
F L G R P I E V W K L Y F N H P Q D V P
EcoRV EcoRV
GGCGATTAGAGATCGTATTCGTGCCCACCCGGCGGTAGTGGATATCTATGAGTACGATAT
301
CCGCTAATCTCTAGCATAAGCACGGGTGGGCCGCCATCACCTATAGATACTCATGCTATA
A I R D R I R A H P A V V D I Y E Y D I
CCCGTTCGCAAAAAGATACTTGATTGATAAAGGACTAATCCCGATGGAAGGCGATGAAGA
361
GGGCAAGCGTTTTTCTATGAACTAACTATTTCCTGATTAGGGCTACCTTCCGCTACTTCT
P F A K R Y L I D K G L I P M E G D E E
AgeI
ATTAACCATGTTAGCGTTCTCCATCTCCACCCTGTACCACGAAGGCGAAGAGTTCGGCAC
421
TAATTGGTACAATCGCAAGAGGTAGAGGTGGGACATGGTGCTTCCGCTTCTCAAGCCGTG
L T M L A F S I S T L Y H E G E E F G T
CGGTCCGATTCTGATGATCTCCTACGCAGACGGTAGCGAAGCACGTGTGATAACCTGGAA
481
GCCAGGCTAAGACTACTAGAGGATGCGTCTGCCATCGCTTCGTGCACACTATTGGACCTT
G P I L M I S Y A D G S E A R V I T W K

FIG. 9

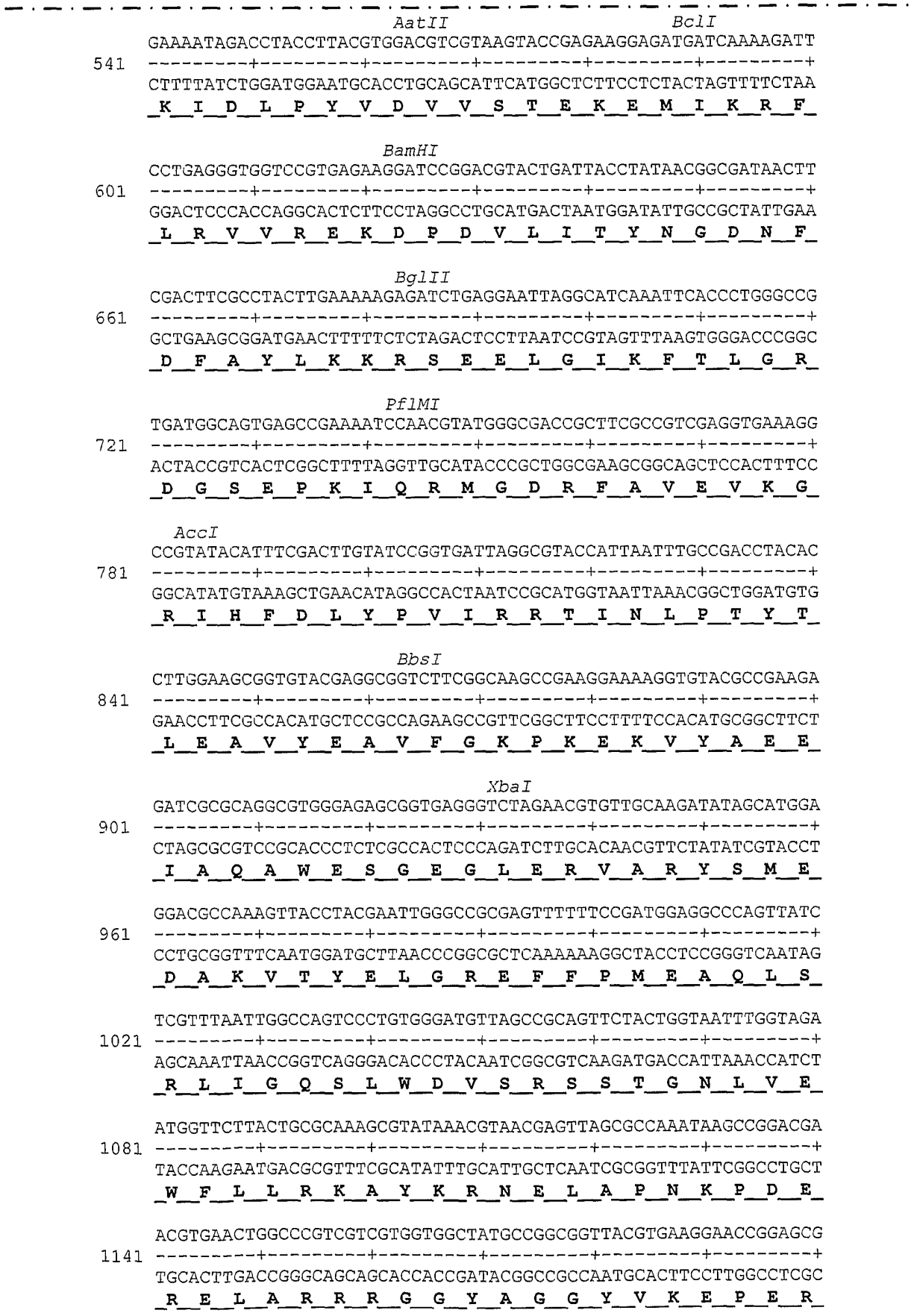

AatII
BclI
GAAAATAGACCTACCTTACGTGGACGTCGTAAGTACCGAGAAGGAGATGATCAAAAGATT
541
CTTTTATCTGGATGGAATGCACCTGCAGCATTCATGGCTCTTCCTCTACTAGTTTTCTAA
K I D L P Y V D V V S T E K E M I K R F
BamHI
CCTGAGGGTGGTCCGTGAGAAGGATCCGGACGTACTGATTACCTATAACGGCGATAACTT
601
GGACTCCCACCAGGCACTCTTCCTAGGCCTGCATGACTAATGGATATTGCCGCTATTGAA
L R V V R E K D P D V L I T Y N G D N F
BglII
CGACTTCGCCTACTTGAAAAAGAGATCTGAGGAATTAGGCATCAAATTCACCCTGGGCCG
661
GCTGAAGCGGATGAACTTTTTCTCTAGACTCCTTAATCCGTAGTTTAAGTGGGACCCGGC
D F A Y L K K R S E E L G I K F T L G R
PflMI
TGATGGCAGTGAGCCGAAAATCCAACGTATGGGCGACCGCTTCGCCGTCGAGGTGAAAGG
721
ACTACCGTCACTCGGCTTTTAGGTTGCATACCCGCTGGCGAAGCGGCAGCTCCACTTTCC
D G S E P K I Q R M G D R F A V E V K G
AccI
CCGTATACATTTCGACTTGTATCCGGTGATTAGGCGTACCATTAATTTGCCGACCTACAC
781
GGCATATGTAAAGCTGAACATAGGCCACTAATCCGCATGGTAATTAAACGGCTGGATGTG
R I H F D L Y P V I R R T I N L P T Y T
BbsI
CTTGGAAGCGGTGTACGAGGCGGTCTTCGGCAAGCCGAAGGAAAAGGTGTACGCCGAAGA
841
GAACCTTCGCCACATGCTCCGCCAGAAGCCGTTCGGCTTCCTTTTCCACATGCGGCTTCT
L E A V Y E A V F G K P K E K V Y A E E
XbaI
GATCGCGCAGGCGTGGGAGAGCGGTGAGGGTCTAGAACGTGTTGCAAGATATAGCATGGA
901
CTAGCGCGTCCGCACCCTCTCGCCACTCCCAGATCTTGCACAACGTTCTATATCGTACCT
I A Q A W E S G E G L E R V A R Y S M E
GGACGCCAAAGTTACCTACGAATTGGGCCGCGAGTTTTTTCCGATGGAGGCCCAGTTATC
961
CCTGCGGTTTCAATGGATGCTTAACCCGGCGCTCAAAAAAGGCTACCTCCGGGTCAATAG
D A K V T Y E L G R E F F P M E A Q L S
TCGTTTAATTGGCCAGTCCCTGTGGGATGTTAGCCGCAGTTCTACTGGTAATTTGGTAGA
1021
AGCAAATTAACCGGTCAGGGACACCCTACAATCGGCGTCAAGATGACCATTAAACCATCT
R L I G Q S L W D V S R S S T G N L V E
ATGGTTCTTACTGCGCAAAGCGTATAAACGTAACGAGTTAGCGCCAAATAAGCCGGACGA
1081
TACCAAGAATGACGCGTTTCGCATATTTGCATTGCTCAATCGCGGTTTATTCGGCCTGCT
W F L L R K A Y K R N E L A P N K P D E
ACGTGAACTGGCCCGTCGTCGTGGTGGCTATGCCGGCGGTTACGTGAAGGAACCGGAGCG
1141
TGCACTTGACCGGGCAGCAGCACCACCGATACGGCCGCCAATGCACTTCCTTGGCCTCGC
R E L A R R R G G Y A G G Y V K E P E R

FIG. 9 CONT'D

```
      TGGCCTATGGGATAACATTGTGTACCTTGACTTTAGAAGCTATGCGGTTAGCATCATCAT
1201  ---------+---------+---------+---------+---------+---------+
      ACCGGATACCCTATTGTAACACATGGAACTGAAATCTTCGATACGCCAATCGTAGTAGTA
      _G__L__W__D__N__I__V__Y__L__D__F__R__S__Y__A__V__S__I__I__I_

                                                          AatII
      CACCCATAATGTTAGTCCGGACACATTGAATCGTGAAGGATGCAAAGAATATGACGTCGC
1261  ---------+---------+---------+---------+---------+---------+
      GTGGGTATTACAATCAGGCCTGTGTAACTTAGCACTTCCTACGTTTCTTATACTGCAGCG
      _T__H__N__V__S__P__D__T__L__N__R__E__G__C__K__E__Y__D__V__A_

      CCCAGAGGTGGGCCACAAATTTTGTAAAGATTTCCCAGGATTCATCCCAAGTTTGTTGGG
1321  ---------+---------+---------+---------+---------+---------+
      GGGTCTCCACCCGGTGTTTAAAACATTTCTAAAGGGTCCTAAGTAGGGTTCAAACAACCC
      _P__E__V__G__H__K__F__C__K__D__F__P__G__F__I__P__S__L__L__G_

      TGATCTGCTGGAAGAACGCCAGAAAATCAAACGTAAGATGAAGGCGACCGTCGATCCACT
1381  ---------+---------+---------+---------+---------+---------+
      ACTAGACGACCTTCTTGCGGTCTTTTAGTTTGCATTCTACTTCCGCTGGCAGCTAGGTGA
      _D__L__L__E__E__R__Q__K__I__K__R__K__M__K__A__T__V__D__P__L_

                                          BclI          EcoRI
      GGAGAAAAAGCTATTGGACTACCGTCAGCGCCTGATCAAGATTTTGGCGAATTCTTTCTA
1441  ---------+---------+---------+---------+---------+---------+
      CCTCTTTTTCGATAACCTGATGGCAGTCGCGGACTAGTTCTAAAACCGCTTAAGAAAGAT
      _E__K__K__L__L__D__Y__R__Q__R__L__I__K__I__L__A__N__S__F__Y_

      TGGATACTACGGCTACGCCAAAGCCCGTTGGTATTGTAAAGAGTGCGCCGAGTCTGTCAC
1501  ---------+---------+---------+---------+---------+---------+
      ACCTATGATGCCGATGCGGTTTCGGGCAACCATAACATTTCTCACGCGGCTCAGACAGTG
      _G__Y__Y__G__Y__A__K__A__R__W__Y__C__K__E__C__A__E__S__V__T_

      TGCCTGGGGTCGTGAATATATCGAAATGGTGATCCGCGAGCTGGAAGAGAAATTTGGATT
1561  ---------+---------+---------+---------+---------+---------+
      ACGGACCCCAGCACTTATATAGCTTTACCACTAGGCGCTCGACCTTCTCTTTAAACCTAA
      _A__W__G__R__E__Y__I__E__M__V__I__R__E__L__E__E__K__F__G__F_

                                                              BsaI
      CAAAGTCTTGTACGCCGATACCGATGGTCTGCACGCGACCATTCCGGGTGCCGATGCCGA
1621  ---------+---------+---------+---------+---------+---------+
      GTTTCAGAACATGCGGCTATGGCTACCAGACGTGCGCTGGTAAGGCCCACGGCTACGGCT
      _K__V__L__Y__A__D__T__D__G__L__H__A__T__I__P__G__A__D__A__E_

      GACCGTGAAGAAAAAGGCGAAAGAGTTTTTGAAATATATCAATCCGAAGTTGCCGGGATT
1681  ---------+---------+---------+---------+---------+---------+
      CTGGCACTTCTTTTTCCGCTTTCTCAAAAACTTTATATAGTTAGGCTTCAACGGCCCTAA
      _T__V__K__K__K__A__K__E__F__L__K__Y__I__N__P__K__L__P__G__L_

      ATTAGAATTGGAATACGAAGGTTTCTATGTTCGCGGCTTTTTCGTGACCAAGAAAAAATA
1741  ---------+---------+---------+---------+---------+---------+
      TAATCTTAACCTTATGCTTCCAAAGATACAAGCGCCGAAAAAGCACTGGTTCTTTTTTAT
      _L__E__L__E__Y__E__G__F__Y__V__R__G__F__F__V__T__K__K__K__Y_

                                                  XbaI
      CGCCGTGATCGACGAGGAAGGAAAAATTACCACCCGTGGTCTAGAGATTGTTCGTCGTGA
1801  ---------+---------+---------+---------+---------+---------+
      GCGGCACTAGCTGCTCCTTCCTTTTTAATGGTGGGCACCAGATCTCTAACAAGCAGCACT
      _A__V__I__D__E__E__G__K__I__T__T__R__G__L__E__I__V__R__R__D_
```

