EUROPEAN PATENT SPECIFICATION

(51) Int Cl.7: C12N 15/54, C12N 9/12, C12Q 1/68 // (C12N9/12, C12R1:145)

(45) Date of publication and mention of the grant of the patent:

(21) Application number: 96944875.2

(22) Date of filing: 13.12.1996

(54) THERMOSTABLE DNA POLYMERASE FROM THERMOANAEROBACTER THERMOHYDROSULFURICUS AND MUTANT ENZYMES THEREOF WITH EXONUCLEASE ACTIVITY REMOVED
THERMOSTABLE DNA POLYMERASE AUS THERMOANAEROBACTER THERMOHYDROSULFURICUS UND DAVON ABGELEITETE MUTIERTE ENZYMEN MIT ENTFERNTER EXONUKLEASEAKTIVITÄT
POLYMERASE D’ADN THERMOSTABLE ISSUE DE THERMOANAEROBACTER THERMOHYDROSULFURICUS ET ENZYMES MUTANTS A ACTIVITE D’EXONUCLEASE SUPPRIMEE AINSI OBTENUS

(84) Designated Contracting States:
AT BE CH DE ES FI FR GB IT LI NL SE


(43) Date of publication of application:

(73) Proprietor: Amersham Life Science Inc
Cleveland, Ohio 44128 (US)

(72) Inventors:
• MAMONE, Joseph, A.
Parma, OH 44134 (US)
• DAVIS, Maria
Twinsburg, OH 44087 (US)
• SHA, Dan
Euclid, OH 44117 (US)

(74) Representative: Rollins, Anthony John
Group Patents,
Amersham International plc,
White Lion Road
Amersham, Bucks HP7 9LL (GB)

(56) References cited:

• BIOCHIMICA ET BIOPHYSICA ACTA, vol. 1307, 1996, pages 178-186, XP000646038 M.G. RIGGS ET AL.: "Construction of single amino acid substitution mutants of cloned Bacillus stearothermophilus DNA polymerase I which lacks 5’-3’ exonuclease activity"
The present invention relates to novel thermostable DNA polymerases obtainable from the thermophilic anaerobe Thermoaerobacter thermohydrosulfuricus, to certain deletants and mutants of this enzyme, to genes and vectors encoding the wild type and mutant polymerases and their use in strand displacement activity, polymerase chain reaction, DNA sequencing and as reverse transcriptases.

The following is a discussion of the relevant art, none of which is admitted to be prior art to the appended claims.

DNA polymerases are a family of enzymes involved in DNA repair and replication. DNA polymerases have been isolated from E.coli (e.g. E.coli DNA polymerase I and the Klenow fragment thereof) and T4 DNA polymerase and more recently thermostable DNA polymerases have been isolated (e.g. from T. aquaticus, US Patent 4,889,818, and from T. litoralis). Thermostable DNA polymerases have been suggested (US Patent 4,683,195) for use in amplifying existing nucleic acid sequences in amounts that are large compared to that originally present.

WO 92/06200 discloses thermostable DNA polymerases from various sources, which exhibit lower levels of exonuclease activity compared with their native polymerases.

The polymerase chain reaction (PCR) and strand displacement amplification (SDA) are two methods of amplifying nucleic acid sequences. PCR is based on the hybridization of oligonucleotide primers to specific sequences on opposite strands of the target DNA molecule, and subsequent extension of these primers with a DNA polymerase to generate two new strands of DNA which themselves can serve as a template for a further round of hybridization and extension. In PCR reactions the product of one cycle serves as the template for the next cycle such that at each repeat of the cycle the amount of the specific sequence present in the reaction doubles leading to an exponential amplification process.

PCR relies on a process of temperature cycling to promote the amplification reaction. This temperature cycling is necessary to separate the strands of DNA formed in one cycle of the reaction to allow hybridization of the oligonucleotides required to initiate the next cycle. DNA strand separation is usually achieved by melting the DNA at temperatures in the range 90-100°C followed by cooling to a lower temperature to allow oligonucleotide hybridization followed by extension or ligation, depending on the reaction process. This cycling process may be repeated 20-50 times, again depending on the process and the degree of amplification required.

Temperature cycling is normally achieved by the use of specialized equipment, commonly termed thermocyclers, which can operate by a wide variety of mechanical, electrical or hydraulic means, but serve a common purpose in heating and cooling a small container or a number of such containers in which the amplification reaction is performed.

In order for the amplification reactions to proceed with the desired efficiency and specificity it is necessary to perform the temperature cycling process within strictly defined and reproducible limits of temperature and time. Failure of the temperature cycling apparatus to achieve the required conditions will result in the partial or total failure of the amplification reaction. These strict requirements on time and temperature of cycling can impose severe restrictions if the handling of large numbers of reactions is required. If it is desired to perform several hundred or more reactions simultaneously use of conventional thermocyclers would be extremely expensive in terms of the capital investment required in equipment, and would in any case be prone to variations between individual thermocyclers. Alternative solutions for large scale use have been constructed based on fan assisted oven or multiple water baths, but such solutions are inevitably cumbersome in use and in any case still require the user to take the normal stringent steps to avoid evaporation which become extremely difficult to apply to a large number of cycles.

In reverse transcription/polymerase chain reaction (RT/PCR), a DNA primer is hybridized to a strand of the target RNA molecule, and subsequent extension of this primer with a reverse transcriptase generates a new strand of DNA which can serve as a template for PCR. Preparation of the DNA template is preferably carried out at an elevated temperature to avoid early termination of the reverse transcriptase reaction caused by RNA secondary structure. There is a lack of efficient reverse transcriptases that act at elevated temperatures, e.g. above 50°C.

SDA differs from PCR in being an isothermal amplification process, i.e. all reactions occur at the same temperature without the need for elevated temperature to melt DNA strands. This is made possible by adoption of a reaction scheme which uses the ability of certain DNA polymerases when extending along a DNA template strand to displace any DNA molecules already hybridized to the template. In SDA this strand displacement is used to separate the double stranded DNA produced during the reaction process and hence to maintain continuous amplification of the target DNA sequence (Walker, G.T., Little, M.C., Nadeau, J.G. and Shank D.D. (1992) Proc. Natl. Acad. Sci. USA 89:392-396). SDA is therefore in principle more suited to use with large numbers of samples than PCR as the isothermal process, which is performed at temperatures of 37°C to 60°C, does not require stringent precautions to be taken to avoid evaporation and can be performed with simple temperature control equipment, for example in a standard laboratory incubator.

DNA polymerases, e.g. Sequenase, Klenow, Taq, etc, have also been extensively used in DNA sequencing,
The present invention provides a thermostable DNA polymerase from *Thermoanaerobacter thermohydrorsulfuricus*. This enzyme is useful for procedures requiring strand-displacing DNA synthesis such as SDA, for DNA sequencing and/or for reverse transcription. Included within the scope of the present invention are various mutants (deletion and substitution) that retain thermostability and the ability to replicate DNA with substantially the same efficiency as the native *Thermoanaerobacter thermohydrorsulfuricus* polymerase.

In a first aspect, the present invention provides a purified DNA polymerase or fragment thereof having the DNA polymerase activity of *Thermoanaerobacter thermohydrorsulfuricus* and having at least 80% amino acid homology, preferably at least 90% homology, to at least a contiguous 40 amino acid sequence shown in Fig. 2A-D (SEQ. ID. NO. 2). Fig. 2A-D represents the translation of an open reading frame spanning nucleotides 1056-3674 of genomic DNA encoding a thermostable DNA polymerase from *Thermoanaerobacter thermohydrorsulfuricus* (Fig. 1A-E) (SEQ. ID. NO. 1) potentially encoding the native polymerase.

When used herein, the term amino acid homology means the amino acid identity of the parent enzyme or conservative amino acid changes thereon.

The purified enzyme of the present invention has a molecular weight of approximately 90,000 daltons when measured on SDS-PAGE. It possesses a 5'-3' exonuclease activity. The temperature optimum of DNA synthesis is near 75°C under assay conditions. The optimum magnesium ion and manganese ion concentrations for DNA synthesis are 1mM and 0.5 mM respectively.

The term thermostable polymerase means an enzyme which is stable to heat (and heat resistant) and is suitable for use in SDA and/or sequencing at an elevated temperature, for example 70°C. The thermostable enzyme herein must satisfy a single criterion to be effective for the amplification reaction, i.e., the enzyme must not become irreversibly denatured (inactivated) when subjected to the elevated temperatures for the time necessary to effect amplification. Irreversible denaturation for purposes herein refers to permanent and complete loss of enzymatic activity. Preferably, the enzyme will not become irreversibly denatured at about 70°C but may become denatured at higher temperatures. The thermostable enzyme herein preferably has an optimum temperature at which it functions that is higher than about 40°C, which is the temperature below which hybridization of primer to template is promoted, although, depending on (1) salt concentration and composition and (2) composition and length of primer, hybridization can occur at higher temperature (e.g., 45-70°C). The higher the temperature optimum for the enzyme, the greater the specificity and/or selectivity of the primer-directed extension process. However, enzymes that are active below 40°C, e.g., at 37°C, are also within the scope of this invention provided they are heat stable. Preferably, the optimum temperature ranges from about 50 to 80°C, more preferably 50-70°C.

When used herein, the term a DNA polymerase or fragment thereof having the DNA polymerase activity of *Thermoanaerobacter thermohydrorsulfuricus* means a DNA polymerase or fragment thereof (as hereinafter defined) which has the ability to replicate DNA with substantially the same efficiency as the enzyme encoded by the SEQ. ID. NO. 1. By substantially the same efficiency is meant at least 80% and preferably at least 90% of the efficiency of the enzyme encoded by SEQ. ID. NO. 1 to incorporate deoxynucleotides.

The invention also encompasses a stable enzyme composition which comprises a purified thermostable DNA polymerase from *Thermoanaerobacter thermohydrorsulfuricus* in a buffer.

The DNA polymerases of the present invention are preferably in a purified form. By purified is meant that the DNA polymerase is isolated from a majority of host cell proteins normally associated with it; preferably the polymerase is at least 10% (w/w), e.g. at least 50% (w/w), of the protein of a preparation, even more preferably it is provided as a homogeneous preparation, e.g., homogeneous solution. Preferably the DNA polymerase is a single polypeptide on an SDS polyacrylamide gel.

Buffers around neutral pH (5--9) such as 5-100 mM Tris-HCl, phosphate or MES are suitable for use in the current invention.

The present invention also provides a gene encoding a polymerase of the present invention. Fig. 1A-E represents nucleotides 1-5300 of a cloned genomic sequence encompassing the gene encoding the polymerase of the present invention (SEQ. ID. NO. 1). Fig. 7A-C represents the DNA sequence of a full length native polymerase from *Thermoanaerobacter thermohydrorsulfuricus* (SEQ. ID. NO. 23).

It has been found that the entire amino acid sequence of the polymerase is not required for enzymatic activity. Thus, for example, the exonuclease domain of the enzyme has been deleted to give an enzyme of molecular weight of approximately 64,000 daltons when measured on SDS-PAGE which retains enzyme activity and also has reverse transcriptase activity making it useful for making cDNA. This exonuclease-free enzyme is analogous to the Klenow fragment of *E. coli* DNA polymerase I. Thus, the present invention also provides fragments of the polymerase which retain the DNA polymerase activity of *Thermoanaerobacter thermohydrorsulfuricus* but have one or more amino acids
deleted, preferably from the amino-terminus, while still having at least 80% amino acid homology to at least a 40 contiguous amino acid sequence shown in Fig. 2A-D (SEQ. ID. NO. 2).

[0024] In a further aspect, the present invention provides a thermostable DNA polymerase which corresponds to the DNA polymerase from *Thermoanaerobacter thermohydrosulfuricus* in which up to one third of the amino acid sequence at the amino-terminus has been deleted. In particular, fragments of *Thermoanaerobacter thermohydrosulfuricus* having N-terminal deletions to give 578 amino acids (See Fig. 4A-C) (SEQ. ID. NO. 4) and 608 amino acids (See Fig. 3) (SEQ. ID. NO. 3) have been found to retain enzyme activity.

[0025] It is preferred that the 5'-3' exonuclease activity of the DNA polymerase is removed or reduced. This may be achieved by deleting the amino acid region of the enzyme responsible for this activity, e.g. by deleting up to one third of the amino acid sequence at the amino terminus, or by appropriate amino acid changes.

[0026] In addition to the N-terminal deletions and amino acid changes to remove the exonuclease activity, the enzyme may have conservative amino acid changes compared with the native enzyme which do not significantly influence thermostability or enzyme activity. Such changes include substitution of like charged amino acids for one another or amino acids with small side chains for other small side chains, e.g. ala for val. More drastic changes may be introduced at non-critical regions where little or no effect on polymerase activity is observed by such a change.

[0027] Joyce and Steitz, *Annu. Rev. Biochem.*, 63:777-822, 1994, discuss various functions of DNA polymerases including the catalytic center, the binding site for the 3' terminus of the primer, and the dNTP binding site. In particular, it mentions mutations that affect the binding of dNTP in the ternary complex. European Patent Application 0655506A discloses that the presence of a polar, hydroxyl containing amino acid residue at a position near the binding site for the dNTP substrate is important for the polymerase being able to efficiently incorporate a dideoxynucleotide. Applicant has discovered that the modification of the dNTP binding site for the dNTP substrate in DNA polymerase obtained from *Thermoanaerobacter thermohydrosulfuricus* by the inclusion of a polar, hydroxyl containing amino acid residue at a position near the binding site increases the efficiency of the polymerase to incorporate a dideoxynucleotide. Preferably the polar, hydroxyl containing amino acid is tyrosine. It has also been found that replacing the phenylalanine at the position corresponding to 706 of the native enzyme with tyrosine improves the incorporation of dideoxynucleotides when the enzyme is used for sequencing. In particular, a polymerase from *Thermoanaerobacter thermohydrosulfuricus* in which the exonuclease activity has been deleted e.g. by point mutation or deletion and which has the phenylalanine at the position corresponding to 706 of the native enzyme replaced by an amino acid which increases the efficiency of the enzyme to incorporate dideoxynucleotides at least 20 fold compared to the wild type enzyme, e.g. tyrosine, is a particularly preferred enzyme for use in sequencing. Preferably, this modified enzyme has from between 540 and 582 amino acids, for example 560 to 582 amino acids and preferably 578 amino acids.

[0028] The DNA polymerases of the present invention can be constructed using standard techniques familiar to those who practice the art. By way of example, in order to prepare a polymerase with the phenylalanine to tyrosine mutation, mutagenic PCR primers can be designed to incorporate the desired Phe to Tyr amino acid change (FY mutation in one primer). Deletion of the exonuclease function is carried out by PCR to remove the amino terminus, or standard techniques of site directed mutagenesis to generate point mutations.

[0029] Improved expression of the DNA polymerases of the present invention can be achieved by introducing silent codon changes (i.e., the amino acid encoded is not changed). Such changes can be introduced by the use of mutagenic PCR primers. Silent codon changes such as the following increase protein production in *E. coli*:

- substitution of the codon GAG for GAA;
- substitution of the codon AGG, AGA, CGG or CGA for CGT or CGC;
- substitution of the codon CTT, CTC, CTA, TTG or TTA for CTG;
- substitution of the codon ATA for ATT or ATC;
- substitution of the codon GGG or GGA for GGT or GGC.

[0030] Genes encoding the DNA polymerase from *Thermoanaerobacter thermohydrosulfuricus* polymerases in which up to one third of the amino acid sequence at the amino terminus has been deleted and such polymerases which incorporate the phenylalanine to tyrosine modification are also provided by the present invention.

[0031] In a yet further aspect, the present invention provides a host cell comprising a vector containing the gene encoding the DNA polymerase activity of the present invention, e.g., encoding an amino acid sequence corresponding to native *Thermoanaerobacter thermohydrosulfuricus* or differentiated from this in that it lacks up to one third of the N-terminal amino acids and optionally has phenylalanine at position 706 replaced by tyrosine.