FIG. 9 CONT'D

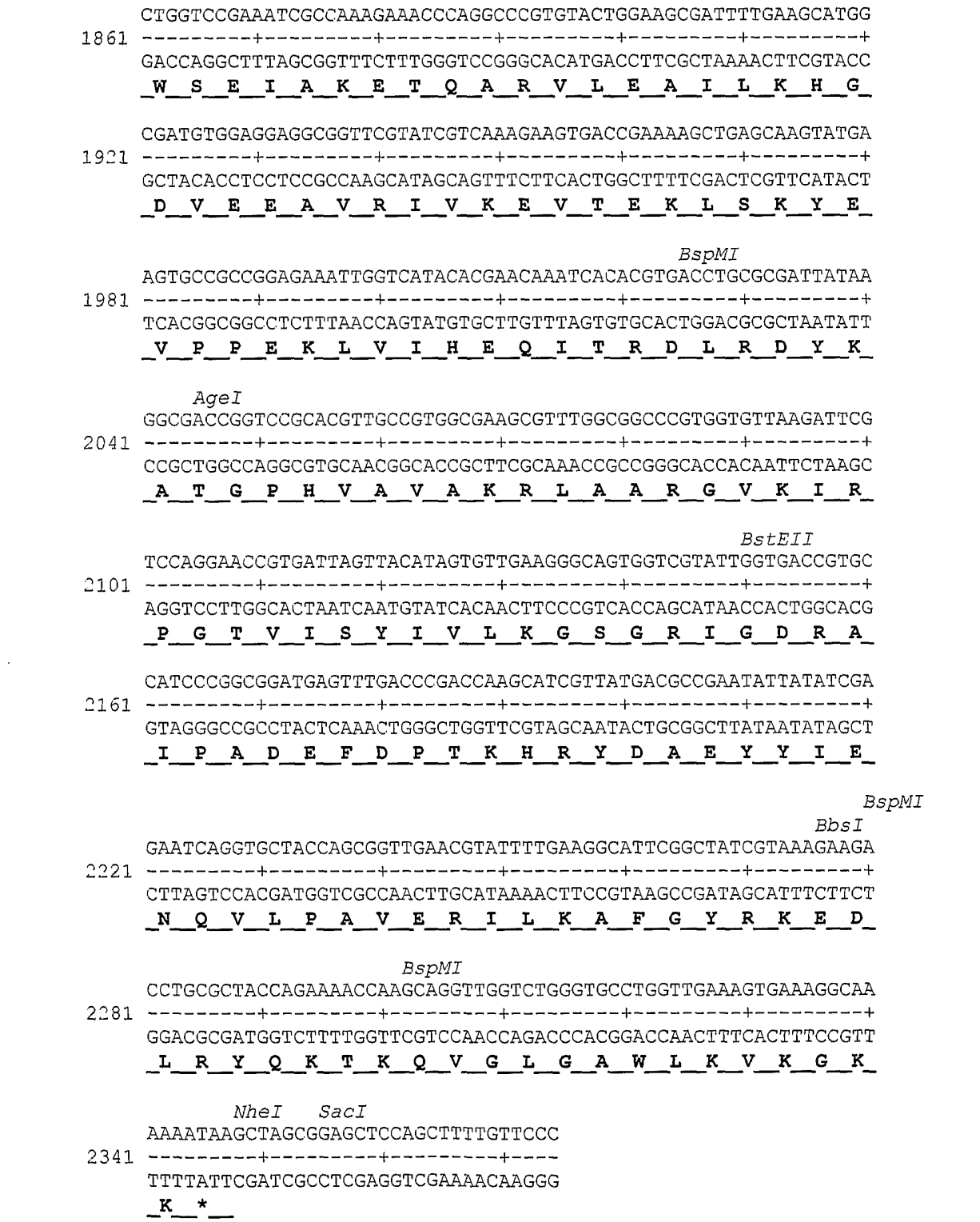
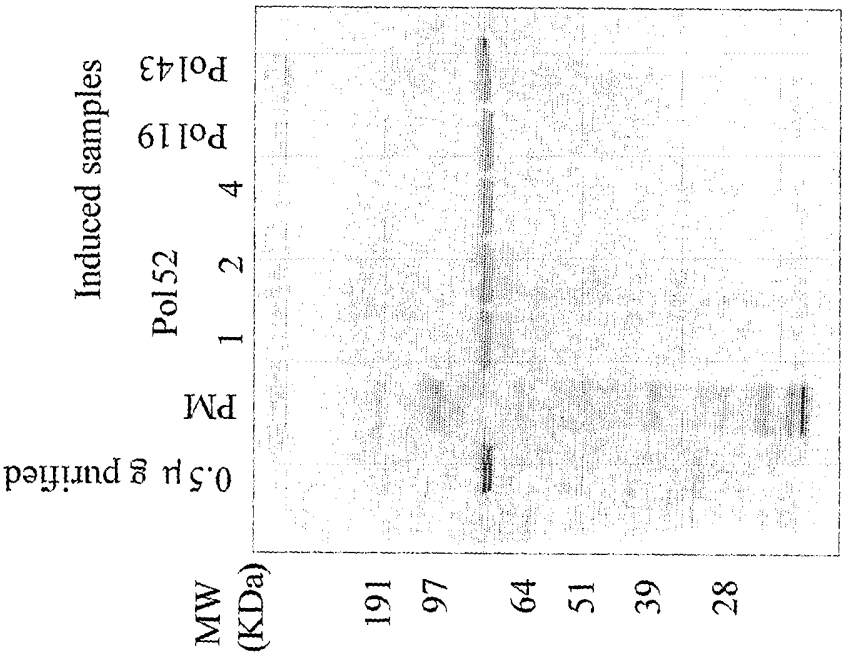

CTGGTCCGAAATCGCCAAAGAAACCCAGGCCCGTGTACTGGAAGCGATTTTGAAGCATGG
1861
GACCAGGCTTTAGCGGTTTCTTTGGGTCCGGGCACATGACCTTCGCTAAAACTTCGTACC
_W__S__E__I__A__K__E__T__Q__A__R__V__L__E__A__I__L__K__H__G_
CGATGTGGAGGAGGCGGTTCGTATCGTCAAAGAAGTGACCGAAAAGCTGAGCAAGTATGA
1921
GCTACACCTCCTCCGCCAAGCATAGCAGTTTCTTCACTGGCTTTTCGACTCGTTCATACT
_D__V__E__E__A__V__R__I__V__K__E__V__T__E__K__L__S__K__Y__E_
BspMI
AGTGCCGCCGGAGAAATTGGTCATACACGAACAAATCACACGTGACCTGCGCGATTATAA
1981
TCACGGCGGCCTCTTTAACCAGTATGTGCTTGTTTAGTGTGCACTGGACGCGCTAATATT
_V__P__P__E__K__L__V__I__H__E__Q__I__T__R__D__L__R__D__Y__K_
AgeI
GGCGACCGGTCCGCACGTTGCCGTGGCGAAGCGTTTGGCGGCCCGTGGTGTTAAGATTCG
2041
CCGCTGGCCAGGCGTGCAACGGCACCGCTTCGCAAACCGCCGGGCACCACAATTCTAAGC
_A__T__G__P__H__V__A__V__A__K__R__L__A__A__R__G__V__K__I__R_
BstEII
TCCAGGAACCGTGATTAGTTACATAGTGTTGAAGGGCAGTGGTCGTATTGGTGACCGTGC
2101
AGGTCCTTGGCACTAATCAATGTATCACAACTTCCCGTCACCAGCATAACCACTGGCACG
_P__G__T__V__I__S__Y__I__V__L__K__G__S__G__R__I__G__D__R__A_
CATCCCGGCGGATGAGTTTGACCCGACCAAGCATCGTTATGACGCCGAATATTATATCGA
2161
GTAGGGCCGCCTACTCAAACTGGGCTGGTTCGTAGCAATACTGCGGCTTATAATATAGCT
_I__P__A__D__E__F__D__P__T__K__H__R__Y__D__A__E__Y__Y__I__E_
BspMI
BbsI
GAATCAGGTGCTACCAGCGGTTGAACGTATTTTGAAGGCATTCGGCTATCGTAAAGAAGA
2221
CTTAGTCCACGATGGTCGCCAACTTGCATAAAACTTCCGTAAGCCGATAGCATTTCTTCT
_N__Q__V__L__P__A__V__E__R__I__L__K__A__F__G__Y__R__K__E__D_
BspMI
CCTGCGCTACCAGAAAACCAAGCAGGTTGGTCTGGGTGCCTGGTTGAAAGTGAAAGGCAA
2281
GGACGCGATGGTCTTTTGGTTCGTCCAACCAGACCCACGGACCAACTTTCACTTTCCGTT
_L__R__Y__Q__K__T__K__Q__V__G__L__G__A__W__L__K__V__K__G__K_
NheI SacI
AAAATAAGCTAGCGGAGCTCCAGCTTTTGTTCCC
2341
TTTTATTCGATCGCCTCGAGGTCGAAAACAAGGG
_K__*__

FIG. 9 CONT'D


Induced samples
Pol43
Pol19
Pol152
4
2
1
PM
0.5μg purified cl 9
MW
(KDa)
191
97
64
51
39
28
Gel II