[0032] The DNA polymerases of the present invention are suitably used in SDA, preferably in combination with a thermostable restriction enzyme. Accordingly, the present invention provides a composition which comprises a DNA polymerase of the present invention in combination with a thermostable restriction enzyme, for example BsoBI from *Bacillus steaothermophilus*. The invention also features a kit or solution for SDA comprising a DNA polymerase of the present invention in combination and a thermostable restriction enzyme.
The polymerases of the present invention are also useful in methods for generating and amplifying a nucleic acid fragment via a strand displacement amplification (SDA) mechanism. The method generally comprises:

a) specifically hybridizing a first primer 5' to a target nucleic acid sequence, the first primer containing a restriction enzyme recognition sequence 5' to a target binding region;
b) extending the 3' ends of the hybridized material using a DNA polymerase of the present invention, preferably one in which the exonuclease activity has been removed, in the presence of three dNTPs and one dNTP; S;
c) nicking at the hemiphosphorothioate recognition site with a restriction enzyme, preferably;
d) extending the 3' end at the nick using a DNA polymerase of the present invention, displacing the downstream complement of the target strand; and
e) repeating steps (c) and (d).

This SDA method proceeds at a linear amplification rate if one primer is used as above. However, if two primers are used which hybridize to each strand of a double-stranded DNA fragment, then the method proceeds exponentially (Walker, G.T., Little, M.C., Nadeau, J.G. and Shank D.D. (1992) Proc. Natl. Acad. Sci. USA 89:392-396).

The present invention also provides a method for determining the nucleotide base sequence of a DNA molecule. The method includes providing a DNA molecule, annealing with a primer molecule able to hybridize to the DNA molecule; and incubating the annealed molecules in a vessel containing at least one, and preferably four deoxynucleotide triphosphate, and a DNA polymerase of the present invention preferably one containing the phenylalanine to tyrosine mutation. Also provided is at least one DNA synthesis terminating agent which terminates DNA synthesis at a specific nucleotide base. The method further includes separating the DNA products of the incubating reaction according to size, whereby at least a part of the nucleotide base sequence of the DNA molecule can be determined.

In preferred embodiments, the sequencing is performed at a temperature between 40 and 75°C.

In other preferred embodiments, the DNA polymerase has less than 1000, 250, 100, 50, 10 or even 2 units of exonuclease activity per mg of polymerase (measured by standard procedure, see below) and is able to utilize primers having only 4, 6 or 10 bases; and the concentration of all four deoxynucleoside triphosphates at the start of the incubating step is sufficient to allow DNA synthesis to continue until terminated by the agent, e.g. a ddNTP.

Preferably, more than 2, 5, 10 or even 100 fold excess of a dNTP is provided to the corresponding ddNTP.

In a related aspect, the invention features a kit or solution for DNA sequencing including a DNA polymerase of the present invention and a reagent necessary for the sequencing such as dITP, deaza dGTP, a chain terminating agent such as a ddNTP, and optionally a pyrophosphatase.

The DNA polymerases of the present invention containing the phenylalanine to tyrosine mutation are suitably used in sequencing, preferably in combination with a pyrophosphatase. Accordingly, the present invention provides a composition which comprises a DNA polymerase of the present invention containing the phenylalanine to tyrosine mutation in combination with a pyrophosphatase, preferably a thermostable pyrophosphatase from Thermoplasma acidophilum.

In another related aspect, the invention features a method for sequencing a strand of DNA essentially as described above with one or more (preferably 2, 3 or 4) deoxyribonucleoside triphosphates, a DNA polymerase of the present invention, and a first chain terminating agent. The DNA polymerase causes the primer to be elongated to form a first series of first DNA products differing in the length of the elongated primer, each first DNA product having a chain terminating agent at its elongated end, and the number of molecules of each first DNA products being approximately the same for substantially all DNA products differing in length by no more than 20 bases. The method also features providing a second chain terminating agent in the hybridized mixture at a concentration different from the first chain terminating agent, wherein the DNA polymerase causes production of a second series of second DNA products differing in length of the elongated primer, with each second DNA product having the second chain terminating agent at its elongated end. The number of molecules of each second DNA product is approximately the same for substantially all second DNA products differing in length from each other by from 1 to 20 bases, and is distinctly different from the number of molecules of all the first DNA products having a length differing by no more than 20 bases from that of said second DNA products.

In preferred embodiments, three or four such chain terminating agents can be used to make different products and the sequence reaction is provided with a magnesium ion, or even a manganese or iron ion (e.g. at a concentration between 0.05 and 100 mM, preferably between 1 and 10mM); and the DNA products are separated according to molecular weight in four or less lanes of a gel.

In another related aspect, the invention features a method for sequencing a nucleic acid by combining an oligonucleotide primer, a nucleic acid to be sequenced, between one and four deoxyribonucleoside triphosphates, a DNA polymerase of the present invention, and at least two chain terminating agents in different amounts, under conditions favoring extension of the oligonucleotide primer to form nucleic acid fragments complementary to the nucleic acid to be sequenced. The method further includes separating the nucleic acid fragments by size and determining the
nucleic acid sequence. The agents are differentiated from each other by intensity of a label in the primer extension products.

[0043] In a further aspect, the present invention provides a method for preparing complementary DNA by combining an oligonucleotide primer, a sample of RNA, a DNA polymerase of the present invention, and between one and four deoxyribonucleoside phosphates, under conditions favoring preparation of the complementary DNA.

[0044] The DNA polymerases of the present invention which act as reverse transcriptases lack appreciable, and preferably have no RNaseH activity and, as such, are useful in RT/PCR, the generation of hybridization probes and RNA sequencing. In a yet further aspect, the present invention provides a purified thermostable reverse transcriptase having a reverse transcriptase activity of greater than 1000 units per milligram. Preferably, the reverse transcriptase lacks RNaseH activity; the reverse transcriptase is from *Thermoanaerobacter thermohydrodsulfuricus*; the reverse transcriptase has an N-terminal deletion or amino acid changes that remove the exonuclease function. In a further aspect, the invention features a method for reverse transcription/polymerase chain reaction (RT/PCR) which utilizes a DNA polymerase of the present invention and a DNA polymerase suitable for PCR in the same reaction vessel. Preferably, the DNA polymerase of the present invention is from *Thermoanaerobacter thermohydrodsulfuricus*, the polymerase has one or more amino acids deleted from the amino terminus or amino acid changes to remove the exonuclease activity, the DNA polymerase suitable for PCR is Taq DNA polymerase. In another aspect, the present invention features a kit or solution for RT/PCR comprising a DNA polymerase of the present invention and a DNA polymerase suitable for PCR. Preferably, the DNA polymerase of the present invention is from *Thermoanaerobacter thermohydrodsulfuricus*, the polymerase has one or more amino acids deleted from the amino terminus or amino acid changes to remove the exonuclease function, the DNA polymerase suitable for PCR is Taq DNA polymerase.

[0045] Other features and advantages of the invention will be apparent from the following description of the preferred embodiments thereof, and from the claims.

Description of the Preferred Embodiments

[0046] The drawings will first be briefly described.

Drawings

[0047] Fig. 1A-E is the DNA sequence for 5300 bp of genomic DNA flanking and encoding the full length DNA polymerase from *Thermoanaerobacter thermohydrodsulfuricus* (SEQ. ID. NO. 1).

[0048] Fig. 2A-D is a contiguous open reading frame capable of encoding the full length polymerase from *Thermoanaerobacter thermohydrodsulfuricus* (SEQ. ID. NO. 2). Translation is of an open reading frame spanning nucleotides 1056-3674 of genomic DNA in Fig. 1A-E, potentially encoding native polymerase.

[0049] Fig. 3 is the contiguous open reading frame of nucleotides 796-2616, amino acids 266-872 plus an initiator methionine at the amino terminus, the exo mutant of the enzyme containing 608 amino acids (SEQ. ID. NO. 3).

[0050] Fig. 4A-C is the contiguous open reading frame of the F412Y mutation of the exo-mutant polymerase, containing 578 amino acids with. (SEQ. ID. NO. 4). At amino acid 412 there is a tyrosine replacing phenylalanine.

[0051] Fig. 5 is a schematic of cloned polymerases. Letters A-D represent start sites of the various constructs shown in the table in Example 2.

[0052] Fig. 6 is a graphical representation comparing relative RT activity of the exo-mutant enzyme and other enzymes that have reverse transcription activity. The x-axis indicates the polymerase (MMLV, AMV, Tth, USB). The y-axis represents reverse transcriptase activity in cpm.

[0053] Fig. 7A-C is the DNA sequence encoding a full length native polymerase from *Thermoanaerobacter thermohydrodsulfuricus* (SEQ. ID. NO. 23).

Examples

[0054] The following examples serve to illustrate the DNA polymerases of the present invention.

Example 1: General Methods Used During Purification and Characterization

Assay of DNA Polymerase Activity

[0055] DNA polymerase activity was assayed by measuring the amount of incorporation of a radiolabelled deoxy-nucleotide into acid precipitable material using activated salmon sperm DNA as a template (Richardson, C. C. (1966) DNA polymerase from *Escherichia coli*, pp. 263-276 in G. L. Cantoni and D. R. Davies (ed.), *Procedures in nucleic acid research*. Harper and Row, New York). A unit of DNA polymerase is the amount of enzyme that catalyzes incor-
poration of 10 nmoles of deoxynucleotide triphosphate into acid-precipitable material in 30 minutes at 70°C. An assay consists of 2-10 μl of protein solution containing DNA polymerase being incubated with 50 μl of assay mix (20 mM Tris HCl pH 8.5 @ room temperature, 200 μM each deoxynucleotide triphosphate, 10 mM KCl, 5 mM MgCl₂, 0.2 mg/ml activated salmon sperm DNA, 1 μCi 3000 Ci/mmole [α-33P]dATP) at 70°C for 10 minutes. The reaction is stopped by adding the contents of the assay to a tube containing 1 ml each of carrier (50 mg/ml fish sperm DNA, 2 mM EDTA) and precipitant (20% (w/v) trichloroacetic acid, 2% (w/v) sodium pyrophosphate). After incubation on ice for at least 5 minutes, the solution is filtered through a glass fiber filter (e.g. GF/B, Whatman), washed with several milliliters ice-cold wash solution (1 N hydrochloric acid, 100 mM sodium pyrophosphate), placed in a counting vial with aqueous liquid scintillation cocktail (e.g. Biosafe II, Research Products International Corp.) and counted in a liquid scintillation counter. DNA polymerase activity of the test solution is calculated from the measured specific activity of the assay mix. Only activity values measured to be within the linear range of 0.01 to 0.2 units per assay reaction were accepted as significant.

Measurement of Protein Concentration

Solution protein concentrations are measured spectrophotometrically by determining the absorbance of the test solution at a wavelength of 205 nm (A₂₀₅) (Scopes, R. K. (1994) pp. 46-48 Protein Purification. Springer-Verlag, New York). The extinction coefficient of polypeptides is taken as ε₂₀₅(1mg/ml) = 31.

Assay of Exonuclease Activity

Exonuclease activity was measured as described (Kong, H., Kucera, R. B. and Jack, W. E. (1993) J. Biol. Chem. 268, 1965-1975) using, as substrate, HindIII digested phage λ DNA labeled with 3H by the action of CpG methylase or a 500 bp PCR product generated from pBR322 labeled with 3H by incorporation of tritiated TTP. Exonuclease assays were performed in the same buffer (20 mM Tris•HCl pH 8.5 @ room temperature, 10 mM KCl, 5 mM MgCl₂) and at the same temperature (70°C) as polymerase assays.

Assay for Directionality of Exonuclease

The directionality of the associated exonuclease activity was determined by monitoring the time-dependent release of radioactivity from a DNA substrate differentially labeled on the 5’ and 3’ ends as described (Kong et al., supra).

Assay for Strand Displacement Activity

An extension assay has been developed to evaluate polymerases for their utility in the SDA (strand displacement amplification) reaction (Becton-Dickinson). This extension reaction tests the ability of the polymerase to displace a downstream strand and its ability to initiate extension at a nick site. This displacement and initiation at a nick is staged by annealing two primers labeled at their 5’ ends immediately adjacent: to each other. If the polymerase has strand displacement activity and is able to initiate at a nick, both primers are extended and two extension products can be visualized. If the polymerase cannot initiate at a nick (that is, it has a preference for initiation at a gap) or it lacks displacement activity then only the downstream primer is extended and only one product is detected.

Using the plasmid pBR322 as the target, two primers pBR1 and pBR2 were synthesized that anneal immediately adjacent to each other on the plasmid. Primer pBR1 is the upstream primer and corresponds to bases 3661 to 3690 in pBR322 (GGTTCCCAACGATCAAGGCGAGTTACATGA)(SEQ. ID. NO. 5). Primer pBR2 anneals downstream and corresponds to pBR322 3691-3720 (TCCCCCATGGTGCAAAAAAGCGGTTAGC)(SEQ. ID. NO. 6). These 30 mer oligos were purified by electrophoresis from a denaturing polyacrylamide gel. The primers are labeled in a kinase reaction using 10 mM primer, 50 mM Tris•HCl pH 7.5, 10 mM MgCl₂, 70 μCi [γ-32P] ATP and 10 units T4 polynucleotide kinase for 30 minutes at 37°C. A final concentration of 0.2 μM (each) of these two labeled primers are combined with 200 ng of PstI / HindIII digested pBR322. These are denatured at 98°C for 3 minutes in a buffer containing 25 mM potassium phosphate pH 7.4, 2-8 mM MgCl₂, 0.2 to 1 mM dNTPs and cooled to the reaction temperature (50-70°C) for 2 minutes. One unit of the polymerase is added and the reaction is continued for 10 minutes. The reactions are stopped by addition of an equal volume of 95% formamide, 50 mM EDTA, bromophenol blue dye. Twenty-five microliters are electrophoresed over a 6% denaturing gel and the gel is exposed to autoradiographic film.
Example 2: Purification of Native Enzyme

Crude Purification

Growth of *Thermoanaerobacter thermohydrosulfuricus* Cells

[0061] A lyophilized culture of *Thermoanaerobacter thermohydrosulfuricus* strain JW102 was purchased from the Deutsche Sammlung von Mikroorganismen und Zellkulturen, GmbH (DSM). The growth medium used was *Clostridium Thermohydrosulfuricum* Medium (medium #61 in the DSM catalog of strains, 1993 edition) which consists of (per liter):

- 10 g tryptone
- 10 g sucrose
- 2 g yeast extract
- 0.2 g FeSO$_4$7H$_2$O
- 0.2 g Na$_2$SO$_3$
- 0.08 g Na$_2$S$_2$O$_3$·5H$_2$O
- 1 mg Resazurin

The culture was grown at 70°C under anaerobic conditions. After growth, the media was cooled to approximately room temperature and the culture harvested by continuous flow centrifugation. Cell paste was stored at -80°C prior to subsequent procedures.

Preparation of Extract

[0062] Frozen cell paste (5 grams) was combined with 100 ml of buffer A (50 mM Tris·HCl pH 8.0, 10 mM NaCl, 1 mM DTT, 1 mM EDTA, 10% (w/v) glycerol). The paste was stirred at room temperature until resuspension was complete. The resulting cell suspension was cooled to 0°C by incubation in an ice bath and subjected to approximately 5 minutes of sonication to lyse the cells. Cell debris was removed by centrifugation at 70,000 xg for 20 minutes at 4°C. The supernatant cleared lysate was designated fraction 1.

Liquid Chromatography Purification

[0063] Fraction 1 was loaded onto a low-pressure liquid chromatography column (70 ml bed volume, 26 mm Ø) of DEAE cellulose (DE-52, Whatman) equilibrated with buffer A. The column was washed with several column volumes of buffer A and the DNA polymerase eluted with a linear salt gradient from 10 mM to 1000 mM NaCl in buffer A at a flow rate of 0.9 ml/min. This and all subsequent chromatography steps were carried out at ambient room temperature (approximately 25°C). The peak of activity eluted at approximately 32 mM NaCl. Active fractions were pooled and designated fraction 2.

[0064] Fraction 2 was loaded directly onto a column (22 ml bed volume, 16 mm Ø) of heparin sepharose (Heparin Sepharose CL-6B, Pharmacia) equilibrated with buffer A. The column was washed with several column volumes of buffer A and the DNA polymerase activity eluted with a linear salt gradient from 10 mM to 700 mM NaCl in buffer A at a flow rate of 0.4 ml/min. The peak of activity eluted at approximately 400 mM NaCl. Active fractions were pooled and designated fraction 3.