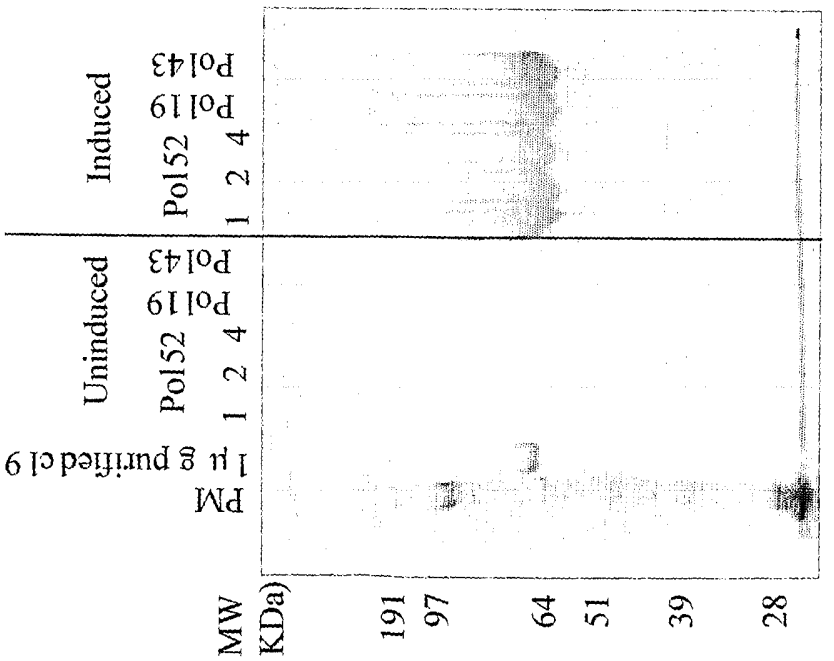

Induced
Pol43
Pol19
Pol152
4
2
1
Uninduced
Pol43
Pol19
Pol152
4
2
1
1μg purified cl 9
PM
MW
(KDa)
191
97
64
51
39
28
Gel I

FIG. 10

SEQUENCE LISTING

<110> Solexa Limited

<120> Improved polymerases

<130> P79219WO00

<160> 21

<170> PatentIn version 3.3

<210> 1
<211> 775
<212> PRT
<213> Artificial sequence

<220>
<223> Mutant polymerase

<400> 1

Met Ile Leu Asp Thr Asp Tyr Ile Thr Glu Asn Gly Lys Pro Val Ile
1 5 10 15

Arg Val Phe Lys Lys Glu Asn Gly Glu Phe Lys Ile Glu Tyr Asp Arg
20 25 30

Thr Phe Glu Pro Tyr Phe Tyr Ala Leu Leu Lys Asp Asp Ser Ala Ile
35 40 45

Glu Asp Val Lys Lys Val Thr Ala Lys Arg His Gly Thr Val Val Lys
50 55 60

Val Lys Arg Ala Glu Lys Val Gln Lys Lys Phe Leu Gly Arg Pro Ile
65 70 75 80

Glu Val Trp Lys Leu Tyr Phe Asn His Pro Gln Asp Val Pro Ala Ile
85 90 95

Arg Asp Arg Ile Arg Ala His Pro Ala Val Val Asp Ile Tyr Glu Tyr
100 105 110

Asp Ile Pro Phe Ala Lys Arg Tyr Leu Ile Asp Lys Gly Leu Ile Pro
115 120 125

Met Glu Gly Asp Glu Glu Leu Thr Met Leu Ala Phe Ala Ile Ala Thr
130 135 140

```
Leu Tyr His Glu Gly Glu Glu Phe Gly Thr Gly Pro Ile Leu Met Ile
145                 150                 155                 160

Ser Tyr Ala Asp Gly Ser Glu Ala Arg Val Ile Thr Trp Lys Lys Ile
                165                 170                 175

Asp Leu Pro Tyr Val Asp Val Val Ser Thr Glu Lys Glu Met Ile Lys
            180                 185                 190

Arg Phe Leu Arg Val Val Arg Glu Lys Asp Pro Asp Val Leu Ile Thr
        195                 200                 205

Tyr Asn Gly Asp Asn Phe Asp Phe Ala Tyr Leu Lys Lys Arg Ser Glu
    210                 215                 220

Glu Leu Gly Ile Lys Phe Thr Leu Gly Arg Asp Gly Ser Glu Pro Lys
225                 230                 235                 240

Ile Gln Arg Met Gly Asp Arg Phe Ala Val Glu Val Lys Gly Arg Ile
                245                 250                 255

His Phe Asp Leu Tyr Pro Val Ile Arg Arg Thr Ile Asn Leu Pro Thr
            260                 265                 270

Tyr Thr Leu Glu Ala Val Tyr Glu Ala Val Phe Gly Lys Pro Lys Glu
        275                 280                 285

Lys Val Tyr Ala Glu Glu Ile Ala Gln Ala Trp Glu Ser Gly Glu Gly
    290                 295                 300

Leu Glu Arg Val Ala Arg Tyr Ser Met Glu Asp Ala Lys Val Thr Tyr
305                 310                 315                 320

Glu Leu Gly Arg Glu Phe Phe Pro Met Glu Ala Gln Leu Ser Arg Leu
                325                 330                 335

Ile Gly Gln Ser Leu Trp Asp Val Ser Arg Ser Ser Thr Gly Asn Leu
            340                 345                 350

Val Glu Trp Phe Leu Leu Arg Lys Ala Tyr Lys Arg Asn Glu Leu Ala
        355                 360                 365

Pro Asn Lys Pro Asp Glu Arg Glu Leu Ala Arg Arg Arg Gly Gly Tyr
    370                 375                 380
```

```
Ala Gly Gly Tyr Val Lys Glu Pro Glu Arg Gly Leu Trp Asp Asn Ile
385                 390                 395                 400

Val Tyr Leu Asp Phe Arg Ser Tyr Ala Val Ser Ile Ile Ile Thr His
                405                 410                 415

Asn Val Ser Pro Asp Thr Leu Asn Arg Glu Gly Cys Lys Glu Tyr Asp
            420                 425                 430

Val Ala Pro Glu Val Gly His Lys Phe Cys Lys Asp Phe Pro Gly Phe
        435                 440                 445

Ile Pro Ser Leu Leu Gly Asp Leu Leu Glu Glu Arg Gln Lys Ile Lys
    450                 455                 460

Arg Lys Met Lys Ala Thr Val Asp Pro Leu Glu Lys Lys Leu Leu Asp
465                 470                 475                 480

Tyr Arg Gln Arg Leu Ile Lys Ile Leu Ala Asn Ser Phe Tyr Gly Tyr
                485                 490                 495

Tyr Gly Tyr Ala Lys Ala Arg Trp Tyr Cys Lys Glu Cys Ala Glu Ser
            500                 505                 510

Val Thr Ala Trp Gly Arg Glu Tyr Ile Glu Met Val Ile Arg Glu Leu
        515                 520                 525

Glu Glu Lys Phe Gly Phe Lys Val Leu Tyr Ala Asp Thr Asp Gly Leu
    530                 535                 540

His Ala Thr Ile Pro Gly Ala Asp Ala Glu Thr Val Lys Lys Lys Ala
545                 550                 555                 560

Lys Glu Phe Leu Lys Tyr Ile Asn Pro Lys Leu Pro Gly Leu Leu Glu
                565                 570                 575

Leu Glu Tyr Glu Gly Phe Tyr Val Arg Gly Phe Phe Val Thr Lys Lys
            580                 585                 590

Lys Tyr Ala Val Ile Asp Glu Glu Gly Lys Ile Thr Thr Arg Gly Leu
        595                 600                 605
```

Glu Ile Val Arg Arg Asp Trp Ser Glu Ile Ala Lys Glu Thr Gln Ala
610 615 620

Arg Val Leu Glu Ala Ile Leu Lys His Gly Asp Val Glu Glu Ala Val
625 630 635 640

Arg Ile Val Lys Glu Val Thr Glu Lys Leu Ser Lys Tyr Glu Val Pro
645 650 655

Pro Glu Lys Leu Val Ile His Glu Gln Ile Thr Arg Asp Leu Arg Asp
660 665 670

Tyr Lys Ala Thr Gly Pro His Val Ala Val Ala Lys Arg Leu Ala Ala
675 680 685

Arg Gly Val Lys Ile Arg Pro Gly Thr Val Ile Ser Tyr Ile Val Leu
690 695 700

Ala Gly Ser Gly Arg Ile Gly Asp Arg Ala Ile Pro Ala Asp Glu Phe
705 710 715 720

Asp Pro Thr Lys His Arg Tyr Asp Ala Glu Tyr Tyr Ile Glu Asn Gln
725 730 735

Val Leu Pro Ala Val Glu Arg Ile Leu Lys Ala Phe Gly Tyr Arg Lys
740 745 750

Glu Asp Leu Arg Tyr Gln Lys Thr Lys Gln Val Gly Leu Gly Ala Trp
755 760 765

Leu Lys Val Lys Gly Lys Lys
770 775

<210> 2
<211> 2328
<212> DNA
<213> Artificial sequence

<220>
<223> Mutant polymerase

<400> 2
atgattctcg ataccgacta catcaccgag aacgggaagc ccgtgataag ggtcttcaag 60

aaggagaacg gcgagtttaa aatcgagtac gacagaacct tcgagcccta cttctacgcc 120

cttctgaagg acgattctgc gatagaggac gtcaagaagg taaccgcaaa gaggcacgga 180

acggttgtca aggtgaagcg cgccgagaag gtgcagaaga agttcctcgg caggccgata 240

gaggtctgga agctctactt caaccatcct caggacgtcc cggcgattcg agacaggata 300

cgtgcccacc ccgctgtcgt tgacatctac gagtacgaca tacccttcgc caagcgctac 360

ctcatcgaca agggcctgat tccgatggag ggcgacgagg agcttacgat gctcgccttc 420

gcgatcgcaa ccctctatca cgagggcgag gagttcggaa ccgggccgat tctcatgata 480

agctacgccg acgggagcga ggcgagggtg ataacctgga agaagattga ccttccgtac 540

gttgacgtcg tctcgaccga gaaggagatg attaagcgct tcctccgcgt cgtcagggag 600

aaggaccccg acgtgctcat cacctacaac ggcgacaact tcgacttcgc ctacctgaag 660

aagcgctctg aggaactcgg aataaagttc acactcggca gggacgggag cgagccgaag 720

atacagcgaa tgggcgaccg ctttgccgtt gaggtgaagg gcaggattca cttcgacctc 780

taccccgtca taaggcgcac gataaacctc ccgacctaca cccttgaggc cgtttacgag 840

gccgtctttg gaaagcccaa ggagaaggtt tacgcagagg agatagcgca ggcctgggag 900

agcggggagg gccttgaaag ggttgcaaga tactcgatgg aggacgctaa ggtgacctac 960

gagctgggaa gggagttctt cccgatggag gcccagcttt cgaggcttat aggccagagc 1020

ctctgggacg tctcgcgctc gagcaccgga aatttggtgg agtggttcct cctgcggaag 1080

gcctacaaga ggaacgagct cgccccaaac aagcccgacg agagggagct cgcgagacgg 1140

cgcgggggct acgctggcgg gtacgttaag gaaccagagc ggggattgtg ggacaacatt 1200

gtgtatctag acttccgctc gtatgcggtt tcaatcatca taacccacaa cgtctcgccg 1260

gataccctca accgcgaggg ctgtaaagag tacgacgtcg cccctgaggt tggacacaag 1320

ttctgcaagg acttccccgg cttcatacca agcctcctgg gagatttgct cgaggagagg 1380

cagaagataa agcggaagat gaaggcaacg gttgacccgc tggagaagaa actcctcgat 1440

tacaggcaga ggctgatcaa aatcctcgcc aacagcttct acggctacta cggctacgcc 1500

aaggcccggt ggtactgcaa ggagtgcgcc gagagcgtta cggcctgggg aagggagtat 1560

atagaaatgg ttatccggga actcgaagaa aaattcggtt ttaaagttct ctatgccgat 1620

acagacggtc tccatgctac cattcccgga gcagacgctg aaacagtcaa gaaaaaagca 1680

aaggagttct taaaatacat taatccaaaa ctgcccggcc tgctcgaact tgagtacgag 1740

ggcttctacg tgaggggctt cttcgtcacg aagaagaagt acgctgtgat agacgaggag 1800

ggcaagataa ccacgagggg tcttgagatt gtgaggcgcg actggagcga gatagcgaag 1860

gagacccagg ccagggtctt agaggcgata ctcaagcacg gtgacgtcga ggaggccgtt 1920

aggatagtca aggaagtgac ggaaaagctg agcaagtatg aggtcccgcc cgagaagctg 1980

gtaatccacg agcagataac gcgcgatttg agggattaca aagccaccgg cccgcacgtt 2040

gccgttgcga agaggctcgc ggcgcgtgga gtgaaaatcc ggcccggcac ggtgataagc 2100

tacatcgtcc tagcgggctc tggaaggata ggcgacaggg cgattccagc tgatgagttc 2160

gacccgacga agcaccgcta cgatgcggaa tactacatcg agaaccaggt tctcccggcg 2220

gtggagagga ttctaaaagc cttcggctat cggaaggagg atttgcgcta ccagaagacg 2280

aagcaggtcg gcttgggcgc gtggctgaag gtgaagggga agaagtga 2328

<210> 3
<211> 775
<212> PRT
<213> Artificial sequence

<220>
<223> Mutant polymerase

<400> 3

Met Ile Leu Asp Thr Asp Tyr Ile Thr Glu Asn Gly Lys Pro Val Ile
1 5 10 15

Arg Val Phe Lys Lys Glu Asn Gly Glu Phe Lys Ile Glu Tyr Asp Arg
20 25 30

Thr Phe Glu Pro Tyr Phe Tyr Ala Leu Leu Lys Asp Asp Ser Ala Ile
35 40 45

Glu Asp Val Lys Lys Val Thr Ala Lys Arg His Gly Thr Val Val Lys
50 55 60

Val Lys Arg Ala Glu Lys Val Gln Lys Lys Phe Leu Gly Arg Pro Ile
65 70 75 80

Glu Val Trp Lys Leu Tyr Phe Asn His Pro Gln Asp Val Pro Ala Ile
85 90 95

Arg Asp Arg Ile Arg Ala His Pro Ala Val Val Asp Ile Tyr Glu Tyr
100 105 110

Asp Ile Pro Phe Ala Lys Arg Tyr Leu Ile Asp Lys Gly Leu Ile Pro
115 120 125

```
Met Glu Gly Asp Glu Glu Leu Thr Met Leu Ala Phe Ala Ile Ala Thr
    130                 135                 140

Leu Tyr His Glu Gly Glu Glu Phe Gly Thr Gly Pro Ile Leu Met Ile
145                 150                 155                 160

Ser Tyr Ala Asp Gly Ser Glu Ala Arg Val Ile Thr Trp Lys Lys Ile
                165                 170                 175

Asp Leu Pro Tyr Val Asp Val Val Ser Thr Glu Lys Glu Met Ile Lys
            180                 185                 190

Arg Phe Leu Arg Val Val Arg Glu Lys Asp Pro Asp Val Leu Ile Thr
        195                 200                 205

Tyr Asn Gly Asp Asn Phe Asp Phe Ala Tyr Leu Lys Lys Arg Ser Glu
    210                 215                 220

Glu Leu Gly Ile Lys Phe Thr Leu Gly Arg Asp Gly Ser Glu Pro Lys
225                 230                 235                 240

Ile Gln Arg Met Gly Asp Arg Phe Ala Val Glu Val Lys Gly Arg Ile
                245                 250                 255

His Phe Asp Leu Tyr Pro Val Ile Arg Arg Thr Ile Asn Leu Pro Thr
            260                 265                 270

Tyr Thr Leu Glu Ala Val Tyr Glu Ala Val Phe Gly Lys Pro Lys Glu
        275                 280                 285

Lys Val Tyr Ala Glu Glu Ile Ala Gln Ala Trp Glu Ser Gly Glu Gly
    290                 295                 300

Leu Glu Arg Val Ala Arg Tyr Ser Met Glu Asp Ala Lys Val Thr Tyr
305                 310                 315                 320

Glu Leu Gly Arg Glu Phe Phe Pro Met Glu Ala Gln Leu Ser Arg Leu
                325                 330                 335

Ile Gly Gln Ser Leu Trp Asp Val Ser Arg Ser Ser Thr Gly Asn Leu
            340                 345                 350

Val Glu Trp Phe Leu Leu Arg Lys Ala Tyr Lys Arg Asn Glu Leu Ala
```

355 360 365

Pro Asn Lys Pro Asp Glu Arg Glu Leu Ala Arg Arg Arg Gly Gly Tyr
370 375 380

Ala Gly Gly Tyr Val Lys Glu Pro Glu Arg Gly Leu Trp Asp Asn Ile
385 390 395 400

Val Tyr Leu Asp Phe Arg Ser Tyr Ala Val Ser Ile Ile Ile Thr His
405 410 415

Asn Val Ser Pro Asp Thr Leu Asn Arg Glu Gly Cys Lys Glu Tyr Asp
420 425 430

Val Ala Pro Glu Val Gly His Lys Phe Cys Lys Asp Phe Pro Gly Phe
435 440 445

Ile Pro Ser Leu Leu Gly Asp Leu Leu Glu Glu Arg Gln Lys Ile Lys
450 455 460

Arg Lys Met Lys Ala Thr Val Asp Pro Leu Glu Lys Lys Leu Leu Asp
465 470 475 480

Tyr Arg Gln Arg Leu Ile Lys Ile Leu Ala Asn Ser Phe Tyr Gly Tyr
485 490 495

Tyr Gly Tyr Ala Lys Ala Arg Trp Tyr Cys Lys Glu Cys Ala Glu Ser
500 505 510

Val Thr Ala Trp Gly Arg Glu Tyr Ile Glu Met Val Ile Arg Glu Leu
515 520 525

Glu Glu Lys Phe Gly Phe Lys Val Leu Tyr Ala Asp Thr Asp Gly Leu
530 535 540

His Ala Thr Ile Pro Gly Ala Asp Ala Glu Thr Val Lys Lys Lys Ala
545 550 555 560

Lys Glu Phe Leu Lys Tyr Ile Asn Pro Lys Leu Pro Gly Leu Leu Glu
565 570 575

Leu Glu Tyr Glu Gly Phe Tyr Val Arg Gly Phe Phe Val Thr Lys Lys
580 585 590

Lys Tyr Ala Val Ile Asp Glu Glu Gly Lys Ile Thr Thr Arg Gly Leu
595 600 605

Glu Ile Val Arg Arg Asp Trp Ser Glu Ile Ala Lys Glu Thr Gln Ala
610 615 620

Arg Val Leu Glu Ala Ile Leu Lys His Gly Asp Val Glu Glu Ala Val
625 630 635 640

Arg Ile Val Lys Glu Val Thr Glu Lys Leu Ser Lys Tyr Glu Val Pro
645 650 655

Pro Glu Lys Leu Val Ile His Glu Gln Ile Thr Arg Asp Leu Arg Asp
660 665 670

Tyr Lys Ala Thr Gly Pro His Val Ala Val Ala Lys Arg Leu Ala Ala
675 680 685

Arg Gly Val Lys Ile Arg Pro Gly Thr Val Ile Ser Tyr Ile Val Leu
690 695 700

Lys Gly Ser Gly Arg Ile Gly Asp Ala Ala Ile Pro Ala Asp Glu Phe
705 710 715 720

Asp Pro Thr Lys His Arg Tyr Asp Ala Glu Tyr Tyr Ile Glu Asn Gln
725 730 735

Val Leu Pro Ala Val Glu Arg Ile Leu Lys Ala Phe Gly Tyr Arg Lys
740 745 750

Glu Asp Leu Arg Tyr Gln Lys Thr Lys Gln Val Gly Leu Gly Ala Trp
755 760 765

Leu Lys Val Lys Gly Lys Lys
770 775

<210> 4
<211> 2328
<212> DNA
<213> Artificial sequence

<220>
<223> Mutant polymerase

<400> 4
atgattctcg ataccgacta catcaccgag aacgggaagc ccgtgataag ggtcttcaag 60
aaggagaacg gcgagtttaa aatcgagtac gacagaacct tcgagcccta cttctacgcc 120
cttctgaagg acgattctgc gatagaggac gtcaagaagg taaccgcaaa gaggcacgga 180
acggttgtca aggtgaagcg cgccgagaag gtgcagaaga agttcctcgg caggccgata 240
gaggtctgga agctctactt caaccatcct caggacgtcc cggcgattcg agacaggata 300
cgtgcccacc ccgctgtcgt tgacatctac gagtacgaca tacccttcgc caagcgctac 360
ctcatcgaca agggcctgat tccgatggag ggcgacgagg agcttacgat gctcgccttc 420
gcgatcgcaa ccctctatca cgagggcgag gagttcggaa ccgggccgat tctcatgata 480
agctacgccg acgggagcga ggcgagggtg ataacctgga agaagattga ccttccgtac 540
gttgacgtcg tctcgaccga gaaggagatg attaagcgct tcctccgcgt cgtcagggag 600
aaggaccccg acgtgctcat cacctacaac ggcgacaact tcgacttcgc ctacctgaag 660
aagcgctctg aggaactcgg aataaagttc acactcggca gggacgggag cgagccgaag 720
atacagcgaa tgggcgaccg ctttgccgtt gaggtgaagg gcaggattca cttcgacctc 780
taccccgtca taaggcgcac gataaacctc ccgacctaca cccttgaggc cgtttacgag 840
gccgtctttg gaaagcccaa ggagaaggtt tacgcagagg agatagcgca ggcctgggag 900
agcggggagg gccttgaaag ggttgcaaga tactcgatgg aggacgctaa ggtgacctac 960
gagctgggaa gggagttctt cccgatggag gcccagcttt cgaggcttat aggccagagc 1020
ctctgggacg tctcgcgctc gagcaccgga aatttggtgg agtggttcct cctgcggaag 1080
gcctacaaga ggaacgagct cgccccaaac aagcccgacg agagggagct cgcgagacgg 1140
cgcgggggct acgctggcgg gtacgttaag gaaccagagc ggggattgtg ggacaacatt 1200
gtgtatctag acttccgctc gtatgcggtt tcaatcatca taacccacaa cgtctcgccg 1260
gataccctca accgcgaggg ctgtaaagag tacgacgtcg cccctgaggt tggacacaag 1320
ttctgcaagg acttccccgg cttcatacca agcctcctgg gagatttgct cgaggagagg 1380
cagaagataa agcggaagat gaaggcaacg gttgacccgc tggagaagaa actcctcgat 1440
tacaggcaga ggctgatcaa aatcctcgcc aacagcttct acggctacta cggctacgcc 1500
aaggcccggt ggtactgcaa ggagtgcgcc gagagcgtta cggcctgggg aagggagtat 1560
atagaaatgg ttatccggga actcgaagaa aaattcggtt ttaaagttct ctatgccgat 1620
acagacggtc tccatgctac cattcccgga gcagacgctg aaacagtcaa gaaaaaagca 1680