[0065] Fraction 3 was loaded directly onto a column (2 ml bed volume, 16 mm Ø) of hydroxylapatite (BioGel HA, Bio-Rad) equilibrated with buffer B (50 mM potassium phosphate pH 7.4, 1 mM DTT, 10% (w/v) glycerol). The column was washed with several column volumes of buffer B and the DNA polymerase eluted with a linear gradient from 50 mM to 700 mM potassium phosphate in buffer A at a flow rate of 0.4 ml/min. The DNA polymerase containing fractions were pooled and concentrated by ultrafiltration (Centricon-50, Amicon) to a volume of 100 ml to which 100 ml of 100% glycerol was added to obtain the final product. This final preparation was determined to contain about 1000 units of DNA polymerase which was judged to be approximately 80% pure on a coomasie stained denaturing polyacrylamide gel (SDS-PAGE).

Optimized Purification

Preparation of Extract

[0066] Frozen cell paste (45 grams) was resuspended in 250 ml of lysis buffer (50 mM Tris·HCl pH 8, 10 mM NaCl, 5 mM EDTA, 0.2% (v/v) NP40, 1 mM DTT). Lysozyme was added to 200 mg/ml final concentration and the solution was stirred at room temperature for 60 min. DNase I was added to 10 mg/ml final concentration and incubation was continued for 10 minutes. Insoluble material was removed by ultracentrifugation at 70,000 xg for 40 min. The decanted supernatant is designated fraction I.

Polyethyleneimine Precipitation

[0067] A pilot polyethyleneimine (PEI) precipitation experiment was done to determine the final PEI concentration for this lysate to precipitate all polymerase activity with the least coprecipitate protein content. PEI was added from a
10% (v/v) stock to Fraction I to 0.4% final concentration. The solution was stirred for 20 minutes at 4°C. The PEI precipitate was collected by centrifugation at 8,000 xg for 20 minutes. The polymerase activity was extracted from the pellet with 150 ml of 300 mM NaCl in buffer A (50 mM Tris-HCl pH 8, 1 mM DTT, 1 mM EDTA and 10% glycerol). This solution was centrifuged again at 17,000 xg. Ammonium sulfate was added to the supernatant to 80% saturation and the solution was stirred for one hour at 4°C to precipitate proteins from residual PEI. The ammonium sulfate fraction was centrifuged at 48,000 xg for 30 min and the pellet resuspended in 400 ml 10 mM NaCl, 50 mM Tris-HCl pH 8. This is designated Fraction II.

Liquid Chromatography Purification

A column (150 ml bed volume) of Heparin-Sepharose CL-6B (Pharmacia) was equilibrated with 150 mM NaCl in buffer A. Fraction II was loaded at 150 ml/h and washed with two bed volumes of buffer A. The polymerase activity was then eluted with a linear gradient of 150 mM to 700 mM NaCl in buffer A yielding a polymerase peak at about 400 mM NaCl. The peak fractions were pooled and diluted fivefold with 100 mM NaCl in buffer B (50 mM Tris-HCl pH 7.5, 0.1 mM DTT, 1 mM EDTA and 5% glycerol). This pool is designated Fraction III.

A column (4 ml bed volume) of ssDNA agarose (Pharmacia) was equilibrated with 100 mM NaCl in buffer B. One-third of the Fraction III was loaded at a flow rate of 12 ml/hr. The column was washed with the same buffer and polymerase activity eluted with a linear gradient of 100 mM to 1000 mM NaCl in buffer B. The polymerase activity peak eluted at approximately 375 mM NaCl. The peak fractions were pooled and concentrated with Centricon-50 (Amicon) and stored by adding an equal volume of 100% glycerol.

Preparation of exo-Enzyme by Subtilisin Cleavage

Approximately 850 units of purified native Thermoanaerobacter thermohydrosulfuricus DNA polymerase I was exchanged into 150 mM potassium phosphate pH 6.5 by three rounds of dilution in this buffer and concentration by ultrafiltration yielding a solution containing about 250 mg of protein in 550 µl. Subtilisin (10 µl of 0.06 mg/ml solution) was added to the polymerase and the mixture was incubated at 37°C for 1 hour. After incubation, the reaction was diluted to 10 ml final volume with buffer A and loaded onto a column of heparin sepharose (4 ml bed volume, 16 mm Ø) equilibrated in buffer A containing 150 mM NaCl. The column was washed with several column volumes of this buffer and the polymerase activity eluted with a linear gradient of 150 mM to 700 mM NaCl in buffer A. Active fractions were pooled.

Example 3: Characterization of Native Enzyme

General properties

The native enzyme migrates as an approximately 90,000 Da polypeptide (just below phosphorylase B) on SDS-PAGE. The enzyme possesses strand displacement activity. It possesses a 5'-3' exonuclease activity at approximately 1/80 the level of polymerase activity (unit:unit). The temperature optimum of DNA synthesis is near 75°C under assay conditions. The optimum Mg²⁺ concentration of DNA synthesis is 1 mM. The optimum Mn²⁺ concentration of DNA synthesis is 0.5 mM.

Assay for thermal stability

To assess the thermal stability of the enzyme, a sample of native Thermoanaerobacter thermohydrosulfuricus DNA polymerase was incubated at elevated temperatures in several buffers. The buffers all consisted of 25 mM Tris-HCl pH 8.5, 50 mM KCl, 5 mM MgCl₂, 1 mM 2-mercaptoethanol, 0.5% (w/v) Nonidet P-40 and 100 mg/ml BSA with the following additions:

- BSA buffer : no additions.
- GLY buffer : add glycerol to 40% (v/v)
- MMO buffer : add N-methylmorpholine-N-oxide (MMO) to 3 M
- TMANO buffer : add Trimethylamine-N-oxide (TMANO) to 4 M

Native enzyme (0.02 units/µl) was incubated in these buffers at various temperatures. At 5, 10 and 15 minutes, 5 µl aliquots were withdrawn and placed into a fresh tube on ice. When all samples had been acquired, the heated aliquots were assayed in the usual manner. Results are expressed as the fraction activity remaining at a given time point at a given temperature.
The enzyme was least stable in BSA buffer, retaining half the initial measured polymerase activity after about 5 minutes at 80°C. In GLY buffer, the enzyme retained half the initial activity after about 5 minutes at 85°C. In MMO and TMANO buffers, incubation at 85 to 90°C was required to eliminate half the initial measured activity in 5 minutes. Incubation in any of the these buffers at 75°C for 15 minutes resulted in not more than 20% loss of initial DNA polymerase activity.

Example 4: Characterization of Subtilisin-Cleaved Enzyme

General properties

The cleaved enzyme migrates as a doublet consisting of polypeptides migrating at approximately 64 kDa and 55 kDa on SDS-PAGE. It is free of detectable exonuclease activity. The enzyme possesses strand displacement activity. The enzyme has similar thermal stability to the native enzyme.

Example 5: Preparation of Cloned 64 kDa Enzyme

Preparation of genomic DNA

Frozen cell paste of *Thermoanaerobacter thermohydrosulfuricus* (approximately 350 mg) was resuspended in 700 μl of lysis buffer (50 mM Tris•HCl pH 8.0, 50 mM EDTA, 3% SDS (w/v), 1% 2-mercaptoethanol (v/v)) and incubated for 1 hour at 65°C. This solution was extracted three times with 25:24:1 phenol:chloroform:isoamyl alcohol, twice with chloroform and precipitated with two volumes of 100% ethanol. The pellet was dried briefly in vacuo and dissolved in 700 μl TE (10 mM Tris•HCl pH 8.0, 1 mM EDTA). The concentration of genomic DNA was determined spectrophotometrically by measuring the absorbance at 260 nm (1A_{260}=50 μg/ml) and by comparison of UV fluorescence of bands on an ethidium bromide stained agarose gel relative to standard concentration markers.

Preparation of DNA polymerase I specific probe

Genomic DNA was used as a substrate for a PCR reaction using degenerate oligonucleotides designed against conserved regions of DNA polymerase I family polymerases (Uemori, T., Ishino, Y., Fujita, K., Asada, K. and Kato, I. (1993) *J. Biochem.* 113, 401-410). A standard Taq PCR reaction (20 μl) contained about 1 ng genomic DNA and 1μM of each primer EPOLF (GACCC(ATC)AAC(CT)T(CG)CA(AG)A A(CT)AT(ATC)CC (SEQ.ID.NO. 7) and EPOLR ((GT)A(CG) (CG)A(GT) (CT)TC (AG) TCGTG(GATC)AC(CT)TG) (SEQ. ID. NO. 8). After one round of PCR, the product was resolved on an agarose gel. The product around the expected 600 bp region was excised from the gel, crushed, soaked in TE and used as substrate for a second round of PCR using the same conditions. The resulting ca. 600 bp product was directly sequenced using the EPOLF and EPOLR primers (Sequenase PCR Product Sequencing Kit, Amersham). One open reading frame was found to be homologous to the targeted region of DNA polymerases I of known sequence as judged by BLAST analysis.

Preparation of subgenomic libraries and screening

This PCR product was labeled using a radioactive or non-radioactive hybridization kit (Gene Images, USB) and used to probe blots of agarose gels of genomic DNA digested with various restriction enzymes. Digestion of genomic DNA with HindIII produced a unique hybridizing fragment of approximately 1.7 kb. DNA fragments of approximately 1.7 kb were excised from an agarose gel and cloned into pBluescript II KS (+) (Stratagene). Individual colonies that hybridized to the probe were purified and sequenced. Similarly, double digestion of *Thermoanaerobacter thermohydrosulfuricus* genomic DNA with EcoRI and XhoI produced a unique hybridizing fragment of approximately 1.9 kb and digestion with EcoRI produced a unique hybridizing fragment of approximately 3 kb. The sequences of these clones were merged and formed a coherent open reading frame which had high homology to the C-terminal portion DNA polymerases I of known sequence as judged by BLAST analysis. The clone containing the 1.7 kb HindIII fragment was digested with EcoRI and XhoI to produce a 1.4 kb fragment which was gel purified, ^32P labeled by nick-translation and used to probe a blot of *Thermoanaerobacter thermohydrosulfuricus* genomic DNA digested with *PstI* and *XhoI*. This identified a unique hybridizing fragment of 3 kb that contained the remaining N-terminal portion of the gene. The nucleotide sequence is presented in Fig. 1A-E (SEQ. ID. NO. 1) and a contiguous open reading frame potentially encoding the native polymerase is presented in Fig. 2A-D (SEQ. ID. NO. 2).
Preparation of expression vectors

Several PCR products were made from the genomic DNA that allowed cloning portions of polymerase coding sequence not including 5'-3' exonuclease domain (See Fig. 5). Letters A-D represent start sites of the various constructs outlined in the table below. These truncated polymerase coding sequences were placed under the control of the $$\lambda$$ PL promoter of the expression vector pRE2 (Reddy, P., Peterkofsky, A. and McKenney, K. (1989) Nucleic Acids Res. 17, 10473-10488). These vectors were propagated in E. coli strain DH-1 $$\lambda$$+ ($$\lambda$$ cI$$^+$$).

To minimize errors introduced by Taq DNA polymerase in PCR, a modified long PCR was carried out. The PCR reaction contained: 20 mM Tricine (pH 8.8), 85 mM KOAc, 200 mM each dNTP, 5% DMSO, 0.5 mM each primer, 1.5 mM MgOAc (added as Hot-Start). 2.5 units Hot Tub (or rTth) DNA polymerase per 100 ml reaction (Amersham), 0.025 U Deep Vent DNA (New England Biolabs) polymerase per 100 ml reaction, 20-100 ng Thermoanaerobacter thermohydrosulfuricus genomic DNA per 100 ml reaction. The cycling parameters consisted of: 94°C 30s, then 68°C 10m40s x 8 cycles; 94°C 30s, then 68°C 12m00s x 8 cycles; 94°C 30s, then 68°C 13m20s x 8 cycles; 94°C 30s, then 68°C 14m40s x 8 cycles.

All PCR products were generated with the reverse primer DDREV(GGGGTACCC TAGATCAGTTGGCCAAAAACCA) (SEQ. ID. NO. 9) at the 3' end and various primers utilized at the 5' end:

**DDMET** (GGAATTCATATGCGCTTTCTGAAACCCGCTGACTCAAC) (SEQ. ID. NO. 10);

**DD55kd** (GGAATTCATATGGCTTTTTAAAGGAATAGAG) (SEQ. ID. NO. 11);

**DD64kd** (GGAATTCATATGAAAGTTGAAAGATGGCAA) (SEQ. ID. NO. 12);

**DD64pref** (GGAATTCATATGAAAGTTGAAAGATGGCTAAAAGTAGACATCAAAGAT) (TAG) (SEQ. ID. NO. 13);

**DDHF** (GGAATTCATATGAAGCTTTAGAGCTTTT) (SEQ. ID. NO. 14);

**DD67kd** (GGAATTCATATGAAACCTTTTAGAGCTTTTCAAAGATTAGAATTC) (SEQ. ID. NO. 15).

The resultant PCR products were digested with Ndel and Kpnl and cloned into pRE2 digested with the same restriction enzymes.

Silent changes (see Table below) denotes that the 5' primer encodes some codons which are synonymous to the native sequences, but are more frequently utilized by E. coli. By this method, six different clones were constructed which encode four different polypeptides.

<table>
<thead>
<tr>
<th>Site</th>
<th>Plasmid</th>
<th>MW</th>
<th>codons</th>
<th>5' Primer</th>
<th>Note</th>
</tr>
</thead>
<tbody>
<tr>
<td>D</td>
<td>pMET11</td>
<td>55 kDa</td>
<td>483</td>
<td>DDMET</td>
<td>Silent changes, native MET</td>
</tr>
<tr>
<td>C</td>
<td>pRED5504</td>
<td>58 kDa</td>
<td>502</td>
<td>DD55kd</td>
<td>Native sequence + MET</td>
</tr>
<tr>
<td>B</td>
<td>pRED6404</td>
<td>64 kDa</td>
<td>579</td>
<td>DD64kd</td>
<td>Native sequence + MET</td>
</tr>
<tr>
<td>B</td>
<td>pPREF2</td>
<td>64 kDa</td>
<td>579</td>
<td>DD64pref</td>
<td>Silent changes + MET</td>
</tr>
<tr>
<td>A</td>
<td>pRED4</td>
<td>67 kDa</td>
<td>609</td>
<td>DDHF</td>
<td>Native sequence + MET</td>
</tr>
<tr>
<td>A</td>
<td>pRED6706</td>
<td>67 kDa</td>
<td>609</td>
<td>DD67kd</td>
<td>Silent changes + MET</td>
</tr>
</tbody>
</table>
Expression of Cloned 64 kDa Enzyme

[0084] E. coli strain MZ-1 (cI857 lysogen) harboring the plasmid pRED6404 was cultured in LB medium (per liter: 10 g tryptone, 5 g yeast extract, 10 g NaCl) supplemented with 50 μg/ml ampicillin. The culture was grown at 30-32°C until the OD_{600} reached 1.5 at which time the culture temperature was raised to 40-42°C. Growth continued at this temperature for 2 hours before the culture was harvested in a continuous flow centrifuge. Cell paste was stored at -80°C prior to subsequent procedures.

Example 6: Preparation of Cloned Full-length Enzyme

[0085] Full length enzyme was cloned by PCR from the clone of genomic DNA (SEQ. ID. NO. 1) (Fig. 1A-E). A 5’ primer DDF-FOR (GGAATTCCATATGGCTTATAAATTTTAATCGTATGATGAGTGACG) (SEQ. ID.NO. 24) was synthesized. This primer encodes an NdeI site, inserts an alanine codon after the initiating methionine and changes two codons to synonymous codons favored by E. coli (ATA -> ATC, Ile at aa position 6 and GGA -> GGT, Gly at aa position 9). This primer was used in conjunction with a vector primer from the assembled clone of genomic DNA to produce a PCR product which was subsequently digested with NdeI and XhoI and subcloned into similarly digested pRED6404 to form pLS-3. The DNA for the full length enzyme is shown in Fig. 7 A-C (SEQ. ID. NO. 23).

Example 7: Purification of Cloned 64 kDa Enzyme

Preparation of Extract

[0086] Frozen cell paste (370 grams) was combined with 1800 ml of buffer A (50 mM Tris•HCl pH 8.0, 10 mM NaCl, 1 mM DTT, 1 mM EDTA, 10% (w/v) glycerol). The paste was stirred at room temperature until resuspension was complete. The cells were lysed using a French pressure cell. Cell debris was removed by centrifugation at 70,000 xg for 20 minutes at 4°C.

Liquid chromatography purification

[0087] The cleared lysate was passed through a column of DEAE cellulose at 300 mM salt. The collected flow through was heated to 70°C for 10 minutes, cleared again by centrifugation and loaded onto a column heparin sepharose. The DNA polymerase was eluted from this column with a salt gradient as before and was found in most cases to be pure. If not, column chromatography on hydroxylapatite was performed.