aaggagttct taaaatacat taatccaaaa ctgcccggcc tgctcgaact tgagtacgag 1740

ggcttctacg tgaggggctt cttcgtcacg aagaagaagt acgctgtgat agacgaggag 1800

ggcaagataa ccacgagggg tcttgagatt gtgaggcgcg actggagcga gatagcgaag 1860

gagacccagg ccagggtctt agaggcgata ctcaagcacg gtgacgtcga ggaggccgtt 1920

aggatagtca aggaagtgac ggaaaagctg agcaagtatg aggtcccgcc cgagaagctg 1980

gtaatccacg agcagataac gcgcgatttg agggattaca aagccaccgg cccgcacgtt 2040

gccgttgcga agaggctcgc ggcgcgtgga gtgaaaatcc ggcccggcac ggtgataagc 2100

tacatcgtcc taaagggctc tggaaggata ggcgacgcgg cgattccagc tgatgagttc 2160

gacccgacga agcaccgcta cgatgcggaa tactacatcg agaaccaggt tctcccggcg 2220

gtggagagga ttctaaaagc cttcggctat cggaaggagg atttgcgcta ccagaagacg 2280

aagcaggtcg gcttgggcgc gtggctgaag gtgaagggga agaagtga 2328

<210> 5
<211> 775
<212> PRT
<213> Artificial sequence

<220>
<223> Mutant polymerase

<400> 5

Met Ile Leu Asp Thr Asp Tyr Ile Thr Glu Asn Gly Lys Pro Val Ile
1 5 10 15

Arg Val Phe Lys Lys Glu Asn Gly Glu Phe Lys Ile Glu Tyr Asp Arg
20 25 30

Thr Phe Glu Pro Tyr Phe Tyr Ala Leu Leu Lys Asp Asp Ser Ala Ile
35 40 45

Glu Asp Val Lys Lys Val Thr Ala Lys Arg His Gly Thr Val Val Lys
50 55 60

Val Lys Arg Ala Glu Lys Val Gln Lys Lys Phe Leu Gly Arg Pro Ile
65 70 75 80

Glu Val Trp Lys Leu Tyr Phe Asn His Pro Gln Asp Val Pro Ala Ile
85 90 95

Arg Asp Arg Ile Arg Ala His Pro Ala Val Val Asp Ile Tyr Glu Tyr

100 105 110

Asp Ile Pro Phe Ala Lys Arg Tyr Leu Ile Asp Lys Gly Leu Ile Pro
115 120 125

Met Glu Gly Asp Glu Glu Leu Thr Met Leu Ala Phe Ala Ile Ala Thr
130 135 140

Leu Tyr His Glu Gly Glu Glu Phe Gly Thr Gly Pro Ile Leu Met Ile
145 150 155 160

Ser Tyr Ala Asp Gly Ser Glu Ala Arg Val Ile Thr Trp Lys Lys Ile
165 170 175

Asp Leu Pro Tyr Val Asp Val Val Ser Thr Glu Lys Glu Met Ile Lys
180 185 190

Arg Phe Leu Arg Val Val Arg Glu Lys Asp Pro Asp Val Leu Ile Thr
195 200 205

Tyr Asn Gly Asp Asn Phe Asp Phe Ala Tyr Leu Lys Lys Arg Ser Glu
210 215 220

Glu Leu Gly Ile Lys Phe Thr Leu Gly Arg Asp Gly Ser Glu Pro Lys
225 230 235 240

Ile Gln Arg Met Gly Asp Arg Phe Ala Val Glu Val Lys Gly Arg Ile
245 250 255

His Phe Asp Leu Tyr Pro Val Ile Arg Arg Thr Ile Asn Leu Pro Thr
260 265 270

Tyr Thr Leu Glu Ala Val Tyr Glu Ala Val Phe Gly Lys Pro Lys Glu
275 280 285

Lys Val Tyr Ala Glu Glu Ile Ala Gln Ala Trp Glu Ser Gly Glu Gly
290 295 300

Leu Glu Arg Val Ala Arg Tyr Ser Met Glu Asp Ala Lys Val Thr Tyr
305 310 315 320

Glu Leu Gly Arg Glu Phe Phe Pro Met Glu Ala Gln Leu Ser Arg Leu
325 330 335

Ile Gly Gln Ser Leu Trp Asp Val Ser Arg Ser Ser Thr Gly Asn Leu
340 345 350

Val Glu Trp Phe Leu Leu Arg Lys Ala Tyr Lys Arg Asn Glu Leu Ala
355 360 365

Pro Asn Lys Pro Asp Glu Arg Glu Leu Ala Arg Arg Arg Gly Gly Tyr
370 375 380

Ala Gly Gly Tyr Val Lys Glu Pro Glu Arg Gly Leu Trp Asp Asn Ile
385 390 395 400

Val Tyr Leu Asp Phe Arg Ser Tyr Ala Val Ser Ile Ile Ile Thr His
405 410 415

Asn Val Ser Pro Asp Thr Leu Asn Arg Glu Gly Cys Lys Glu Tyr Asp
420 425 430

Val Ala Pro Glu Val Gly His Lys Phe Cys Lys Asp Phe Pro Gly Phe
435 440 445

Ile Pro Ser Leu Leu Gly Asp Leu Leu Glu Glu Arg Gln Lys Ile Lys
450 455 460

Arg Lys Met Lys Ala Thr Val Asp Pro Leu Glu Lys Lys Leu Leu Asp
465 470 475 480

Tyr Arg Gln Arg Leu Ile Lys Ile Leu Ala Asn Ser Phe Tyr Gly Tyr
485 490 495

Tyr Gly Tyr Ala Lys Ala Arg Trp Tyr Cys Lys Glu Cys Ala Glu Ser
500 505 510

Val Thr Ala Trp Gly Arg Glu Tyr Ile Glu Met Val Ile Arg Glu Leu
515 520 525

Glu Glu Lys Phe Gly Phe Lys Val Leu Tyr Ala Asp Thr Asp Gly Leu
530 535 540

His Ala Thr Ile Pro Gly Ala Asp Ala Glu Thr Val Lys Lys Lys Ala
545 550 555 560

Lys Glu Phe Leu Lys Tyr Ile Asn Pro Lys Leu Pro Gly Leu Leu Glu

565 570 575

Leu Glu Tyr Glu Gly Phe Tyr Val Arg Gly Phe Phe Val Thr Lys Lys
580 585 590

Lys Tyr Ala Val Ile Asp Glu Glu Gly Lys Ile Thr Thr Arg Gly Leu
595 600 605

Glu Ile Val Arg Arg Asp Trp Ser Glu Ile Ala Lys Glu Thr Gln Ala
610 615 620

Arg Val Leu Glu Ala Ile Leu Lys His Gly Asp Val Glu Glu Ala Val
625 630 635 640

Arg Ile Val Lys Glu Val Thr Glu Lys Leu Ser Lys Tyr Glu Val Pro
645 650 655

Pro Glu Lys Leu Val Ile His Glu Gln Ile Thr Arg Asp Leu Arg Asp
660 665 670

Tyr Lys Ala Thr Gly Pro His Val Ala Val Ala Lys Arg Leu Ala Ala
675 680 685

Arg Gly Val Lys Ile Arg Pro Gly Thr Val Ile Ser Tyr Ile Val Leu
690 695 700

Lys Gly Ser Gly Arg Ile Gly Asp Arg Ala Ile Pro Ala Asp Glu Phe
705 710 715 720

Asp Pro Thr Lys His Arg Tyr Asp Ala Glu Tyr Tyr Ile Glu Asn Gln
725 730 735

Val Leu Pro Ala Val Glu Ala Ile Leu Lys Ala Phe Gly Tyr Arg Lys
740 745 750

Glu Asp Leu Arg Tyr Gln Lys Thr Lys Gln Val Gly Leu Gly Ala Trp
755 760 765

Leu Lys Val Lys Gly Lys Lys
770 775

<210> 6
<211> 2328
<212> DNA

<213> Artificial sequence

<220>
<223> Mutant polymerase

<400> 6

atgattctcg ataccgacta catcaccgag aacgggaagc ccgtgataag ggtcttcaag 60

aaggagaacg gcgagtttaa aatcgagtac gacagaacct tcgagccCta cttctacgcc 120

cttctgaagg acgattctgc gatagaggac gtcaagaagg taaccgcaaa gaggcacgga 180

acggttgtca aggtgaagcg cgccgagaag gtgcagaaga agttcctcgg caggccgata 240

gaggtctgga agctctactt caaccatcct caggacgtcc cggcgattcg agacaggata 300

cgtgcccacc ccgctgtcgt tgacatctac gagtacgaca tacccttcgc caagcgctac 360

ctcatcgaca agggcctgat tccgatggag ggcgacgagg agcttacgat gctcgccttc 420

gcgatcgcaa ccctctatca cgagggcgag gagttcggaa ccgggccgat tctcatgata 480

agctacgccg acgggagcga ggcgagggtg ataacctgga agaagattga ccttccgtac 540

gttgacgtcg tctcgaccga gaaggagatg attaagcgct tcctccgcgt cgtcagggag 600

aaggaccccg acgtgctcat cacctacaac ggcgacaact tcgacttcgc ctacctgaag 660

aagcgctctg aggaactcgg aataaagttc acactcggca gggacgggag cgagccgaag 720

atacagcgaa tgggcgaccg ctttgccgtt gaggtgaagg gcaggattca cttcgacctc 780

taccccgtca taaggcgcac gataaacctc ccgacctaca cccttgaggc cgtttacgag 840

gccgtctttg gaaagcccaa ggagaaggtt tacgcagagg agatagcgca ggcctgggag 900

agcggggagg gccttgaaag ggttgcaaga tactcgatgg aggacgctaa ggtgacctac 960

gagctgggaa gggagttctt cccgatggag gcccagcttt cgaggcttat aggccagagc 1020

ctctgggacg tctcgcgctc gagcaccgga aatttggtgg agtggttcct cctgcggaag 1080

gcctacaaga ggaacgagct cgccccaaac aagcccgacg agagggagct cgcgagacgg 1140

cgcgggggct acgctggcgg gtacgttaag gaaccagagc ggggattgtg ggacaacatt 1200

gtgtatctag acttccgctc gtatgcggtt tcaatcatca taacccacaa cgtctcgccg 1260

gataccctca accgcgaggg ctgtaaagag tacgacgtcg cccctgaggt tggacacaag 1320

ttctgcaagg acttccccgg cttcatacca agcctcctgg gagatttgct cgaggagagg 1380

cagaagataa agcggaagat gaaggcaacg gttgacccgc tggagaagaa actcctcgat 1440

tacaggcaga ggctgatcaa aatcctcgcc aacagcttct acggctacta cggctacgcc 1500

aaggcccggt ggtactgcaa ggagtgcgcc gagagcgtta cggcctgggg aagggagtat 1560

atagaaatgg ttatccggga actcgaagaa aaattcggtt ttaaagttct ctatgccgat 1620

acagacggtc tccatgctac cattcccgga gcagacgctg aaacagtcaa gaaaaaagca 1680

aaggagttct taaaatacat taatccaaaa ctgcccggcc tgctcgaact tgagtacgag 1740

ggcttctacg tgaggggctt cttcgtcacg aagaagaagt acgctgtgat agacgaggag 1800

ggcaagataa ccacgagggg tcttgagatt gtgaggcgcg actggagcga gatagcgaag 1860

gagacccagg ccagggtctt agaggcgata ctcaagcacg gtgacgtcga ggaggccgtt 1920

aggatagtca aggaagtgac ggaaaagctg agcaagtatg aggtcccgcc cgagaagctg 1980

gtaatccacg agcagataac gcgcgatttg agggattaca aagccaccgg cccgcacgtt 2040

gccgttgcga agaggctcgc ggcgcgtgga gtgaaaatcc ggcccggcac ggtgataagc 2100

tacatcgtcc taaagggctc tggaaggata ggcgacaggg cgattccagc tgatgagttc 2160

gacccgacga agcaccgcta cgatgcggaa tactacatcg agaaccaggt tctcccggcg 2220

gtggaggcga ttctaaaagc cttcggctat cggaaggagg atttgcgcta ccagaagacg 2280

aagcaggtcg gcttgggcgc gtggctgaag gtgaagggga agaagtga 2328

<210> 7
<211> 704
<212> PRT
<213> Artificial sequence

<220>
<223> Mutant polymerase

<400> 7

Met Ile Leu Asp Thr Asp Tyr Ile Thr Glu Asn Gly Lys Pro Val Ile
1 5 10 15

Arg Val Phe Lys Lys Glu Asn Gly Glu Phe Lys Ile Glu Tyr Asp Arg
20 25 30

Thr Phe Glu Pro Tyr Phe Tyr Ala Leu Leu Lys Asp Asp Ser Ala Ile
35 40 45

Glu Asp Val Lys Lys Val Thr Ala Lys Arg His Gly Thr Val Val Lys
50 55 60

Val Lys Arg Ala Glu Lys Val Gln Lys Lys Phe Leu Gly Arg Pro Ile
65 70 75 80

Glu Val Trp Lys Leu Tyr Phe Asn His Pro Gln Asp Val Pro Ala Ile

85 90 95

Arg Asp Arg Ile Arg Ala His Pro Ala Val Val Asp Ile Tyr Glu Tyr
100 105 110

Asp Ile Pro Phe Ala Lys Arg Tyr Leu Ile Asp Lys Gly Leu Ile Pro
115 120 125

Met Glu Gly Asp Glu Glu Leu Thr Met Leu Ala Phe Ala Ile Ala Thr
130 135 140

Leu Tyr His Glu Gly Glu Glu Phe Gly Thr Gly Pro Ile Leu Met Ile
145 150 155 160

Ser Tyr Ala Asp Gly Ser Glu Ala Arg Val Ile Thr Trp Lys Lys Ile
165 170 175

Asp Leu Pro Tyr Val Asp Val Val Ser Thr Glu Lys Glu Met Ile Lys
180 185 190

Arg Phe Leu Arg Val Val Arg Glu Lys Asp Pro Asp Val Leu Ile Thr
195 200 205

Tyr Asn Gly Asp Asn Phe Asp Phe Ala Tyr Leu Lys Lys Arg Ser Glu
210 215 220

Glu Leu Gly Ile Lys Phe Thr Leu Gly Arg Asp Gly Ser Glu Pro Lys
225 230 235 240

Ile Gln Arg Met Gly Asp Arg Phe Ala Val Glu Val Lys Gly Arg Ile
245 250 255

His Phe Asp Leu Tyr Pro Val Ile Arg Arg Thr Ile Asn Leu Pro Thr
260 265 270

Tyr Thr Leu Glu Ala Val Tyr Glu Ala Val Phe Gly Lys Pro Lys Glu
275 280 285

Lys Val Tyr Ala Glu Glu Ile Ala Gln Ala Trp Glu Ser Gly Glu Gly
290 295 300

Leu Glu Arg Val Ala Arg Tyr Ser Met Glu Asp Ala Lys Val Thr Tyr
305 310 315 320

Glu Leu Gly Arg Glu Phe Phe Pro Met Glu Ala Gln Leu Ser Arg Leu
325 330 335

Ile Gly Gln Ser Leu Trp Asp Val Ser Arg Ser Ser Thr Gly Asn Leu
340 345 350

Val Glu Trp Phe Leu Leu Arg Lys Ala Tyr Lys Arg Asn Glu Leu Ala
355 360 365

Pro Asn Lys Pro Asp Glu Arg Glu Leu Ala Arg Arg Arg Gly Gly Tyr
370 375 380

Ala Gly Gly Tyr Val Lys Glu Pro Glu Arg Gly Leu Trp Asp Asn Ile
385 390 395 400

Val Tyr Leu Asp Phe Arg Ser Tyr Ala Val Ser Ile Ile Ile Thr His
405 410 415

Asn Val Ser Pro Asp Thr Leu Asn Arg Glu Gly Cys Lys Glu Tyr Asp
420 425 430

Val Ala Pro Glu Val Gly His Lys Phe Cys Lys Asp Phe Pro Gly Phe
435 440 445

Ile Pro Ser Leu Leu Gly Asp Leu Leu Glu Glu Arg Gln Lys Ile Lys
450 455 460

Arg Lys Met Lys Ala Thr Val Asp Pro Leu Glu Lys Lys Leu Leu Asp
465 470 475 480

Tyr Arg Gln Arg Leu Ile Lys Ile Leu Ala Asn Ser Phe Tyr Gly Tyr
485 490 495

Tyr Gly Tyr Ala Lys Ala Arg Trp Tyr Cys Lys Glu Cys Ala Glu Ser
500 505 510

Val Thr Ala Trp Gly Arg Glu Tyr Ile Glu Met Val Ile Arg Glu Leu
515 520 525

Glu Glu Lys Phe Gly Phe Lys Val Leu Tyr Ala Asp Thr Asp Gly Leu
530 535 540

His Ala Thr Ile Pro Gly Ala Asp Ala Glu Thr Val Lys Lys Lys Ala

545 550 555 560

Lys Glu Phe Leu Lys Tyr Ile Asn Pro Lys Leu Pro Gly Leu Leu Glu
565 570 575

Leu Glu Tyr Glu Gly Phe Tyr Val Arg Gly Phe Phe Val Thr Lys Lys
580 585 590

Lys Tyr Ala Val Ile Asp Glu Glu Gly Lys Ile Thr Thr Arg Gly Leu
595 600 605

Glu Ile Val Arg Arg Asp Trp Ser Glu Ile Ala Lys Glu Thr Gln Ala
610 615 620

Arg Val Leu Glu Ala Ile Leu Lys His Gly Asp Val Glu Glu Ala Val
625 630 635 640

Arg Ile Val Lys Glu Val Thr Glu Lys Leu Ser Lys Tyr Glu Val Pro
645 650 655

Pro Glu Lys Leu Val Ile His Glu Gln Ile Thr Arg Asp Leu Arg Asp
660 665 670

Tyr Lys Ala Thr Gly Pro His Val Ala Val Ala Lys Arg Leu Ala Ala
675 680 685

Arg Gly Val Lys Ile Arg Pro Gly Thr Val Ile Ser Tyr Ile Val Leu
690 695 700

<210> 8
<211> 2328
<212> DNA
<213> Artificial sequence

<220>
<223> Mutant polymerase

<400> 8
atgattctcg ataccgacta catcaccgag aacgggaagc ccgtgataag ggtcttcaag 60

aaggagaacg gcgagtttaa aatcgagtac gacagaacct tcgagcccta cttctacgcc 120

cttctgaagg acgattctgc gatagaggac gtcaagaagg taaccgcaaa gaggcacgga 180

acggttgtca aggtgaagcg cgccgagaag gtgcagaaga agttcctcgg caggccgata 240

gaggtctgga agctctactt caaccatcct caggacgtcc cggcgattcg agacaggata 300

cgtgcccacc ccgctgtcgt tgacatctac gagtacgaca taccttcgc caagcgctac 360

ctcatcgaca agggcctgat tccgatggag ggcgacgagg agcttacgat gctcgccttc 420

gcgatcgcaa ccctctatca cgagggcgag gagttcggaa ccgggccgat tctcatgata 480

agctacgccg acgggagcga ggcgagggtg ataacctgga agaagattga ccttccgtac 540

gttgacgtcg tctcgaccga gaaggagatg attaagcgct tcctccgcgt cgtcagggag 600

aaggaccccg acgtgctcat cacctacaac ggcgacaact tcgacttcgc ctacctgaag 660

aagcgctctg aggaactcgg aataaagttc acactcggca gggacgggag cgagccgaag 720

atacagcgaa tgggcgaccg ctttgccgtt gaggtgaagg gcaggattca cttcgacctc 780

taccccgtca taaggcgcac gataaacctc ccgacctaca cccttgaggc cgtttacgag 840

gccgtctttg gaaagcccaa ggagaaggtt tacgcagagg agatagcgca ggcctgggag 900

agcggggagg gccttgaaag ggttgcaaga tactcgatgg aggacgctaa ggtgacctac 960

gagctgggaa gggagttctt cccgatggag gcccagcttt cgaggcttat aggccagagc 1020

ctctgggacg tctcgcgctc gagcaccgga aatttggtgg agtggttcct cctgcggaag 1080

gcctacaaga ggaacgagct cgccccaaac aagcccgacg agagggagct cgcgagacgg 1140

cgcgggggct acgctggcgg gtacgttaag gaaccagagc ggggattgtg ggacaacatt 1200

gtgtatctag acttccgctc gtatgcggtt tcaatcatca taacccacaa cgtctcgccg 1260

gataccctca accgcgaggg ctgtaaagag tacgacgtcg cccctgaggt tggacacaag 1320

ttctgcaagg acttccccgg cttcatacca agcctcctgg gagatttgct cgaggagagg 1380

cagaagataa agcggaagat gaaggcaacg gttgacccgc tggagaagaa actcctcgat 1440

tacaggcaga ggctgatcaa aatcctcgcc aacagcttct acggctacta cggctacgcc 1500

aaggcccggt ggtactgcaa ggagtgcgcc gagagcgtta cggcctgggg aagggagtat 1560

atagaaatgg ttatccggga actcgaagaa aaattcggtt ttaaagttct ctatgccgat 1620

acagacggtc tccatgctac cattcccgga gcagacgctg aaacagtcaa gaaaaaagca 1680

aaggagttct taaaatacat taatccaaaa ctgcccggcc tgctcgaact tgagtacgag 1740

ggcttctacg tgaggggctt cttcgtcacg aagaagaagt acgctgtgat agacgaggag 1800

ggcaagataa ccacgagggg tcttgagatt gtgaggcgcg actggagcga gatagcgaag 1860

gagacccagg ccagggtctt agaggcgata ctcaagcacg gtgacgtcga ggaggccgtt 1920

aggatagtca aggaagtgac ggaaaagctg agcaagtatg aggtcccgcc cgagaagctg 1980

gtaatccacg agcagataac gcgcgatttg agggattaca aagccaccgg cccgcacgtt 2040

gccgttgcga agaggctcgc ggcgcgtgga gtgaaaatcc ggcccggcac ggtgataagc 2100

tacatcgtcc tgacgggctc tggaaggata ggcgacaggg cgattccagc tgatgagttc 2160

gacccgacga agcaccgcta cgatgcggaa tactacatcg agaaccaggt tctcccggcg 2220

gtggagagga ttctaaaagc cttcggctat cggaaggagg atttgcgcta ccagaagacg 2280

aagcaggtcg gcttgggcgc gtggctgaag gtgaagggga agaagtga 2328

<210> 9
<211> 28
<212> DNA
<213> Artificial sequence

<220>
<223> Fwd primer

<400> 9
cccggcggtg gaggcgattc taaaagcc 28

<210> 10
<211> 28
<212> DNA
<213> Artificial sequence

<220>
<223> Rev primer

<400> 10
gggccgccac ctccgctaag attttcgg 28

<210> 11
<211> 30
<212> DNA
<213> Artificial sequence

<220>
<223> Fwd primer

<400> 11
gaaggatagg cgacgcggcg attccagctg 30

<210> 12
<211> 30
<212> DNA
<213> Artificial sequence

<220>
<223> Rev primer

<400> 12
cttcctatcc gctgcgccgc taaggtcgac 30

<210> 13
<211> 29
<212> DNA
<213> Artificial sequence

<220>
<223> Fwd primer

<400> 13
gctacatcgt cctagcgggc tctggaagg 29

<210> 14
<211> 29
<212> DNA
<213> Artificial sequence

<220>
<223> Rev primer

<400> 14
cgatgtagca ggatcgcccg agaccttcc 29

<210> 15
<211> 29
<212> DNA
<213> Artificial sequence

<220>
<223> Fwd primer

<400> 15
gctacatcgt cctatgaggc tctggaagg 29

<210> 16
<211> 29
<212> DNA
<213> Artificial sequence

<220>
<223> Rev primer

<400> 16
cgatgtagca ggatactccg agaccttcc 29

<210> 17
<211> 84
<212> DNA
<213> Artificial sequence

<220>
<223> Template DNA

<400> 17

cgatcacgat cacgatcacg atcacgatca cgatcacgct gatgtgcatg ctgttgtttt 60

tttacaacag catgcacatc agcg 84

<210> 18
<211> 84
<212> DNA
<213> Artificial sequence

<220>
<223> NH2 coupled template

<220>
<221> misc_feature
<222> (1)..(1)
<223> NH2 group attached to 5' end

<400> 18
cgatcacgat cacgatcacg atcacgatca cgatcacgct gatgtgcatg ctgttgtttt 60

tttacaacag catgcacatc agcg 84

<210> 19
<211> 2328
<212> DNA
<213> Artificial sequence

<220>
<223> Codon optimised polymerase

<400> 19
atgatcttag ataccgacta tatcaccgag aacggtaaac cggtgataag ggtgttcaaa 60

aaggaaaatg gcgaattcaa gatcgagtat gatagaacct tcgaaccgta cttctacgcc 120

ttgttgaagg acgatagtgc catcgaagat gtgaaaaaag ttaccgccaa acgtcacggc 180

accgtggtaa aggttaaacg cgccgaaaag gttcagaaga agttcctagg ccgtccgatc 240

gaggtgtgga aattgtactt taaccatccg caggatgtcc cggcgattag agatcgtatt 300

cgtgcccacc cggcggtagt ggatatctat gagtacgata tcccgttcgc aaaaagatac 360