Example 8: Characterization of Cloned 64 kDa Enzyme

General properties

[0088] Cloned enzyme refers to the purified gene product of pRED6404. The cloned enzyme migrates as an approximately 64,000 Da polypeptide (just below bovine serum albumin, MW 67,000) on SDS-PAGE. The specific activity of the cloned enzyme is 28,000 units/mg. It is free of detectable exonuclease activity. The temperature optimum of DNA synthesis is near 72°C under assay conditions. The optimum Mg^{2+} concentration of DNA synthesis is 2 mM. The optimum Mn^{2+} concentration of DNA synthesis is 0.7 mM.

Reverse transcriptase activity

[0089] The cloned enzyme functioned as a reverse transcriptase in an assay containing poly(rA):oligo(dT_{12-18}) and [3H]TTP at 42°C. This enzyme has RT activity in the presence of Mg^{2+} with or without the addition of Mn^{2+}.

Measurement of reverse transcriptase activities of different enzymes using an SPA format

[0090] Amersham reverse transcriptase scintillation proximity assay (SPA) kit (NK9020), and MMLV (E 70456), and AMV (E 70041) enzymes were supplied by Amersham International. The cloned 64 kD enzyme prepared above was designated as USB enzyme. Tth (N808-0098) enzyme was obtained from Perkin-Elmer.

[0091] MMLV and AMV enzymes were tested at 37°C and 42°C respectively in the presence of magnesium. The Tth and USB enzymes were both used at 60°C with manganese present as the co-factor. The assay buffer used was the same as supplied in the SPA kit except the appropriate co-factor was substituted in the buffer mix. Essentially, the assay was set up as follows in duplicate using the protocol provided in the SPA kit in sarstedt tubes in a total reaction
volume of 100 µl.

[0092] 80 µl of 0.77 µM [3H]TTP in a buffer consisting of:

- 50 mM Tris•HCl pH 8
- 80 mM KCl
- 10 mM MgCl₂ or 4 mM MnCl₂
- 10 mM DTT
- 0.05% w/v NP40
- 10 µl of 0.2 µM primer/template (biotinylated (dT₁₆) oligonucleotide annealed to poly r(A)) in

- 10 mM Tris•HCl pH 8
- 1 mM MgCl₂

[0093] 10 µl of 1 unit (as defined by the manufacturer) of enzyme diluted in 50 mM Tris•HCl pH 8, 10 mM DTT was added to each tube to start the reaction and incubated at the indicated temperature for 30 minutes after which the reaction was terminated with the addition of 10 µl 0.56 M EDTA pH 8. 10 µl of the supplied streptavidin SPA beads were then added to each tube, and after incubating for a further 10 minutes at room temperature, were counted in a LKB 1205 scintillation counter.

[0094] Appropriate controls without enzyme were also performed. The data obtained is shown in Fig. 6. This experiment does not take into account the melting temperature of the primer used in this experiment which at 60°C would theoretically leave only a minority of the primer still being annealed to the template. This may actually result in an underestimate of the activity observed with the USB and Tth enzymes.

Use of 64 kDa polymerase in Strand Displacement Amplification

[0095] The exonuclease deficient cloned enzyme described above was used in a Strand Displacement Amplification (SDA) reaction (Walker et al., (1992) Proc. Natl. Acad. Sci. USA 89:392-396). The target DNA to be amplified is a 1.2 kb clone of the insertion element IS 6110 from Mycobacterium tuberculosis [GenBank accession no. M29899]. The target DNA is cloned into pBluescript KS (Stratagene, La Jolla, CA). The reaction mixture (final concentrations: 25mM potassium phosphate pH 7.4, 0.1 mg/ml BSA, 0.05 mM each primers B1 (CGATCGAGCAAGCCA) (SEQ. ID. NO. 16) and B2 (CGAGCCGCTCGCTGA) (SEQ. ID. NO. 17), 1.4 mM dCTP aS, 0.5 mM each dATP, dGTP and dCTP and 10⁷ copies of target plasmid) was heated to 100°C for 3 minutes to denature the target and then placed at 55°C. After about 5 minutes at 55°C, enzyme mix was added (final concentrations: 35 Units 64 kD polymerase, 160 Units BsoBI restriction enzyme (New England Biolabs, Beverly, MA) and 6.9 mM MgCl₂ (total reaction volume 50 µl) and incubation was continued for 10 minutes. An equal volume of stop solution (95% formamide, 20 mM EDTA and 0.05% each bromophenol blue and xylene cyanol FF) was added and 5 µl of this was electrophoresed on a 10% denaturing polyacrylamide gel. DNA was visualized by staining with ethidium bromide. The 64 kDa polymerase produced the expected 105 nucleotide full length product (major product) and the expected 86 nucleotide nicked product (minor product).

Use of 64 kDa polymerase in RT-PCR

[0096] The 64 kDa polymerase was used as a reverse transcriptase in conjunction with AmpliTaq in RT-PCR. Using rabbit globin mRNA as a template, 10⁶ copies were detectable under the following protocol. In a reaction volume of 100 µl, the following components were assembled: 1X PCR buffer (10 mM Tris•HCl pH 8.3, 50 mM KCl), 2 mM MgCl₂, 200 µM each dNTP, 100 U human placental RNAse inhibitor (Amersham), 10⁶ copies rabbit globin mRNA (BRL), 30 pmoles globin-down primer (TCAGTGGTATTTGTGAGCCAGGGCATTGGCCACACCACCTTCT) (SEQ. ID. NO. 25), 15 pmole globin-up primer (ATCCCCAAAACAGAGAATGCTGGCATCTGTC) (SEQ. ID. NO. 26), 2.5 U AmpliTaq polymerase and 5 U 64 kD polymerase. The reaction was incubated at 65°C for 30 minutes followed by 5 minutes at 95°C, 35 cycles of [95°C for 1 minute, 65°C for 1 minute, 70°C for 2 minutes] followed by 72°C for 5 minutes. The product was resolved on a 2% ethidium bromide stained agarose gel and the expected 463 bp product was observed.

[0097] Using total RNA from rat liver, an RT-PCR product from the albumin gene was detected from 250 ng of template using the following protocol. In a reaction volume of 100 µl, the following components were assembled: 1X PCR buffer (10 mM Tris•HCl pH 8.3, 50 mM KCl), 2 mM MgCl₂, 200 µM each dNTP, 100 U human placental RNAse inhibitor (Amersham), 250 ng rat liver RNA, 30 pmoles albumin-down primer (GTCTCTCTGGAGAGGGATTGACCATGACCT) (SEQ. ID. NO. 27), 15 pmole albumin-up primer (CACACAGAGTGAGATCGCCATCTGGTTAAAGGA) (SEQ. ID. NO.
28), 2.5 U AmpliTaq polymerase and 40 U 64 kD polymerase. The reaction was incubated at 65°C for 30 minutes followed by 5 minutes at 95°C, 35 cycles of [95°C for 1 minute, 65°C for 1 minute, 70°C for 2 minutes] followed by 72°C for 5 minutes. The product was resolved on a 2% ethidium bromide stained agarose gel and the expected 386 bp product was observed.

Example 9: Construction of FY DNA polymerase from *Thermoanaerobacter thermohydrosulfuricus*

Starting Material

- DNA constructs that were used to generate the FY polymerase were as follows:
  - pRED6706 is an expression vector which encodes a polypeptide of 608 amino acids. After an engineered initiating methionine, the remaining 607 amino acids encoded represent the C-terminal portion of the DNA polymerase I obtainable from *Thermoanaerobacter thermohydrosulfuricus*.
  - pRED6404 is an expression vector which encodes a polypeptide of 578 amino acids. After an engineered initiating methionine, the remaining 577 amino acids encoded represent the C-terminal portion of the DNA polymerase I obtainable from *Thermoanaerobacter thermohydrosulfuricus*.
  - pPREF2 encodes the same polypeptide as pRED6404, but has three silent mutations of the *Thermoanaerobacter thermohydrosulfuricus* sequence at codons 7, 10 and 11 to eliminate codons rarely utilized by *E. coli*.

- All of these constructs are derived from expression vector pRE2 (*Nucleic Acids Res.* (1989) 17, 10473-88) which utilizes the $\lambda$ P$_L$ promoter and $\lambda$ cII ribosome binding site.

Construction of F412Y

**Thermoanaerobacter thermohydrosulfuricus** DNA polymerase was searched for the sequence (RXXXKXXXFXXXYG) (SEQ. ID. NO. 20) to locate the motif B or O-helix region. The sequence (RRAAKAVNFGIIYG) (SEQ. ID. NO. 21) was found at amino acid residues numbered 433 through 466 of the predicted pRED6706 gene product. A mutagenic oligonucleotide, DDFY (AAGGTCAGAAAGCCCATAATCGCTTATGCCATATATTATG CCATAATTTACGGC) (SEQ. ID. NO. 22) was designed anti parallel to nucleotides 1315-1371 of the pRED6706 coding sequence except for the change (AAA-ATA) at nt 1325. This change corresponds to phenylalanine being changed to tyrosine in the expressed polymerase. This oligonucleotide also overlaps a unique XhoI restriction site at nucleotide 1361.

Oligonucleotide DDFY together with DD64pref (SEQ. ID. NO. 13) which is complementary to nucleotides 94-133 of the pRED6706 coding sequence, contains an NdeI restriction site, were used to direct PCR amplification and simultaneous mutagenesis of pRED6706. The resulting PCR product was digested with NdeI and XhoI, purified by agarose gel electrophoresis and ligated into the large fragment of similarly treated pRED6404. The ligated construct was used to transform competent *E. coli* strain DH1 $\lambda$ (+) (lcI $\lambda$ +). Clones yielding an appropriate restriction enzyme digestion pattern were sequenced to confirm the mutation. Spurious mutations were found in addition to the desired mutation in all clones examined, so an AflIII/KpnI restriction fragment containing exclusively the F412Y mutation (pRED6404 numbering) was subcloned in pRED6404 to yield pLS-1.

Example 10: Purification of F412Y DNA polymerase

Plasmid pLS-1 was transformed into *E. coli* strain MZ-1 (lcI857 lysogen). One liter of culture in LB medium containing 100 μg/ml Ampicillin was grown at 30°C until OD$_{600}$ = 1.1. The temperature was raised to 42°C and growth continued for 2 hours. The cells were harvested by centrifugation (4 g wet weight) and resuspended in 15 ml buffer A (50 mM Tris-HCl pH 8.0, 1 mM DTT, 1 mM PMSF, 1 mM EDTA, 10% (v/v) glycerol) containing 50 mM NaCl. The resuspension was sonicated four times for 30 seconds and the lysate was cleared by centrifugation. The cleared lysate was made 300 mM in NaCl and passed through a column of DEAE cellulose (65 ml bed volume) equilibrated in buffer A containing 300 mM NaCl. The flow through fraction (45 ml) was collected, heated to 70°C for 15 minutes and cleared by centrifugation. The cleared supernatant was diluted with buffer A containing 50 mM NaCl to a total volume of 120 ml. This solution was loaded onto a column of heparin sepharose (70 ml bed volume) equilibrated with buffer A containing 150 mM NaCl. The column was washed with this same buffer and the activity eluted from the column with a linear salt gradient from 150 mM to 700 mM NaCl in buffer A. Active fractions judged to be greater than 90% pure by SDS-PAGE were pooled, concentrated by ultrafiltration and dialyzed against 50 mM phosphate 50% (v/v) glycerol. The sequence is provided in Fig. 4A-C (SEQ. ID. NO. 4).

Other embodiments are within the following claims.

[0102] Plasmid pLS-1 was transformed into *E. coli* strain MZ-1 (λ, c857 lysogen). One liter of culture in LB medium containing 100 μg/ml Ampicillin was grown at 30°C until OD$_{600}$ = 1.1. The temperature was raised to 42°C and growth continued for 2 hours. The cells were harvested by centrifugation (4 g wet weight) and resuspended in 15 ml buffer A (50 mM Tris-HCl pH 8.0, 1 mM DTT, 1 mM PMSF, 1 mM EDTA, 10% (v/v) glycerol) containing 50 mM NaCl. The resuspension was sonicated four times for 30 seconds and the lysate was cleared by centrifugation. The cleared lysate was made 300 mM in NaCl and passed through a column of DEAE cellulose (65 ml bed volume) equilibrated in buffer A containing 300 mM NaCl. The flow through fraction (45 ml) was collected, heated to 70°C for 15 minutes and cleared by centrifugation. The cleared supernatant was diluted with buffer A containing 50 mM NaCl to a total volume of 120 ml. This solution was loaded onto a column of heparin sepharose (70 ml bed volume) equilibrated with buffer A containing 150 mM NaCl. The column was washed with this same buffer and the activity eluted from the column with a linear salt gradient from 150 mM to 700 mM NaCl in buffer A. Active fractions judged to be greater than 90% pure by SDS-PAGE were pooled, concentrated by ultrafiltration and dialyzed against 50 mM phosphate 50% (v/v) glycerol. The sequence is provided in Fig. 4A-C (SEQ. ID. NO. 4).

[0103] Other embodiments are within the following claims.
SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Mamone, Joseph A. Davis, Maria Sha, Dan

(ii) TITLE OF INVENTION: THERMOSTABLE DNA POLYMERASE

(iii) NUMBER OF SEQUENCES: 28

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Lyon & Lyon
(B) STREET: 633 West Fifth Street Suite 4700
(C) CITY: Los Angeles
(D) STATE: California
(E) COUNTRY: U.S.A.
(F) ZIP: 90071-2066

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
(B) COMPUTER: IBM Compatible
(C) OPERATING SYSTEM: IBM P.C. DOS 5.0
(D) SOFTWARE: FastSeq Version 2.0

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: To Be Assigned
(B) FILING DATE:
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:

Prior applications total, including application described below: One

(A) APPLICATION NUMBER: 60/008,688
(B) FILING DATE: December 15, 1995

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Warburg, Richard J.
(B) REGISTRATION NUMBER: 32,327
(C) REFERENCE/DOCKET NUMBER: 223/104

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: (213) 489-1600
(B) TELEFAX: (213) 955-0440
(C) TELEX: 67-3510
(2) INFORMATION FOR SEQ ID NO: 1:

[0105]

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 5300 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xii) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