ttgattgata aaggactaat cccgatggaa ggcgatgaag aattaaccat gttagcgttc 420

tccatctcca ccctgtacca cgaaggcgaa gagttcggca ccggtccgat tctgatgatc 480

tcctacgcag acggtagcga agcacgtgtg ataacctgga agaaaataga cctaccttac 540

gtggacgtcg taagtaccga gaaggagatg atcaaaagat tcctgagggt ggtccgtgag 600

aaggatccgg acgtactgat tacctataac ggcgataact tcgacttcgc ctacttgaaa 660

aagagatctg aggaattagg catcaaattc accctgggcc gtgatggcag tgagccgaaa 720

atccaacgta tgggcgaccg cttcgccgtc gaggtgaaag gccgtataca tttcgacttg 780

tatccggtga ttaggcgtac cattaatttg ccgacctaca ccttggaagc ggtgtacgag 840

gcggtcttcg gcaagccgaa ggaaaaggtg tacgccgaag agatcgcgca ggcgtgggag 900

agcggtgagg gtctagaacg tgttgcaaga tatagcatgg aggacgccaa agttacctac 960

gaattgggcc gcgagttttt tccgatggag gcccagttat ctcgtttaat tggccagtcc 1020

ctgtgggatg ttagccgcag ttctactggt aatttggtag aatggttctt actgcgcaaa 1080

gcgtataaac gtaacgagtt agcgccaaat aagccggacg aacgtgaact ggcccgtcgt 1140

cgtggtggct atgccggcgg ttacgtgaag gaaccggagc gtggcctatg ggataacatt 1200

gtgtaccttg actttagaag ctatgcggtt agcatcatca tcacccataa tgttagtccg 1260

gacacattga atcgtgaagg atgcaaagaa tatgacgtcg ccccagaggt gggccacaaa 1320

ttttgtaaag atttcccagg attcatccca agtttgttgg gtgatctgct ggaagaacgc 1380

cagaaaatca aacgtaagat gaaggcgacc gtcgatccac tggagaaaaa gctattggac 1440

taccgtcagc gcctgatcaa gattttggcg aattctttct atggatacta cggctacgcc 1500

aaagcccgtt ggtattgtaa agagtgcgcc gagtctgtca ctgcctgggg tcgtgaatat 1560

atcgaaatgg tgatccgcga gctggaagag aaatttggat tcaaagtctt gtacgccgat 1620

accgatggtc tgcacgcgac cattccgggt gccgatgccg agaccgtgaa gaaaaaggcg 1680

aaagagtttt tgaaatatat caatccgaag ttgccgggat tattagaatt ggaatacgaa 1740

ggtttctatg ttcgcggctt tttcgtgacc aagaaaaaat acgccgtgat cgacgaggaa 1800

ggaaaaatta ccacccgtgg tctagagatt gttcgtcgtg actggtccga aatcgccaaa 1860

gaaacccagg cccgtgtact ggaagcgatt ttgaagcatg gcgatgtgga ggaggcggtt 1920

cgtatcgtca aagaagtgac cgaaaagctg agcaagtatg aagtgccgcc ggagaaattg 1980

gtcatacacg aacaaatcac acgtgacctg cgcgattata aggcgaccgg tccgcacgtt 2040

gccgtggcga agcgtttggc ggcccgtggt gttaagattc gtccaggaac cgtgattagt 2100

tacatagtgt tgaagggcag tggtcgtatt ggtgaccgtg ccatcccggc ggatgagttt 2160

gacccgacca agcatcgtta tgacgccgaa tattatatcg agaatcaggt gctaccagcg 2220

gttgaacgta ttttgaaggc attcggctat cgtaaagaag acctgcgcta ccagaaaacc 2280

aagcaggttg gtctgggtgc ctggttgaaa gtgaaaggca aaaaataa 2328

<210> 20
<211> 2328

<212> DNA
<213> Artificial sequence

<220>
<223> Codon optimised thumb mutant polymerase

<400> 20
atgatcttag ataccgacta tatcaccgag aacggtaaac cggtgataag ggtgttcaaa 60

aaggaaaatg gcgaattcaa gatcgagtat gatagaacct tcgaaccgta cttctacgcc 120

ttgttgaagg acgatagtgc catcgaagat gtgaaaaaag ttaccgccaa acgtcacggc 180

accgtggtaa aggttaaacg cgccgaaaag gttcagaaga agttcctagg ccgtccgatc 240

gaggtgtgga aattgtactt taaccatccg caggatgtcc cggcgattag agatcgtatt 300

cgtgcccacc cggcggtagt ggatatctat gagtacgata tcccgttcgc aaaaagatac 360

ttgattgata aaggactaat cccgatggaa ggcgatgaag aattaaccat gttagcgttc 420

tccatctcca ccctgtacca cgaaggcgaa gagttcggca ccggtccgat tctgatgatc 480

tcctacgcag acggtagcga agcacgtgtg ataacctgga agaaaatага cctaccttac 540

gtggacgtcg taagtaccga gaaggagatg atcaaaagat tcctgagggt ggtccgtgag 600

aaggatccgg acgtactgat tacctataac ggcgataact tcgacttcgc ctacttgaaa 660

aagagatctg aggaattagg catcaaattc accctgggcc gtgatggcag tgagccgaaa 720

atccaacgta tgggcgaccg cttcgccgtc gaggtgaaag gccgtataca tttcgacttg 780

tatccggtga ttaggcgtac cattaatttg ccgacctaca ccttggaagc ggtgtacgag 840

gcggtcttcg gcaagccgaa ggaaaaggtg tacgccgaag agatcgcgca ggcgtgggag 900

agcggtgagg gtctagaacg tgttgcaaga tatagcatgg aggacgccaa agttacctac 960

gaattgggcc gcgagttttt tccgatggag gcccagttat ctcgtttaat tggccagtcc 1020

ctgtgggatg ttagccgcag ttctactggt aatttggtag aatggttctt actgcgcaaa 1080

gcgtataaac gtaacgagtt agcgccaaat aagccggacg aacgtgaact ggcccgtcgt 1140

cgtggtggct atgccggcgg ttacgtgaag gaaccggagc gtggcctatg ggataacatt 1200

gtgtaccttg actttagaag ctatgcggtt agcatcatca tcacccataa tgttagtccg 1260

gacacattga atcgtgaagg atgcaaagaa tatgacgtcg ccccagaggt gggccacaaa 1320

ttttgtaaag atttcccagg attcatccca agtttgttgg gtgatctgct ggaagaacgc 1380

cagaaaatca aacgtaagat gaaggcgacc gtcgatccac tggagaaaaa gctattggac 1440

taccgtcagc gcctgatcaa gattttggcg aattctttct atggatacta cggctacgcc 1500

aaagcccgtt ggtattgtaa agagtgcgcc gagtctgtca ctgcctgggg tcgtgaatat 1560

atcgaaatgg tgatccgcga gctggaagag aaatttggat tcaaagtctt gtacgccgat 1620

accgatggtc tgcacgcgac cattccgggt gccgatgccg agaccgtgaa gaaaaaggcg 1680

aaagagtttt tgaaatatat caatccgaag ttgccgggat tattagaatt ggaatacgaa 1740

ggtttctatg ttcgcggctt tttcgtgacc aagaaaaaat acgccgtgat cgacgaggaa 1800

ggaaaaatta ccacccgtgg tctagagatt gttcgtcgtg actggtccga aatcgccaaa 1860

gaaacccagg cccgtgtact ggaagcgatt ttgaagcatg gcgatgtgga ggaggcggtt 1920

cgtatcgtca aagaagtgac cgaaaagctg agcaagtatg aagtgccgcc ggagaaattg 1980

gtcatacacg aacaaatcac acgtgacctg cgcgattata aggcgaccgg tccgcacgtt 2040

gccgtggcga agcgtttggc ggcccgtggt gttaagattc gtccaggaac cgtgattagt 2100

tacatagtgt tgaagggcag tggtcgtatt ggtgaccgtg ccatcccggc ggatgagttt 2160

gacccgacca agcatcgtta tgacgccgaa tattatatcg agaatcaggt gctaccagcg 2220

gttgaagcta ttttgaaggc attcggctat cgtaaagaag acctgcgcta ccagaaaacc 2280

aagcaggttg gtctgggtgc ctggttgaaa gtgaaaggca aaaaataa 2328

<210> 21
<211> 775
<212> PRT
<213> Artificial sequence

<220>
<223> Codon optimised thumb mutant polymerase

<400> 21

Met Ile Leu Asp Thr Asp Tyr Ile Thr Glu Asn Gly Lys Pro Val Ile
1 5 10 15

Arg Val Phe Lys Lys Glu Asn Gly Glu Phe Lys Ile Glu Tyr Asp Arg
20 25 30

Thr Phe Glu Pro Tyr Phe Tyr Ala Leu Leu Lys Asp Asp Ser Ala Ile
35 40 45

Glu Asp Val Lys Lys Val Thr Ala Lys Arg His Gly Thr Val Val Lys
50 55 60

Val Lys Arg Ala Glu Lys Val Gln Lys Lys Phe Leu Gly Arg Pro Ile
65 70 75 80

```
Glu Val Trp Lys Leu Tyr Phe Asn His Pro Gln Asp Val Pro Ala Ile
                85                  90                  95

Arg Asp Arg Ile Arg Ala His Pro Ala Val Val Asp Ile Tyr Glu Tyr
            100                 105                 110

Asp Ile Pro Phe Ala Lys Arg Tyr Leu Ile Asp Lys Gly Leu Ile Pro
        115                 120                 125

Met Glu Gly Asp Glu Glu Leu Thr Met Leu Ala Phe Ser Ile Ser Thr
    130                 135                 140

Leu Tyr His Glu Gly Glu Glu Phe Gly Thr Gly Pro Ile Leu Met Ile
145                 150                 155                 160

Ser Tyr Ala Asp Gly Ser Glu Ala Arg Val Ile Thr Trp Lys Lys Ile
                165                 170                 175

Asp Leu Pro Tyr Val Asp Val Val Ser Thr Glu Lys Glu Met Ile Lys
            180                 185                 190

Arg Phe Leu Arg Val Val Arg Glu Lys Asp Pro Asp Val Leu Ile Thr
        195                 200                 205

Tyr Asn Gly Asp Asn Phe Asp Phe Ala Tyr Leu Lys Lys Arg Ser Glu
    210                 215                 220

Glu Leu Gly Ile Lys Phe Thr Leu Gly Arg Asp Gly Ser Glu Pro Lys
225                 230                 235                 240

Ile Gln Arg Met Gly Asp Arg Phe Ala Val Glu Val Lys Gly Arg Ile
                245                 250                 255

His Phe Asp Leu Tyr Pro Val Ile Arg Arg Thr Ile Asn Leu Pro Thr
            260                 265                 270

Tyr Thr Leu Glu Ala Val Tyr Glu Ala Val Phe Gly Lys Pro Lys Glu
        275                 280                 285

Lys Val Tyr Ala Glu Glu Ile Ala Gln Ala Trp Glu Ser Gly Glu Gly
    290                 295                 300

Leu Glu Arg Val Ala Arg Tyr Ser Met Glu Asp Ala Lys Val Thr Tyr
```

305 310 315 320

Glu Leu Gly Arg Glu Phe Phe Pro Met Glu Ala Gln Leu Ser Arg Leu
325 330 335

Ile Gly Gln Ser Leu Trp Asp Val Ser Arg Ser Ser Thr Gly Asn Leu
340 345 350

Val Glu Trp Phe Leu Leu Arg Lys Ala Tyr Lys Arg Asn Glu Leu Ala
355 360 365

Pro Asn Lys Pro Asp Glu Arg Glu Leu Ala Arg Arg Arg Gly Gly Tyr
370 375 380

Ala Gly Gly Tyr Val Lys Glu Pro Glu Arg Gly Leu Trp Asp Asn Ile