CTGCAAGAAA AATATAAACAA AGTTTTAAAG AATATGCTTC TATATAACTC TCTCTGAGAT  60
TATTAAATAAA TATTTACCTCA AAAATTGGTT TATCATAAA AATTCTTCAA GAAGGAAAAA  120
GAAGACCTAA AAAAGACGTT TACACACTTG ATACCTTGGC AACCATTTAA TATTTTGAC  180
AGATAATTTAG GAGAATTAGT TATAATATAAC CTTACAATT GGCTTATAAT CTTCTTTCAG  240
TAACCTTGGAA ATGGGACAGA GTAGTGGCTT ACCCTAGAAG AAAGGGACAT ATTATGCAAG  300
AGGATGTCAC GACACTGCCT CAAGGAAAT TACCAAGAGT GACACCTATT CCTGGAGAT  360
AGGAGGGGAC ATGTTGCAAA AGCTTAAATG ATCTTTGCTT GTCGAGGATC AATACCTTTGA  420
TATGAGGAACT TCTGAGGAAA CAGTACCCAG TGAATATATC AAAAGGTTAA ATGGTTTTTG  480
GAAGTGTCGG GAACTGATAT TTGAAACTTG GATTTCCTAA ATATAGATT TTGGAAACTG  540
ACAGCTTTTAA GAGTTGACAG CAAGAGCAAT CAGACGACCT ATATTAGCTT AAAAGCAAAG  600
CTAAAATAAA TGAAGTAGGA GAAACAAAAT GTCGAGACGA TCGTAAAAAT GCTTACATA  660
GTCGCGGAAA TCAGTTTGCAG GAGAGGATCA TTAAACACGT AACGAGTTGA GCTTACAATC  720
AGTTGAGGGAC ATGTTGCAAA AGCTTAAATG ATCTTTGCTT GTCGAGGATC AATACCTTTGA  780
TATGAGGAACT TCTGAGGAAA CAGTACCCAG TGAATATATC AAAAGGTTAA ATGGTTTTTG  840
AAAGTGTTAA AAGAGCATTT TATATAATAT TAAGAATTTC ATATTATTTT ATGAAAATAG  900
TATTAAATAAA TATTTGCTGT AATTTGTCCT TATATGTCAT TGCTCTATAT CTTTTATATA  960
GTTGAAATAA TGAAGTAGGA GAAACAAAAT GTCGAGACGA TCGTAAAAAT GCTTACATA 1020
TTAAATAACA AGAAATTTAA ATGGGAGTGA ATTTAATAGT TAAATTTTTA ATAATTTGAG 1080
GAAGTGTCGG GAACTGATAT TTGAAACTTG GATTTCCTAA ATATAGATT TTGGAAACTG 1140
TCTGGAAACA TGCTCTGATG GGTATCTTAA TGATTGCTAT AAAATCAATC GAGGGGAAA 1200
ACACTGTATA CATGCTATAT CTTTTGAGCA AAAAGGCTTT CATTCTTATA CAAACAAGAT 1260
ATCAAGACATA CAAGAGCTCA AGACAGCAGA TCGTCTAGAA ACCTGCTGAA CAAATGACT 1320
ATTGGGAAA AATATAACTG GCTTCTAATA AAAATACCTA TCAAGACAGA ATGGTTTGA 1380
CTGATGAGAT TATGAGGACAT ATCTGCTTTAC GAGCTCTTTG GAGAGGATCA TTAAACACGT 1440
TATTGAGGAACT TCTGAGGAAA CAGTACCCAG TGAATATATC AAAAGGTTAA ATGGTTTTTG 1500
CTAAAATAAA TGAAGTAGGA GAAACAAAAT GTCGAGACGA TCGTAAAAAT GCTTACATA 1560
GAAGTGTCGG GAACTGATAT TTGAAACTTG GATTTCCTAA ATATAGATT TTGGAAACTG 1620
TCCCTGAGAT CACCTATAAA GGGAAAATA TCTGAGTATA GCTTATTAAG GATTTTTGGA 1680
CAATTTAGGAA TTTATCTCTA AGCTTTAGCA AGCTTTTTAA TCAATTTGCTA GAAAAATAG 1740
AAAATCATAA AGGCTGATCG ATATACTGAA GAGCTCTTTG TACCTTATAC AAAATGACT 1800
CCCTGCAGAT AGAATTGAGT CGTCTGAATA TAAATAAATG AAAATAAAAA TCAAGACAGA 1860
AGCTTTTACA TTTATGAGGA CTTTTTGCAT AATTTGAGCT TAGGTTTTGA TCAAGACAGA 1920
TAGGATGATTG AAGATACTAT AAAAGGTGAA AAGGACCTAA ATGAGTAATA AAGGTTTATG 1980
TAACCTTGTGT CAAAGATACG AGAATTAGA TCTTTTACCA GCTTATTTAA GTATTTTGA 2040
TAAAAATAAT AGCTTTTGCA AGCTTTTGAA TACGCTTTTT TATGAGTTTC TTTGACCAAG 2100
AAGAGTTAAA GAGAGTTTTT GAAAAAAGGT ATTTTGAATT TACACCCCAT GAAAATAGAG 2160
ATTGTTAGGT GAGCTCCTCT TAAATAGGAA TAGAAGTTAA AGCGCATGAC TATGACCTG 2220
CTGAGATGGG CAGTCTGTTT GAAGCTTTTT GAGCTCCTTG TACGTACAGC GCTTTTACT 2280
AAATTTATTT AAGGTTGAA ATGCTTTGTT GCAGTTTGTG TAGCCTGCAAT TCAAGACAGA 2340
AAAAGGAAAAT TAGGAGGACT GAGCAAAGAA TACTCTCCAG TATATGTGAG ATGAGATGTTG 2400
TAGATTCTAT CAAAGATACG AGAATTAGA TCTTTTACCA GCTTATTTAA GTATTTTGA 2460
AGAAGATTAT GTAAGTGTAT AGCATTTTAA GAGCTTTTGCA TACACCCCAT GAAAATAGAG 2520
AGTCTTTATA CTGCTTTTAA GCTTTTACCA GAGCTTTTGA TACACCCCAT GAAAATAGAG 2580
(2) INFORMATION FOR SEQ ID NO: 2:

[0106]

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 872 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xii) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

Met Tyr Lys Phe Leu Ile Ile Asp Gly Ser Ser Leu Met Tyr Arg Ala
<table>
<thead>
<tr>
<th>5</th>
<th>Gly Lys Asp Thr Val Tyr Ile Asp Val Phe Gln Thr Glu Asp Leu Lys</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>Glu Ile Phe Glu Lys Glu Asp Phe Glu Phe Thr Thr His Glu Ile Lys</td>
</tr>
<tr>
<td>15</td>
<td>Asp Phe Leu Val Arg Leu Ser Tyr Lys Gly Ile Glu Cys Lys Ser Lys</td>
</tr>
<tr>
<td>20</td>
<td>Tyr Ile Asp Thr Ala Val Met Ala Tyr Leu Leu Asn Pro Ser Glu Ser</td>
</tr>
<tr>
<td>25</td>
<td>Asn Tyr Asp Leu Asp Arg Val Leu Lys Tyr Leu Lys Val Asp Val</td>
</tr>
<tr>
<td>30</td>
<td>Pro Ser Tyr Glu Gly Ile Phe Gly Lys Gly Arg Asp Lys Lys Lys Ile</td>
</tr>
<tr>
<td>35</td>
<td>Glu Glu Ile Asp Glu Asn Ile Leu Ala Asp Tyr Ile Cys Ser Arg Cys</td>
</tr>
<tr>
<td>40</td>
<td>Val Tyr Leu Phe Asp Leu Lys Glu Lys Leu Met Asn Phe Ile Glu Glu</td>
</tr>
<tr>
<td>45</td>
<td>Met Asp Met Lys Lys Leu Leu Leu Glu Ile Glu Met Pro Leu Val Glu</td>
</tr>
<tr>
<td>50</td>
<td>Val Leu Lys Ser Met Gly Val Ser Gly Phe Thr Leu Asp Lys Glu Val</td>
</tr>
<tr>
<td>55</td>
<td>Leu Lys Glu Leu Ser Gln Lys Ile Asp Asp Arg Ile Gly Glu Ile Leu</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>5</th>
<th>Asp Lys Ile Tyr Lys Glu Ala Gly Tyr Gln Phe Asn Val Asn Ser Pro</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>Lys Glu Leu Ser Glu Phe Leu Phe Glu Lys Leu Asn Leu Pro Val Ile</td>
</tr>
<tr>
<td>15</td>
<td>Lys Lys Thr Lys Thr Gly Tyr Ser Thr Asp Ser Glu Val Leu Glu Gln</td>
</tr>
<tr>
<td>20</td>
<td>Leu Val Pro Tyr Asp Ile Val Ser Asp Ile Ile Glu Tyr Arg Gln</td>
</tr>
<tr>
<td>25</td>
<td>Leu Thr Lys Leu Lys Ser Thr Tyr Ile Asp Gly Phe Leu Pro Leu Met</td>
</tr>
<tr>
<td>30</td>
<td>Asp Glu Asn Arg Val His Ser Asn Phe Lys Glu Met Val Thr Ala</td>
</tr>
<tr>
<td>35</td>
<td>Thr Gly Arg Ile Ser Ser Thr Glu Pro Asn Leu Gln Asn Ile Pro Ile</td>
</tr>
<tr>
<td>40</td>
<td>Arg Glu Glu Phe Gly Arg Gln Ile Arg Arg Ala Phe Ile Pro Arg Ser</td>
</tr>
</tbody>
</table>

19
(2) INFORMATION FOR SEQ ID NO: 3:

[0107]

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 608 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(x) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

Arg Asp Gly Tyr Ile Val Ser Ala Asp Tyr Ser Gln Ile Glu Leu Arg 645 655
Val Leu Ala His Val Ser Gly Asp Glu Leu Ile Glu Ser Phe Met 660 670
Asn Asn Glu Asp Ile His Leu Arg Thr Ala Ser Glu Val Phe Lys Val 675 680
Pro Met Glu Lys Val Thr Pro Glu Met Arg Ala Ala Lys Ala Val 690 695 700
Asn Phe Gly Ile Ile Tyr Gly Ile Ser Asp Tyr Gly Leu Ser Arg Asp 705 710 715 720
Leu Lys Ile Ser Arg Lys Glu Ala Lys Glu Tyr Ile Asn Asn Tyr Phe 725 730 735
Glu Arg Tyr Lys Gly Val Lys Asp Tyr Ile Glu Lys Ile Val Arg Phe 740 745 750
Ala Lys Glu Asn Gly Tyr Val Thr Thr Ile Met Asn Arg Arg Arg Tyr 755 760 765
Ile Pro Glu Ile Asn Ser Arg Asn Phe Thr Gln Arg Ser Gln Ala Glu 770 775 780
Arg Leu Ala Met Asn Ala Pro Ile Gln Gly Ser Ala Ala Asp Ile Ile 785 790 795 800
Lys Met Ala Met Val Lys Val Tyr Asn Leu Lys Leu Lys Leu 805 810 815
Lys Ser Lys Leu Ile Leu Gln Val His Asp Glu Leu Val Val Asp Thr 820 825 830
Tyr Lys Asp Glu Val Asp Ile Ile Lys Lys Leu Lys Glu Asn Met 835 840 845
Glu Asn Val Val Gln Leu Lys Val Pro Leu Val Val Glu Ile Gly Val 850 855 860 865
Gly Pro Asn Trp Phe Leu Ala Lys 870
Met Lys Leu Leu Glu Leu Phe Asn Lys Leu Glu Phe Phe Ser Leu Ile
1  5  10  15
Asp Asn Ile Lys Lys Glu Ser Ser Ile Glu Ile Val Asp Asn His Lys
20  25  30
Val Glu Lys Trp Ser Lys Val Asp Ile Lys Glu Leu Val Thr Leu Leu
35  40  45
Gln Asp Asn Arg Asn Ile Ala Phe Tyr Pro Leu Ile Tyr Glu Gly Glu
50  55  60
Ile Lys Lys Ile Ala Phe Ser Phe Gly Lys Asp Thr Val Tyr Ile Asp
65  70  75  80
Val Phe Glu Thr Glu Asp Leu Lys Glu Ile Phe Glu Lys Glu Asp Phe
85  90  95
Glu Phe Thr Thr His Glu Ile Lys Asp Phe Leu Val Arg Leu Ser Tyr
100 105 110
Lys Gly Ile Glu Cys Lys Ser Lys Tyr Ile Asp Thr Ala Val Met Ala
115 120 125
Tyr Leu Leu Asn Pro Ser Glu Ser Asn Tyr Asp Leu Arg Val Leu
130 135 140
Lys Lys Tyr Leu Lys Val Asp Val Pro Ser Tyr Glu Gly Ile Phe Gly
145 150 155 160
Lys Gly Arg Asp Lys Lys Ile Glu Gly Ile Asp Glu Asn Ile Leu
165 170 175
Ala Asp Tyr Ile Cys Ser Arg Cys Val Tyr Leu Phe Asp Leu Lys Glu
180 185 190
Lys Leu Met Asn Phe Ile Glu Glu Met Asp Met Lys Leu Leu Leu
195 200 205
Glu Ile Glu Met Pro Leu Val Glu Val Leu Lys Ser Met Glu Val Ser
210 215 220
Gly Phe Thr Leu Asp Lys Glu Val Leu Lys Glu Leu Ser Gln Lys Ile
225 230 235 240
Asp Asp Arg Ile Gly Glu Ile Leu Asp Lys Ile Tyr Lys Glu Ala Gly
245 250 255
Tyr Gln Phe Asn Val Asn Ser Pro Lys Gln Leu Ser Glu Phe Leu Phe
260 265 270
Glu Lys Leu Asn Leu Pro Val Ile Lys Thr Lys Thr Gly Tyr Ser
275 280 285
Thr Asp Ser Glu Val Leu Glu Gln Leu Val Pro Tyr Asn Asp Ile Val
290 295 300
Ser Asp Ile Ile Glu Tyr Arg Gln Leu Thr Lys Leu Lys Ser Thr Tyr
305
### INFORMATION FOR SEQ ID NO: 4:

#### [0108]

(i) SEQUENCE CHARACTERISTICS:

<table>
<thead>
<tr>
<th>Description</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(A) LENGTH</td>
<td>578 amino acids</td>
</tr>
<tr>
<td>(B) TYPE</td>
<td>amino acid</td>
</tr>
<tr>
<td>(C) STRANDEDNESS</td>
<td>single</td>
</tr>
<tr>
<td>(D) TOPOLOGY</td>
<td>linear</td>
</tr>
</tbody>
</table>

### Amino Acid Sequence

<table>
<thead>
<tr>
<th>Position</th>
<th>Amino Acid 1</th>
<th>Amino Acid 2</th>
<th>Amino Acid 3</th>
<th>Amino Acid 4</th>
<th>Amino Acid 5</th>
<th>Amino Acid 6</th>
<th>Amino Acid 7</th>
<th>Amino Acid 8</th>
<th>Amino Acid 9</th>
<th>Amino Acid 10</th>
</tr>
</thead>
<tbody>
<tr>
<td>305-310</td>
<td>Ile Asp Gly</td>
<td>Phe Leu Pro</td>
<td>Leu Met Asp</td>
<td>Glu Asn Asn</td>
<td>Arg Val His</td>
<td>Ser 335</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>315-320</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>325</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>330</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>335</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>340-345</td>
<td>Asn Phe Lys</td>
<td>Gln Met Val</td>
<td>Thr Ala Thr</td>
<td>Gly Arg</td>
<td>Ile Ser Ser</td>
<td>Ser Glu 350</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>350</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>355</td>
<td>Ala Asn Leu</td>
<td>Gln Asn Ile</td>
<td>Arg Pro Ile</td>
<td>Glu Glu Phe</td>
<td>Gly Arg Gln</td>
<td>Ile 365</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>360-365</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>370-375</td>
<td>Arg Arg Ala</td>
<td>Phe Ile Pro</td>
<td>Arg Cys Asp</td>
<td>Tyr Ile Val</td>
<td>Ser Ala</td>
<td>380</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>380-390</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>395-400</td>
<td>Glu Lys Leu</td>
<td>Ile Glu Ser</td>
<td>Phe Mtn Asn</td>
<td>Asn Glu Asp</td>
<td>Ile His Leu</td>
<td>Arg 410</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>415-420</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>425-430</td>
<td>Thr Ala Ser</td>
<td>Glu Val Phe</td>
<td>Lys Val Pro</td>
<td>Met Glu Lys</td>
<td>Val Thr Pro</td>
<td>Glu 430</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>440-445</td>
<td>Met Arg Arg</td>
<td>Ala Ala Lys</td>
<td>Ala Val Asn</td>
<td>Phe Gly Ile</td>
<td>Tyr Gly Ile</td>
<td>445</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>450-455</td>
<td>Ser Asp Tyr</td>
<td>Gly Leu Ser</td>
<td>Arg Asp Leu</td>
<td>Lys Ile Ser</td>
<td>Arg Lys Glu</td>
<td>Ala 460</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>465-470</td>
<td>Lys Glu Tyr</td>
<td>Ile Glu Tyr</td>
<td>Phe Glu Arg</td>
<td>Tyr Lys Gly</td>
<td>Val Lys Asp</td>
<td>475</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>480</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>485-490</td>
<td>Tyr Ile Glu</td>
<td>Ile Val Arg</td>
<td>Phe Ala Lys</td>
<td>Glu Asn Gly</td>
<td>Tyr Val Thr</td>
<td>495</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>500-505</td>
<td>Thr Ile Met</td>
<td>Asn Arg Arg</td>
<td>Tyr Ile Pro</td>
<td>Glu Ile Asn</td>
<td>Ser Arg Asn</td>
<td>510</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>515-520</td>
<td>Phe Thr Glu</td>
<td>Arg Ser Glu</td>
<td>Ala Glu Arg</td>
<td>Leu Ala Met</td>
<td>Asn Ala Pro</td>
<td>Ile 525</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>530-535</td>
<td>Gln Gly Ser</td>
<td>Ala Ala Asp</td>
<td>Ile Lys Met</td>
<td>Ala Met Val</td>
<td>Lys Val Tyr</td>
<td>540</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>545-550</td>
<td>Asn Asp Leu</td>
<td>Lys Leu Lys</td>
<td>Leu Lys Ser</td>
<td>Lys Leu Ile</td>
<td>Leu Gln Val</td>
<td>560</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>555-560</td>
<td>His Asp Glu</td>
<td>Leu Val Val</td>
<td>Asp Thr Tyr</td>
<td>Lys Asp Glu</td>
<td>Val Asp Ile</td>
<td>575</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>570-575</td>
<td>Lys Lys Ile</td>
<td>Lys Glu Asn</td>
<td>Met Glu Asn</td>
<td>Val Val Glu</td>
<td>Lys Lys Val</td>
<td>590</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>580-585</td>
<td>Pro Leu Val</td>
<td>Val Glu Ile</td>
<td>Gly Val Gly</td>
<td>Pro Asn Trp</td>
<td>Phe Leu Ala</td>
<td>Lys 605</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>595-600</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
(ii) MOLECULE TYPE: peptide

(x) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

<table>
<thead>
<tr>
<th></th>
<th>Met</th>
<th>Lys</th>
<th>Val</th>
<th>Glu</th>
<th>Lys</th>
<th>Trp</th>
<th>Ser</th>
<th>Lys</th>
<th>Val</th>
<th>Asp</th>
<th>Ile</th>
<th>Lys</th>
<th>Glu</th>
<th>Leu</th>
<th>Val</th>
<th>Thr</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Leu</td>
<td>Leu</td>
<td>Gln</td>
<td>Asp</td>
<td>Asn</td>
<td>Arg</td>
<td>Asn</td>
<td>Ile</td>
<td>Ala</td>
<td>Phe</td>
<td>Tyr</td>
<td>Pro</td>
<td>Leu</td>
<td>Ile</td>
<td>Tyr</td>
<td>Glu</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Gly</td>
<td>Glu</td>
<td>Ile</td>
<td>Lys</td>
<td>Lys</td>
<td>Ile</td>
<td>Ala</td>
<td>Phe</td>
<td>Ser</td>
<td>Phe</td>
<td>Gly</td>
<td>Lys</td>
<td>Asp</td>
<td>Thr</td>
<td>Val</td>
<td>Tyr</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Ile</td>
<td>Asp</td>
<td>Val</td>
<td>Phe</td>
<td>Gln</td>
<td>Thr</td>
<td>Glu</td>
<td>Asp</td>
<td>Leu</td>
<td>Lys</td>
<td>Glu</td>
<td>Ile</td>
<td>Phe</td>
<td>Glu</td>
<td>Lys</td>
<td>Glu</td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Asp</td>
<td>Phe</td>
<td>Glu</td>
<td>Phe</td>
<td>Thr</td>
<td>Thr</td>
<td>His</td>
<td>Glu</td>
<td>Ile</td>
<td>Lys</td>
<td>Asp</td>
<td>Phe</td>
<td>Leu</td>
<td>Val</td>
<td>Arg</td>
<td>Leu</td>
</tr>
<tr>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Ser</td>
<td>Tyr</td>
<td>Lys</td>
<td>Gly</td>
<td>Ile</td>
<td>Glu</td>
<td>Cys</td>
<td>Lys</td>
<td>Ser</td>
<td>Lys</td>
<td>Tyr</td>
<td>Ile</td>
<td>Asp</td>
<td>Thr</td>
<td>Ala</td>
<td>Val</td>
</tr>
<tr>
<td>11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>Met</td>
<td>Ala</td>
<td>Tyr</td>
<td>Leu</td>
<td>Leu</td>
<td>Asn</td>
<td>Pro</td>
<td>Ser</td>
<td>Glu</td>
<td>Ser</td>
<td>Tyr</td>
<td>Asp</td>
<td>Leu</td>
<td>Asp</td>
<td>Arg</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>Val</td>
<td>Leu</td>
<td>Lys</td>
<td>Tyr</td>
<td>Leu</td>
<td>Lys</td>
<td>Val</td>
<td>Asp</td>
<td>Val</td>
<td>Pro</td>
<td>Ser</td>
<td>Tyr</td>
<td>Glu</td>
<td>Gly</td>
<td>Ile</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>Phe</td>
<td>Gly</td>
<td>Lys</td>
<td>Tyr</td>
<td>Leu</td>
<td>Lys</td>
<td>Val</td>
<td>Asp</td>
<td>Arg</td>
<td>Lys</td>
<td>Lys</td>
<td>Ile</td>
<td>Glu</td>
<td>Glu</td>
<td>Ile</td>
<td>Asp</td>
</tr>
<tr>
<td>17</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>Ile</td>
<td>Leu</td>
<td>Ala</td>
<td>Asp</td>
<td>Tyr</td>
<td>Ile</td>
<td>Cys</td>
<td>Ser</td>
<td>Arg</td>
<td>Cys</td>
<td>Val</td>
<td>Tyr</td>
<td>Leu</td>
<td>Phe</td>
<td>Asp</td>
<td>Leu</td>
</tr>
<tr>
<td>19</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>Lys</td>
<td>Glu</td>
<td>Lys</td>
<td>Leu</td>
<td>Met</td>
<td>Asn</td>
<td>Phe</td>
<td>Ile</td>
<td>Glu</td>
<td>Glu</td>
<td>Met</td>
<td>Asp</td>
<td>Met</td>
<td>Lys</td>
<td>Lys</td>
<td>Leu</td>
</tr>
<tr>
<td>21</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>Leu</td>
<td>Leu</td>
<td>Glu</td>
<td>Ile</td>
<td>Glu</td>
<td>Met</td>
<td>Pro</td>
<td>Leu</td>
<td>Val</td>
<td>Glu</td>
<td>Val</td>
<td>Leu</td>
<td>Lys</td>
<td>Ser</td>
<td>Met</td>
<td>Glu</td>
</tr>
<tr>
<td>23</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>Val</td>
<td>Ser</td>
<td>Gly</td>
<td>Phe</td>
<td>Thr</td>
<td>Leu</td>
<td>Asp</td>
<td>Lys</td>
<td>Glu</td>
<td>Val</td>
<td>Leu</td>
<td>Lys</td>
<td>Glu</td>
<td>Leu</td>
<td>Ser</td>
<td>Gln</td>
</tr>
<tr>
<td>25</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>Lys</td>
<td>Ile</td>
<td>Asp</td>
<td>Asp</td>
<td>Arg</td>
<td>Ile</td>
<td>Gly</td>
<td>Glu</td>
<td>Ile</td>
<td>Leu</td>
<td>Asp</td>
<td>Lys</td>
<td>Ile</td>
<td>Tyr</td>
<td>Lys</td>
<td>Glu</td>
</tr>
<tr>
<td>27</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>Ala</td>
<td>Gly</td>
<td>Tyr</td>
<td>Gln</td>
<td>Phe</td>
<td>Asn</td>
<td>Val</td>
<td>Asn</td>
<td>Ser</td>
<td>Pro</td>
<td>Lys</td>
<td>Gln</td>
<td>Leu</td>
<td>Ser</td>
<td>Glu</td>
<td>Phe</td>
</tr>
<tr>
<td>29</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>Leu</td>
<td>Phe</td>
<td>Glu</td>
<td>Lys</td>
<td>Leu</td>
<td>Asn</td>
<td>Leu</td>
<td>Pro</td>
<td>Val</td>
<td>Ile</td>
<td>Lys</td>
<td>Thr</td>
<td>Lys</td>
<td>Thr</td>
<td>Gly</td>
<td></td>
</tr>
<tr>
<td>31</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>Tyr</td>
<td>Ser</td>
<td>Thr</td>
<td>Asp</td>
<td>Ser</td>
<td>Glu</td>
<td>Val</td>
<td>Leu</td>
<td>Glu</td>
<td>Gln</td>
<td>Leu</td>
<td>Val</td>
<td>Pro</td>
<td>Tyr</td>
<td>Asn</td>
<td>Asp</td>
</tr>
</tbody>
</table>
EP 0 866 868 B1

260  265  270
Ile Val Ser Asp Ile Ile Glu Tyr Arg Gln Leu Thr Lys Leu Lys Ser
   280  285
Thr Tyr Ile Asp Gly Phe Leu Pro Leu Met Asp Glu Asn Asn Arg Val
   290  295  300
His Ser Asn Phe Lys Gln Met Val Thr Ala Thr Gly Arg Ile Ser Ser
   305  310  315  320
Ser Glu Ala Asn Leu Gln Asn Ile Pro Ile Arg Glu Phe Gly Arg
   325  330  335
Gln Ile Arg Arg Ala Phe Ile Pro Arg Ser Arg Asp Gly Tyr Ile Val
   340  345  350
Ser Ala Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala His Val Ser
   355  360  365
Gly Asp Glu Lys Leu Ile Glu Ser Phe Met Asn Asn Glu Asp Ile His
   370  375  380
Leu Arg Thr Ala Ser Gln Val Phe Lys Val Pro Met Glu Lys Val Thr
   385  390  395  400
Pro Glu Met Arg Arg Ala Ala Lys Ala Val Asn Tyr Gly Ile Ile Tyr
   405  410  415
Gly Ile Ser Asp Tyr Gly Leu Ser Arg Asp Leu Lys Ile Ser Arg Lys
   420  425  430
Glu Ala Lys Glu Tyr Ile Asn Asn Tyr Phe Glu Arg Tyr Lys Gly Val
   435  440  445
Lys Asp Tyr Ile Glu Lys Ile Val Arg Phe Ala Lys Glu Asn Gly Tyr
   450  455  460
Val Thr Thr Ile Met Asn Arg Arg Arg Tyr Ile Pro Glu Ile Asn Ser
   465  470  475  480
Arg Asn Phe Thr Gln Arg Ser Gln Ala Glu Arg Leu Ala Met Asn Ala
   485  490  495
Pro Ile Gln Gly Ser Ala Ala Asp Ile Ile Lys Met Ala Met Val Lys
   500  505  510
Val Tyr Asn Asp Leu Lys Lys Leu Lys Leu Lys Ser Lys Leu Ile Leu
   515  520  525
Gln Val His Asp Glu Leu Val Asp Thr Tyr Lys Asp Glu Val Asp
   530  535  540
Ile Ile Lys Lys Ile Leu Lys Glu Asn Met Glu Asn Val Val Gln Leu
   545  550  555  560
Lys Val Pro Leu Val Glu Ile Gly Val Gly Pro Asn Trp Phe Leu
   565  570  575
Ala Lys

(2) INFORMATION FOR SEQ ID NO: 5:

[0109]

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

<table>
<thead>
<tr>
<th>GGTTCACAC GATCAAGCCG AGTTACATGA</th>
</tr>
</thead>
</table>

(2) INFORMATION FOR SEQ ID NO: 6:

[0110]

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

<table>
<thead>
<tr>
<th>TCCCATGT TGTGAAAAA AGGTATTAGC</th>
</tr>
</thead>
</table>

(2) INFORMATION FOR SEQ ID NO: 7:

[0111]

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 23 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ix) FEATURE:

(D) OTHER INFORMATION: "H" is A, T or C; "Y" is C or T; "S" is C or G; and "R" is A or G.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

<table>
<thead>
<tr>
<th>GACCHAACY TSCARAAAYAT HCC</th>
</tr>
</thead>
</table>

(2) INFORMATION FOR SEQ ID NO: 8:

[0112]

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ix) FEATURE:
(D) OTHER INFORMATION: "K" is G or T; "S" is C or G; "Y" is C or T; "R" is A or G; and "N" is G, A, T or C.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

KASSAKYTCR TCGTNACYT G

(2) INFORMATION FOR SEQ ID NO: 9:

0113

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 32 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

GGGGTACCTC TAGATCATT GGCCAAAAAC CA

(2) INFORMATION FOR SEQ ID NO: 10:

0114

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 43 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

GGAATTCCTG ATGGCTTACC TGCTGAACCC GTCTGAGTCT AAC

(2) INFORMATION FOR SEQ ID NO: 11:

0115

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 37 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

GGAATTCCTG ATGGCTTTT CTTATAAAGG AATAGAG
(2) INFORMATION FOR SEQ ID NO: 12:

[0116]

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 31 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xii) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

GGAATTCAT ATGAAAGTG AAAATGGTC A 31

(2) INFORMATION FOR SEQ ID NO: 13:

[0117]

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 53 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xii) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

GGAATTCAT ATGAAAGTG AAAATGGTC TAAAGTAGAC ATCAAAGAT TAG 53

(2) INFORMATION FOR SEQ ID NO: 14:

[0118]

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xii) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

GGAATTCAT ATGAAAGTTT TAGAGCTTTT 30

(2) INFORMATION FOR SEQ ID NO: 15:

[0119]

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 46 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xii) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

GGAATTCCAT ATGAAACTTT TAGAGCTTTT CAACAAATTA GAATTC

(2) INFORMATION FOR SEQ ID NO: 16:

[0120]

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 15 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xii) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

CGATCGAGCA AGCCA

(2) INFORMATION FOR SEQ ID NO: 17:

[0121]

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 15 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xii) SEQUENCE DESCRIPTION: SEQ ID NO: 17:

CGAGCCGCTC GCTGA

(2) INFORMATION FOR SEQ ID NO: 18:

[0122]

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 40 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xii) SEQUENCE DESCRIPTION: SEQ ID NO: 18:

ACCGCATCGA ATGCATGTCT CGGTAAGGC GTACTCGACC
(2) INFORMATION FOR SEQ ID NO: 19:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 40 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:

CGATTCGCT CCAGACTTCT CGGGTGTA CT GAGATCCCT

40

(2) INFORMATION FOR SEQ ID NO: 20:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 14 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:
(D) OTHER INFORMATION: "Xaa" is any amino acid.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 20:

Arg Xaa Xaa Xaa Lys Xaa Xaa Xaa Xaa Phe Xaa Xaa Tyr Gly
1 5 10

(2) INFORMATION FOR SEQ ID NO: 21:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 14 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 21:

Arg Arg Ala Ala Lys Ala Val Asn Phe Gly Ile Ile Tyr Gly
1 5 10
(2) INFORMATION FOR SEQ ID NO: 22:

[0126]

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 54 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 22:

AAGGTCAGAA AGGCCATAAT CGCTTATGCC ATATATTATG CCATAATTTA CGGC 54

(2) INFORMATION FOR SEQ ID NO: 23:

[0127]

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2622 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 23:

ATGCTTATA AATTTTTAAT CATTGATGGT AGTAGCCTCA TGACAGAGC CTATTATGCC
60
TTGCCCATGC TTACTACAAG TGAGGGATTG CCTACAATATG CTCTGTATGG TTTTACTATG
120
ATGCTTTATA AACTTATCGA GAGGAAAAA CCTGATTACA TAGCTATGC TTTGACAAA
180  AAGCTCCTA CTTTAGACCA CAAAGATATC CAAGACTACA AAGCTCAAG ACAAGCTATG
240  CTTGAGAAC CGTACACT AATGAGCTG TTGAAAGAAA TTATAGATGG CTTTTAATATA
300  AAGACATTAG AATTAGAAGG TTATGAAGCT GATGCACATA TTGGGACATG TTCAAGAAGCTG
360  GCGAGGAAA AAGGAATGGG AGTGCCCTGTG CCTACAGAGG AGCAGAGATGC TCTCCAATTA
420  GTTTTCAGATA AATGGAAGAT AAAATAATCT AAAAAAGGTTA TTACTCAGAT GGAAGACTTT
480  GACGAAAAGG CTAATTTAGA AAGGTATGGA ATAAACCCCTG ACAGGTATAG AGATTAAAAAA
540  GCCCTTATGG GAGATAAATC TGTAATAATC CCTGGGACTAC CTAATATAGG CAGAAAAACT
600  GCGATTACG TATTAAGGA TTTTGGAAAC ATTTGAAATT TAAATCCAAA TCTTTTCAG
660  CTTAAGGATA AATAAAAGA AATAATAGAA AACAATAAAG AGTGAATGAT AATGAGTAAG
720  AGGCATTAGTCA TATAAAGAG AGACATTCCC AGTGAATAGT ATTTTTGAGA GTAAAAGTTA
780  AAAAAATTTA ATGAGGAAAG GCTTTTAGAG CTTTTTAAATA AATTAGAATT TTATAGTTTA
840  ATGTGAACAA TAAAGAACAG AAGTAGACAT GAGATTGATG ATAAACCCATA AGTTGAAAAA
900  TGGTCAAAAG TAGATATAAA AGAATATTAGA ACTGTTGATGC AGATAACAG AAATATGTCT
960  TTATACCGGT TAATATAGA AGGGGAATAA AAAAAATAG GCCTTTCTTT TTGGAAGGAT
1020  AGCCTTATATA TTGAGCTTTTT CCAACACGAA GATTTAAGAG AGATTTGAGA AAAAGAAGAT
1080  TTTGAAATTCA CAACCCATGA AATAAAAGAT TTTTTAGTGGA GCCTTTCTTA TAAAGCATAA
1140  GTATTGAAAAG GCAAGTACTAT AGATACTGCT GGAATGCTTG ATCTCTCTGA TCCCTTGAG
1200  TCTAATGATT AGCTTAGACCC TGCTGAAAA AAAAAATTTA AGGATGATGT GCCTTTCTTT
1260  GAGGAAATAT TTGGAACAGG TAGGATATAA AAGAAAATGT AAGAGATTGA CGAAAACATA
1320  CTTGCCTGTAT ATATTTTGAG TAGATGCTTG TATCTATATTG ATTAAAAGAA AAGCCTGATG
1380  ATATTATTAG AAGAGAATGGG TATGAAAAAA TCTCTATATT AATAGAATAT GCCTCTTGTAT
1440  GAAATTTTAA AATCAATGGG GTTAAGTGGTT TTATATTGGA ATAAAGAAGAT TCTAAAGAG
1500  CTTTCACAAA AGATAGATGA TAGAATAGGA GAAATACTAG ATAAAATTTA TAAAGAGGCA
1560  GGAATACTAT TAAATGAAAA TTCACTCTAG CAATTAAGTG ATATTGTGGT TGAAGAGTTA
1620  AACATTACAG TAATAAGGAA ACAAAACACA GGAATCTCTA CGGATCTCAG AGTTTTGAAA
1680  CAAATGTTTGTC TTTATAATGA TTTTGTCAG GATATAATAG AGTATCGGGA ACTTACAAAA
1740  CTTAATATCTA CTTATATAGA TGGATTTTTG CCTCTATGGA AGAAGAACA TAGAATCAT
1800  TCTAATTTTTA AACAAATGTT TATGTGCTACA GGTGAATATA CGACACACAG GCAAAATCTA
1860  CAAAATATAC ATATAAGGAGG AGATTTGACG AAGACAAATTA GAAGGCGTTT TATTTAGG
1920
AGTAGAGATG GATATATTGT TTCCAGCGAT TATTTTCAGA TTGAACTGAG GGT TTAGCA 1980
CATGTTCCGG GAGATGAAAA GCTAAAAGAA TCTTTTATGA ATAATGAAGA TATACATTTA 2040
AGGACAGCTT CGGAGGTTTT TAAAGTTCT ATGGAAAAG TTACACCGGA GATGAGAAGA 2100
GCAGCAAAAG CCGTAAATTT TGGCATAATA TATGGCATAA GCGATTATGG GCTT TCTGGA 2160
GACCCTAAAA TATCAAGAAA AGAAGCCAAA GAGTACATAA ATAATTTATG TGAAAGATAT 2220
AAAGAGGATTTA AAGATTATAT TGA AAAAATAA GTACGATTGG CAAAGAAAA TGGCTATGTG 2280
ACTACAATAA TGAACAGAAG GAGATATATT CCTGAAATAA ACTCAAGAAA TTTTACTCAA 2340
AGATGCGAGG CCGGAAGGT TTAGCAATGA GCTCCGATAC AGGGAAGTGCC GGCTGATATA 2400
ATAAAAAATGG CAATGGTTAA GGTATACAC GATTTAAAAA AATTAAAGCT TAAGTCTAAAG 2460
CTTATATTGC AAGTTCATGA CGAGCTTGTA GTGGATACTT ATAAGGATGA AGTAGATATC 2520
ATAAAAAAGA TACCTAAAGA AAATATGGGA AATGTAGTGC AATTTAAAAT TCCTCTGGTTT 2580
GTTGAAATTT GCGTAGGGCC TAATTGTGTT TTGGCAAGT GA 2622

[0128] INFORMATION FOR SEQ ID NO: 24:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 46 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 24:

GGAATTCAT ATGGCTTATA AATTTTTAAT CATTGAGGT AGTAGC 46

[0129] INFORMATION FOR SEQ ID NO: 25:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 50 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 25:

TCAGTGGTAT TTGTGAGCCA GGGCATGGCC CACACAGGC ACCACCTTCT 50
(2) INFORMATION FOR SEQ ID NO: 26:

[0130]

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 34 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 26:

ATCCCCAAA ACAGACAGAA TGGTGATCT GTCC

(2) INFORMATION FOR SEQ ID NO: 27:

[0131]

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 27:

GTTCCTCTGG AAGGAGTGC ACATGGCCCT

(2) INFORMATION FOR SEQ ID NO: 28:

[0132]

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 34 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 28:

CACACAAGAG TGAGATGCCC CATCGTTTTA AGGA

Claims

1. An enzymatically active DNA polymerase or fragment thereof having at least 80% homology in its amino acid sequence to at least a contiguous 40 amino acid sequence as given in SEQ.ID.NO.2 of the DNA polymerase of Thermoanaerobacter thermohydrosulfuricus.
2. The polymerase of claim 1 wherein said DNA polymerase has an exonuclease activity removed.

3. The polymerase of claim 1 or 2 wherein said polymerase has one or more amino acids deleted from the amino
   terminus or amino acid changes to remove said exonuclease activity.

4. The polymerase of claims 1 to 3 wherein said polymerase has the phenylalanine at position 706 replaced with
   tyrosine.

5. A stable enzymatic composition comprising a purified DNA polymerase of any claims 1-4 in a buffer.

6. Purified nucleic acid encoding the DNA polymerase of any claims 1-4.

7. Host cell comprising a vector containing a nucleic acid sequence encoding the DNA polymerase of any of claims 1-4.

8. A DNA polymerase of any claims 1-4 and a thermostable restriction enzyme for use in strand displacement amplification.

9. Method for generating and amplifying a nucleic acid fragment by strand displacement utilising a DNA polymerase
   or fragment thereof having at least 80% homology in its amino acid sequence to at least a contiguous 40 amino
   acid sequence as given in SEQ.ID.NO.2 of the DNA polymerase of Thermoanaerobacter thermohydrodsulfuricus.

10. Kit for strand displacement amplification of DNA comprising a DNA polymerase of claim 1-4 and a thermostable
    restriction enzyme.

11. Method for sequencing DNA whereby at least a part of the nucleotide base sequence of a DNA template can be
    determined comprising the step of generating chain terminated fragments from said DNA template with a DNA
    polymerase of claim 4, in the presence of at least one chain terminating agent and determining the sequence of
    at least part of said DNA template from the sizes of said fragments.

12. Kit for sequencing DNA comprising a DNA polymerase of claim 4 and pyrophosphatase.

13. Method for preparing cDNA by providing an oligonucleotide primer, a polymerase of any claims 1-4 and between
    one and four deoxyribonucleoside phosphates.

14. A purified DNA polymerase according to claim 1 or 2 which has reverse transcriptase activity.

15. A DNA polymerase of claim 1 or 2, useful as a reverse transcriptase.

16. Method for reverse transcription/polymerase chain reaction utilizing a DNA polymerase of claim 1 or 2 and a
    polymerase suitable for carrying out polymerase chain reaction in a single reaction vessel.

17. The method of claim 16, wherein said polymerase suitable for carrying out polymerase chain reaction is Taq DNA
    polymerase.

18. Kit for reverse transcription/polymerase chain reaction comprising a DNA polymerase of claim 1 or 2 and a polymer-
   ase suitable for carrying out polymerase chain reaction.

19. Solution for use in RT/PCR comprising DNA polymerase of claim 1 or 2 and a polymerase suitable for carrying
    our polymerase chain reaction.

20. The solution of claim 19, wherein said polymerase suitable for carrying out polymerase chain reaction is Taq DNA
    polymerase.

21. Solution for strand displacement amplification of DNA comprising a DNA polymerase of claims 1 - 4 and a ther-
    mostable restriction enzyme.

22. Solution for sequencing DNA comprising a DNA polymerase of claims 1 - 4 and pyrophosphatase.
Patentansprüche

1. Enzymatisch aktive DNA-Polymerase oder ein Fragment hiervon mit einer Homologie von mindestens 80% in ihrer/seiner Aminosäuresequenz zu mindestens einer zusammenhängenden Sequenz von 40 Aminosäuren, wie in SEQ ID Nr. 2 angegeben, der DNA-Polymerase von Thermoanaerobacter thermohydrosulfuricus.

2. Polymerase nach Anspruch 1, wobei die Exonucleaseaktivität der DNA-Polymerase entfernt wurde.

3. Polymerase nach Anspruch 1 oder 2, wobei die Polymerase eine oder mehrere Aminosäuredeletionen am Aminoterminus oder einen oder mehrere Aminosäureaustausche aufweist, um die Exonucleaseaktivität zu entfernen.

4. Polymerase nach Anspruch 1 bis 3, wobei in der Polymerase das Phenylalanin an Position 706 durch Tyrosin ersetzt wurde.

5. Stabile enzymatische Zusammensetzung, umfassend eine gereinigte DNA-Polymerase nach einem der Ansprüche 1 bis 4 in einem Puffer.

6. Gereinigte Nucleinsäure, welche die DNA-Polymerase nach einem der Ansprüche 1 bis 4 codiert.

7. Wirtszenile, umfassend einen Vektor, der eine Nucleinsäuresequenz enthält, welche die DNA-Polymerase nach einem der Ansprüche 1 bis 4 codiert.

8. DNA-Polymerase nach einem der Ansprüche 1 bis 4 und ein thermostabiles Restriktionsenzym zur Verwendung in einer Strangverdrängungsamplifikation.


10. Kit zur Strangverdrängungsamplifikation von DNA, umfassend eine DNA-Polymerase nach Anspruch 1 bis 4 und ein thermostabiles Restriktionsenzym.


12. Kit zur Sequenzierung von DNA, umfassend eine DNA-Polymerase nach Anspruch 4 und eine Pyrophosphatase.

13. Verfahren zur Herstellung von cDNA durch Vorsehen eines Oligonucleotidprimers, einer Polymerase nach einem der Ansprüche 1 bis 4 und zwischen einem und vier Desoxyribonucleosidphosphaten.

14. Gereinigte DNA-Polymerase nach Anspruch 1 oder 2, welche eine reverse Transkriptaseaktivität aufweist.

15. DNA-Polymerase nach Anspruch 1 oder 2, welche als eine reverse Transkriptase verwendbar ist.

16. Verfahren zur reversen Transkription/Polymerasekettenreaktion unter Verwendung einer DNA-Polymerase nach Anspruch 1 oder 2 und einer Polymerase, welche zur Durchführung der Polymerasekettenreaktion in einem Einzelreaktionsgefäße geeignet ist.

17. Verfahren nach Anspruch 16, wobei die zur Durchführung der Polymerasekettenreaktion geeignete Polymerase eine Taq-DNA-Polymerase ist.

18. Kit zur reversen Transkription/Polymerasekettenreaktion, umfassend eine DNA-Polymerase nach Anspruch 1 oder 2 und eine zur Durchführung der Polymerasekettenreaktion geeignete Polymerase.

19. Lösung zur Verwendung bei der RT/PCR, umfassend eine DNA-Polymerase nach Anspruch 1 oder 2 und eine zur
Durchführung der Polymerasekettenreaktion geeignete Polymerase.

20. Lösung nach Anspruch 19, wobei die zur Durchführung der Polymerasekettenreaktion geeignete Polymerase eine Taq-DNA-Polymerase ist.

21. Lösung zur Strangverdrängungsamplifikation von DNA, umfassend eine DNA-Polymerase nach Anspruch 1 bis 4 und ein thermostabiles Restrikionsenzym.

22. Lösung zur Sequenzierung von DNA, umfassend eine DNA-Polymerase nach Anspruch 1 bis 4 und eine Pyrophosphatase.

**Revendications**

1. ADN polymérase enzymatiquement active ou l'un de ses fragments ayant au moins 80 % d'homologie dans sa séquence d'acides aminés avec au moins une séquence de 40 acides aminés contigus, comme donnée dans SEQ.ID.NO.2, de l'ADN polymérase de *Thermoanaerobacter thermohydrosulfuricus*.

2. Polymérase selon la revendication 1, l'activité exonucléase de ladite ADN polymérase ayant été éliminée.

3. Polymérase selon la revendication 1 ou la revendication 2, ladite polymérase ayant un ou plusieurs acides aminés supprimés de son extrémité amino-terminale ou un ou plusieurs changements d'acides aminés pour éliminer ladite activité exonucléase.

4. Polymérase selon les revendications 1 à 3, ladite polymérase ayant la phénylalanine de la position 706 remplacée par une tyrosoine.

5. Composition enzymatique stable comprenant une ADN polymérase purifiée de l'une quelconque des revendications 1 à 4 dans un tampon.


7. Cellule hôte comprenant un vecteur contenant une séquence d'acide nucléique codant l'ADN polymérase de l'une quelconque des revendications 1 à 4.

8. ADN polymérase de l'une quelconque des revendications 1 à 4 et enzyme de restriction thermostable destinées à être utilisées dans de l'amplification par déplacement de brin.

9. Méthode pour générer et amplifier un fragment d'acide nucléique par déplacement de brin utilisant une ADN polymérase ou l'un de ses fragments ayant au moins 80 % d'homologie dans sa séquence d'acides aminés avec au moins une séquence de 40 acides aminés contigus, comme donnée dans SEQ.ID.NO.2, de l'ADN polymérase de *Thermoanaerobacter thermohydrosulfuricus*.

10. Kit pour l'amplification d'ADN par déplacement de brin comprenant une ADN polymérase de la revendication 1 à 4 et une enzyme de restriction thermostable.

11. Méthode de séquençage d'ADN par laquelle au moins une partie de la séquence de bases nucléotidiques d'une matrice d'ADN peut être déterminée, comprenant les étapes consistant à générer des fragments à chaîne terminée à partir de ladite matrice d'ADN avec une ADN polymérase de la revendication 4, en présence d'au moins un agent de terminaison de chaîne, et à déterminer la séquence d'au moins une partie de ladite matrice d'ADN d'après les tailes desdits fragments.

12. Kit de séquençage d'ADN comprenant une ADN polymérase de la revendication 4 et une pyrophosphatase.

13. Méthode de préparation d'ADNc en fournissant une amorce oligonucléotidique, une polymérase de l'une quelconque des revendications 1 à 4 et entre un et quatre désoxyribonucléoside phosphates.

14. ADN polymérase purifiée de la revendication 1 ou la revendication 2, qui a une activité de transcriptase inverse.
15. ADN polymérase de la revendication 1 ou la revendication 2, utile comme transcriptase inverse.

16. Méthode de transcription inverse/réaction en chaîne par polymérase utilisant une ADN polymérase de la revendication 1 ou la revendication 2 et une polymérase convenant à la mise œuvre d'une réaction en chaîne par polymérase dans un seul récipient de réaction.

17. Méthode selon la revendication 16, dans laquelle ladite polymérase convenant à la mise œuvre d'une réaction en chaîne par polymérase est l'ADN polymérase Taq.

18. Kit pour transcription inverse/réaction en chaîne par polymérase comprenant une ADN polymérase de la revendication 1 ou la revendication 2 et une polymérase convenant à la mise en œuvre d'une réaction en chaîne par polymérase.

19. Solution destinée à être utilisée en RT/PCR, comprenant une ADN polymérase de la revendication 1 ou la revendication 2 et une polymérase convenant à la mise en œuvre d'une réaction en chaîne par polymérase.

20. Solution de la revendication 19, dans laquelle ladite polymérase convenant à la mise œuvre d'une réaction en chaîne par polymérase est l'ADN polymérase Taq.

21. Solution pour l'amplification d'ADN par déplacement de brin comprenant une ADN polymérase des revendications 1 à 4 et une enzyme de restriction thermostable.