385 390 395 400

Val Tyr Leu Asp Phe Arg Ser Tyr Ala Val Ser Ile Ile Ile Thr His
405 410 415

Asn Val Ser Pro Asp Thr Leu Asn Arg Glu Gly Cys Lys Glu Tyr Asp
420 425 430

Val Ala Pro Glu Val Gly His Lys Phe Cys Lys Asp Phe Pro Gly Phe
435 440 445

Ile Pro Ser Leu Leu Gly Asp Leu Leu Glu Glu Arg Gln Lys Ile Lys
450 455 460

Arg Lys Met Lys Ala Thr Val Asp Pro Leu Glu Lys Lys Leu Leu Asp
465 470 475 480

Tyr Arg Gln Arg Leu Ile Lys Ile Leu Ala Asn Ser Phe Tyr Gly Tyr
485 490 495

Tyr Gly Tyr Ala Lys Ala Arg Trp Tyr Cys Lys Glu Cys Ala Glu Ser
500 505 510

Val Thr Ala Trp Gly Arg Glu Tyr Ile Glu Met Val Ile Arg Glu Leu
515 520 525

Glu Glu Lys Phe Gly Phe Lys Val Leu Tyr Ala Asp Thr Asp Gly Leu
530 535 540

```
His Ala Thr Ile Pro Gly Ala Asp Ala Glu Thr Val Lys Lys Lys Ala
545                 550                 555                 560

Lys Glu Phe Leu Lys Tyr Ile Asn Pro Lys Leu Pro Gly Leu Leu Glu
                565                 570                 575

Leu Glu Tyr Glu Gly Phe Tyr Val Arg Gly Phe Phe Val Thr Lys Lys
            580                 585                 590

Lys Tyr Ala Val Ile Asp Glu Glu Gly Lys Ile Thr Thr Arg Gly Leu
        595                 600                 605

Glu Ile Val Arg Arg Asp Trp Ser Glu Ile Ala Lys Glu Thr Gln Ala
    610                 615                 620

Arg Val Leu Glu Ala Ile Leu Lys His Gly Asp Val Glu Glu Ala Val
625                 630                 635                 640

Arg Ile Val Lys Glu Val Thr Glu Lys Leu Ser Lys Tyr Glu Val Pro
                645                 650                 655

Pro Glu Lys Leu Val Ile His Glu Gln Ile Thr Arg Asp Leu Arg Asp
            660                 665                 670

Tyr Lys Ala Thr Gly Pro His Val Ala Val Ala Lys Arg Leu Ala Ala
        675                 680                 685

Arg Gly Val Lys Ile Arg Pro Gly Thr Val Ile Ser Tyr Ile Val Leu
    690                 695                 700

Lys Gly Ser Gly Arg Ile Gly Asp Arg Ala Ile Pro Ala Asp Glu Phe
705                 710                 715                 720

Asp Pro Thr Lys His Arg Tyr Asp Ala Glu Tyr Tyr Ile Glu Asn Gln
                725                 730                 735

Val Leu Pro Ala Val Glu Ala Ile Leu Lys Ala Phe Gly Tyr Arg Lys
            740                 745                 750

Glu Asp Leu Arg Tyr Gln Lys Thr Lys Gln Val Gly Leu Gly Ala Trp
        755                 760                 765

Leu Lys Val Lys Gly Lys Lys
```

770 775

INTERNATIONAL SEARCH REPORT

International application No
PCT/GB2006/001700

A. CLASSIFICATION OF SUBJECT MATTER
INV. C12N9/12

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)
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 practical, search terms used)
EPO-Internal, Sequence Search, EMBASE, WPI Data, PAJ, BIOSIS

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	GARDNER ANDREW F ET AL: "Determinants of nucleotide sugar recognition in an archaeon DNA polymerase" NUCLEIC ACIDS RESEARCH, OXFORD UNIVERSITY PRESS, SURREY, GB, vol. 27, no. 12, 15 June 1999 (1999-06-15), pages 2545-2553, XP002163101 ISSN: 0305-1048 abstract; table 2 ----- -/--	

[X] Further documents are listed in the continuation of Box C.

[X] 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 document 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	Date of mailing of the international search report
25 July 2006	11/08/2006

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL – 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Weinberg, S

2

Form PCT/ISA/210 (second sheet) (April 2005)

INTERNATIONAL SEARCH REPORT

International application No
PCT/GB2006/001700

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	MINNICK D T ET AL: "A thumb subdomain mutant of the large fragment of Escherichia coli DNA polymerase I with reduced DNA binding affinity, processivity, and frameshift fidelity." THE JOURNAL OF BIOLOGICAL CHEMISTRY. 4 OCT 1996, vol. 271, no. 40, 4 October 1996 (1996-10-04), pages 24954-24961, XP002391794 ISSN: 0021-9258	
A	WO 2005/024010 A (SOLEXA LIMITED; SMITH, GEOFFREY PAUL; BAILEY, DAVID, MARK, DUNSTAN; SA) 17 March 2005 (2005-03-17) table 1	
A	RODRIGUEZ A C ET AL: "Crystal structure of a pol alpha family DNA polymerase from the hyperthermophilic archaeon Thermococcus sp. 9<o>N-7" JOURNAL OF MOLECULAR BIOLOGY, LONDON, GB, vol. 299, no. 2, 2 June 2000 (2000-06-02), pages 447-462, XP004469030 ISSN: 0022-2836	
A	SOUTHWORTH M W ET AL: "CLONING OF THERMOSTABLE DNA POLYMERASES FROM HYPERTHERMOPHILIC MARINE ARCHEA WITH EMPHASIS ON THERMOCOCCUS SP. 9 N-7 AND MUTATIONS AFFECTING 3'-5' EXONUCLEASE ACTIVITY" PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA, NATIONAL ACADEMY OF SCIENCE, WASHINGTON, DC, US, vol. 93, no. 11, 28 May 1996 (1996-05-28), pages 5281-5285, XP000652319 ISSN: 0027-8424	

2

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box II.2

Present claims 1-18 relate to a compound which is defined purely by reference to a given desired property or effect, namely a polymerase having a reduced affinity for DNA such that the polymerase has an ability to incorporate a nucleotide or nucleotides into a plurality of separate DNA templates in each reaction cycle compared to a control polymerase. However, the description provides support and disclosure in the sense of Article 6 and 5 PCT for only a limited number of any such compounds having the said property or effect. This non-compliance with the substantive provisions is to such an extent, that the search for these claims only be performed insofar as the claims relate to subject-matter which is disclosed and supported, namely, where the polymerase has a substitution at the positions Lys705, Arg713 or 743 in the 9 degree N polymerase, or at positions Lys790, Lys800, Arg806, Lys844, Lys874 or Lys878 of the RB69 DNA polymerase (PCT Guidelines 9.19 and 9.20).

Claims 61-65 define the subject-matter for which protection is sought by merely referring to the entire description and all figures attached to the application without explicitly referring to any technical feature. Hence, the definition of the subject-matter of these claims is totally unclear, and does not comply with Article 6 PCT and with Rule 6.2(a) PCT. As a result, Claims 61-65 could not be searched.

The applicant's attention is drawn to the fact that claims relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure. If the application proceeds into the regional phase before the EPO, the applicant is reminded that a search may be carried out during examination before the EPO (see EPO Guideline C-VI, 8.5), should the problems which led to the Article 17(2) declaration be overcome.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/GB2006/001700

Box 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. [X] 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:

see FURTHER INFORMATION sheet PCT/ISA/210

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 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. [X] As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

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.:

Remark on Protest

[ ] The additional search fees were accompanied by the applicant's protest.

[ ] 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: 1-22, 24-60 (partially) and Claims 28-35, 37, 38, 41 (completely)

9 degree N and RB69 polymerase mutants with substitution at any one of positions Lys790, Lys 800, Arg806, Lys844, Lys874 or Lys878 (RB69), Lys705, Arg713 or Arg743 (9 degree N); aspects associated therewith.

---

2. claims: 1-22, 25-27, 36, 39, 40, 42-60 (partially) and Claim 23 (completely)

9 degree N polymerase mutants with substitution at either of positions Arg606 or His679; aspects associated therewith.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/GB2006/001700

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2005024010 A	17-03-2005	EP 1664287 A1	07-06-2006