22. Solution pour le séquençage d'ADN comprenant une ADN polymérase des revendications 1 à 4 et une pyrophosphatase.
CTGCAGAAAAAAAAATAATTTTTTTTTTTTTAGCCTCCTATAGAACTTCC
TGAAGATTTTTAATAAAATTACATCA
AAAAATTGTATTCAAAAAATTTTTTTAGAGAAAGAAAGGCTAAAAA
AGACGTATTTAAAAGTAAGTCTTTAGC
ACATCTTTAATTTATGACAGTATTTATAGTAATACATTTAATAATCCTTA
CAAAATTGAGTCTAATACTCTTTAGC
TAAACATTGAAATGGGACAGGTTATGGGCTTACCTCGAGAAAGAAGAGCCTATA
ATGAAAGAAGATGGCCAGCGAGAAACCT
CAAGGAAAAATCAAAAGATGCAACCTTCTTTGGAGATGAAAGAGCCTATA
GGATTATTGATTTAACGTCAGTTTTA
GTAAGACAGAAGGTTTTATTTGAAACCTTCTGAGGAAACAGTATCCAGTGA
TAATATCAAAAAAGGTAAAGTGTGTGT
GAAGGTGCGGGAAGAATAAATTGGAAGGATGTGATTTCCTAATTTAAAGATTGG
CAAGGATGACAGCCTTTAAGGATTGACAG
CAAAGACATTACAGACCAGTTATATGACCTTACAAGCAAGGCTAAAAATTTGG
AAAGTATGCAGCAGCAGATTCACTTTACG
TCATGAATAAGGCTACTACAGTGAAAGAATGCTTTGAAGGTGAAATCTGAAATA
AAGCAGGAAAGATGGAATAATTGAGTC
ATGGAAGGACAGCTAAAATATTTGGAAGAATCTTACGAGTCTGGCAACCATAAA
TCCTCAATAAAAGGATACATCTTTA
TGAGAAACCTTTGGCTATAGTTTTTCTTTTTAAGGGAATTGGTAAGGTATATAAG
AAGCATTTTAAATAATGAAATTTTT
TAAATTTTTTATGAAAAGTTTGATAATAATATTTGTTGATATAATTTTTCCTTATA
GTGTACTGGCTTTAATCGGTTATAAA
GTGTATATTCTTTTCAAAAAAGGAGATGCAATTTACATCTTTTTTTTTGC
GGTATATTTAAATAAAGAATAAAA
AATGGGAGTATTTTAATGTATATAATTTTATTTTATTAATATGAGAAGATTGCCTCATG

**FIG. 1A**
TACAGAGCCTATTTATGCCTGGCCCAT
GCTTACTACAAGTGAGGGATTCCTACAAATGCTCTGTATGTTTTTACTATGAT
GCTTATAAAAACCTTATGCAGGAGGAA
AACCTGATTACATAGCTATTGCTTTTGACAACAAAAAGCTCCTACTTTTAGACACA
AAGAATATCAAGACATAAAGCTACA
AGACAAGCTATGCCGAAAGAAGCTTGTGCTGAAACAGTGAAGCATATTGGAAAGAAAT
TATAGATGCTTTTATATATGAAACATT
AGAATTAGAAGGTTATGAGCTGATGACATTATAGGGACTATTCTAAAGGCTGG
CAGAGGAAAAAGGAATGGGAAGTGCTTG
TAGTTACAGGAGACAGAGATGCTCTTTCAATTAGTGTTCAGATAAAAGTGAAAGATA
AAAATTTTCTAAAAAGGCGTTATATTACTCAG
ATGGAAGAGTGGTTGACGAAAAGGGCTATTTTTAGAAAGGTATGGAATAACTCTCA
GCAGTTTATAGATTTAAAAAGGCTTAT
GGGAAGATAATCTGATAATATCTCTGGAGGTACCTAATATAGGGGAAAAACTG
CGATTAAGCTATTTAAAGGATTTGGGA
CAATTTGAAAAATTAAATCCAAAAATCTTTTCAGCTTAAAGGTAAAATAAAAAAGAA
AATATAGAAAAAACAATAAAAGGAGTTAGC
ATAATGAGTAAGAGGGCTTGCTACTATAAAAAAGAGACATTCCCATTGAGATAGA
TTTGGAGGACTATATAATGAAAAAAATT
TATGAGGAGAAGAAGCTTTTAGAGCTTTTTTTAACAAATAGAATTTCTTTAGTTTAAT
TGATAACATAAAAGGAAGTTGAGA
TAGAGATTGAGATAATCATAAAAGTTGAAATGTTCAAAGGTAGATATAAAA
GAATTAGTAACTTTTGGTGAAGATAAC
AGAAATATTGCTTTTTACCCTTAAATTATAGAAGGGGAAATAAAAAATAAGC
CTTTTCTTTGGAAAGGATACGCTTTA
TATTGACGGTTTTTCAAAACAGAAAGATTAAAAGGAATTTTTGGAAAGAGATT
TTGAATTACAACCCCATGAAATCAAAGG
ATTTTTGAGGCTTTTCTTTATAAAGGAAATAGAGTGTAAAGCAAGTACATA
GATACGTGCTGTAAATGGCTTTATCTCCTG
FIG. 1C
TTCTGAAATAAACTCAAGAAATTATTACTCAAAGATCGAGGCGAAAGGTTA
GCAATGAAATGCTCCGATACAGGGAAAGT
GCGGCTGATATAAAAAATGGCAATGTTAAGGTATACAACGATTAAAAAAA
ATTAAAGCTTAAGTCTAAAGCTTTATATT
GCAAGTTTAGCAGCAGCTTGTAGTGAGACTTATTAAGGATGAAGTAGATATCA
TAAAAAGATACCTTAAAGAAAATATGG
AAAAATGTAGTGCAATTTAAGGAAGTTTCTCTGGTTGAAAAATTGGGCTAGGGCT
AATTTGGTTTGGCCAAGTGAGGGG
CAAAAATGCAAGTAAGTGACTTACTGGCGGCATTGCCTACAAGTTAAAGTAGTT
GTATCCAAAGCTTATTAAAGAAAGATGGGA
GCTGTGTAATTGATGCAGATATTGTATCAAGAGAAATAATGTGAAAAGGAAC
AGAACGATATAATAGAAATTGTCAGAAATA
TTTTGGAAAAAGAAAAATTTTTAAAAGGAAGACCGCGAATTGACAGAAAGAAATTA
GGCAATATTGTGTTTGGCTGACAGAAAA
AGCTAAAAAATCATAAGGAATTACTCATCCCATATAAAATAGAGAGGATAAAAA
GAAGGAATAGAAGAAGAAAGAAAGAAAA
ATCAGCAGAAGGCAATTGTATTAGATGCAGCTTTTTAATTGAGATGAACTTTA
TAAATGCTAGATGAGGTTTTGGTAGT
AGTAGTAGACTCAAAAACTCAAAAAAAGAGGGTTATGGAGAGGGATAAGCTTT
CTTATAAGGATGCGATAAACCAGATAA
AAAGCACAATTGCTTTGGAGCAAAAAATGAAGTACGCAGATTTTTAATAAACA
TAGCAAAAGTTTTAAAGCTATGGAAAA
CAGTACCACATTATTTGGGAGATTGTTCAGAACGTTATTCATGATATCGGAGGA
GTTACTTTGAGAAAAAGTTTCGGCCA
ATTTAAATTACTATTAGTTTTAGTTTACTTACGAGTTAATACTCATTATTTTT
TAAAAAAATTTATCCCTAGTTAAATA
CTATAATTATGATTTTATTACGCAAAGAGTAAGTAGGCACTGACCCCTATCTTAT
TTTTGCTGTGATAAAGAGTGCAAGGA
ATTTTAGAGCGATGCAATTTCCAGCAAAAAATGCTAGAGGGTTATGCAAAATT

FIG. 1D
FIG. 1E
ATG TAT AAA TTT TTA ATA ATT GAT GGA AGT AGC CTC ATG TAC AGA GCC TAT TAT GCC TTG
M Y K F L I I D G S S L M Y R A Y Y A L
61/21
CCC ATG CTT ACT ACA AGT GAG GGA TTG CCT ACA AAT GCT CTG TAT GGT TTT ACT ATG ATG
P M L T T S E G L P T N A L Y G F T M M
121/41
CTT ATA AAA CTT ATC GAG GAG GAA AAA CCT GAT TAC ATA GCT ATT GCT TTT GAC AAA AAA
L I K L I I E E E K P D Y I A I A F D K K
181/61
GCT CCT ACT TTT AGA CAC AAA GAA TAT CAA GAC TAC AAA GCT ACA AGA CAA GCT ATG CCT
A P T F R H K E Y Q D Y K A T R Q A M P
241/81
GAA GAA CTT GCT GAA CAA GTA GAC TAT TTG AAA GAA ATT ATA GAT GGC TTT AAT ATA AAG
E E L A E Q V D Y L K E I I I D G F N I K
301/101
ACA TTA GAA TTA GAA GCT GAT GAC ATT ATA GGG ACT ATT TCA AAG CTG GCA
T L E L E G Y E A D D I I G T I S K L A
361/121
GAG GAA AAA GGA ATG GAA GTG CTT GTA GTT ACA GGA GAC AGA GAT GCT CTT CAA TTA GTT
E E K G M E V L V V T G D R D A L Q L V
421/141
TCA GAT AAA GTG AAG ATA AAA ATT TCT AAA AAG GGT ATT ACT CAG ATG GAA GAG TTT GAC
S D K V K I K I S K K G I T Q M E E F D
481/161
GAA AAG GCT ATT TTA GAA AGG TAT GGA ATA ACT CCT CAG CAG TTT ATA GAT TTA AAA GGG
E K A I L E R Y G I T P Q Q F I D L K G
541/181
CTT ATG GGA GAT AAA TCT GAT AAT ATC CCT GGA GTA CCT AAT ATA GGG GAA AAA ACT GCG
L M G D K S D N I P G V P N I G E K T A
601/201
ATT AAG CTA TTA AAG GAT TTT GGA ACA ATT GAA AAT TTA ATC CAA AAT CTT TCT CAG CTT
I K L L K D F G T I E N L I Q N L S Q L
661/221
AAA GGT AAA ATA AAA GAA AAT ATA GAA AAC AAT AAA GAG TTA GCT ATA ATG AGT AAG AGG

FIG. 2A
FIG. 2B
FIG. 2C
FIG. 2D
ATG AAA GTT GAA AAA TGG TCT AAA GTA GAC ATC AAA GAA TTA GTA ACT TTG TTG CAA GAT
M K V E K W S K V D I K E L V T L L Q D
61/21
AAC AGA AAT ATT GCT TTT TAC CCG TTA ATT TAT GAA GGG GAA ATA AAA AAA ATA GCC TTT
N R N I A F Y P L I Y E G E I K K I A F
121/41
TCT TTT GGA AAG GAT ACG GTT TAT ATT GAC GTT TTC CAA ACA GAA GAT TTA AAG GAG ATT
S F G K D T V Y I D V F Q T E D L K E I
181/61
TTT GAA AAA GAA GAT TTT GAA TTT ACA ACC CAT GAA ATA AAG GAT TTT TTA GTG AGG CTT
F E K E D F E F T T H E I K D F L V R L
241/81
TCT TAT AAA GGA ATA GAG TGT AAA AGC AAG TAC ATA GAT ACT GCT GTA ATG GCT TAT CTT
S Y K G I E C K S K Y I D T A V M A Y L
301/101
CTG AAT CCT TCT GAG TCT AAC TAT GAC TTA GAC CGT GTG CTA AAA AAA TAT TTA AAG GTA
L N P S E S N Y D L D R V L K K Y L K V
361/121
GAT GTG CCT TCT TAT GAA GGA ATA TTT GGC AAA GGT AGG GAT AAA AAG AAA ATT GAA GAG
D V P S Y E G I F G K G R D K K K I E E
421/141
ATT GAC GAA AACATA CTT GCT GAT TAT ATT TGC AGT AGA TGT GTG TAT CTA TTT GAT TTA
I D E N I L A D Y I C S R C V Y L F D L
481/161
AAA GAA AAG CTG AAT TTT ATT GAA GAG ATG GAT ATG AAA AAA AAG CTT CTA TTA GAA ATA
K E K L M N F I E E M D M K L L L L E I
541/181
GAA ATG CCT CTT GTA GAA GTT TTA AAA TCA ATG GAG GTA AGT GGT TTG ACA TTG GAT AAA
E M P L V E V L K S M E V S G F T L D K
601/201
GAA GTT CTA AAA GAG CTT TCA CAA AAG ATA GAT GAT AGA ATA GGA GAA ATA CTA GAT AAA
E V L K E L S Q K I D D R I G E I L D K
661/221
ATT TAT AAA GAG GCA GGA TAT CAA TTT AAT GTA AAT TCA CCT AAG CAA TTA AGT GAA TTT

FIG. 4A
FIG. 4B
FIG. 4C
ATGGCTTATAAATTTTTAATCATTGATGGTAGTAGCCTCATGTCACAGAG
CCTATTATGCCTTGGCCCATGCTTACTACAAGTGAGGATTGCCTACAAATGCTCTGTA
TGTTTTACTATGATGCTTTATA
AAACTTATCGAGGAGGAAAAACCTGATTACATAGCTATTGCTTTTGACAAAAAAGGT
CCTACTTTTAGACACAAAGATA
TCAAGACTACAAAGCTACAAGCAACAGCTATGCCTGAAGAACTTGCTGAACAAGTAGA
GCTTTAAATATAAGACATTAGAGAATTAGTTATGAAAGCTGTAGACATTATAGGGA
CTATTCTAAAGCTGGCAGAGGAA
AAAGGGAATGGAAAGTGGTTGATTAGGAGAACAGAGATGCTCTTCAATTAGTTTCA
GATAAAAGTAGAATAAAATTTTC
TAAAAAGGGTATTACTCACGATGGGAAGTGGTTGACGCAAAGGCTATTAGTTAGAAAT
TGGAATAACTCCTCAGCGAGTTTA
TAGATTAAAGGGCTTTATGGGAGATAAAATCTGATAATATCCCTGGAGTACCTAATAT
AGGGGAAAAACTGCGATTAAG
CTATTTAAGGATTTTGGAACATTGGAAATTTAATCCAAAATCTTTCTCAGCTTAAG
GTAAAAAGAAAAAGGAAATAGA
AAACAAATAAGAGTACTAGTATAGTAGAAGGCTTTGCTACTATAAAAAAGGACAT
TCCCCATGGAGATAATTTTGGG
AGTATAAAAGTAAAAATTTAATGAGGAGGAAAGCTTTTAGAGCTTTTTAAATAATTAGA
ATTCTTTAGTTTAAATGATAAC
ATAAAGAAAGAAGGTAGCATAGAGATTGTAGATAATCATAAAGTTGAAAAATGTTCA
AAAGTAGATAAAGAAATTAGT
AACTTTGTTGCAAGATAAAGAATATTGCTTTTTACCCGTTAATTTTAGAAGGGGA
ATAAAAAATAAGCCTTTTCTT

FIG. 7A
TTGAAAGGATACGGTTATATTGACGTTTCCAAACAGAAGATTAAAGGAGATTT
TGAAAAAGAAGATTTTGAATTT
ACACCCCATGAAATAAGGATTTTTTAGTGAAGCTTTTTTTTTAATAAGGATAGTG
AAAGCAAAGTACATAGATACTGC
TGTAATGGCTTTATCTTCTGAAATCTCTTCTGAGTCTAACTATGACCTAGACCTGCTAA
AAAAATTTAAAGGTAGATG
TGCCCTCTTATGAAAGGAATATTTGGCCAAGGTTAGGGATAAAAAAGAAATTTGAGGAGA
TTGACGAAAACATACTTGTGAT
TATATTTGCAGTAGATGTGTGTATCTATTTGATTAAAAAGGAAGCTGATGAAATTTTAT
TGAAAGAGATGGATATGAAAA
ACTTCTATTAGAAATAGAAATGCCTTCTGTAGAAAGTTTTAAATCAATGGAGGTAAGT
GGTTTTACATTGGATAAAAG
TTCTAAAAGAGCTTTTCACAACAAAGATAGATGATAGAAATAGAGAAATACTAGATAAAA
TTTATAAAGAGGGCAAGGATATCAA
TTTAATGTAAAATTCACCTAGCAATTTAGTGAATTTTTGGAAAAGTTAAACTTACC
AGTAATAAAGAAACAAAAAC
AGGAATCTCTACGGATTCTGAAGTTTTGGAACAATTTGTTCTTTATAATGATATTGTC
AGCGATATAATAGAGTATCGGC
AACTTACAAAACATTTAAATCTACATATTATAGATGGATTTTTGGCTTTATGGATGAAA
CAATAGAGTACTATTTCAATTT
AAACAAATGGTTACTGCATTACAGTAGAAATAAGCAGCACCCGACCCAAATCTACAAAT
ATACCTATAAGAGAAGGTTG
CAGACAAAAATGAGGGCTTTTTATGCGAGGAGTAGAGATGGATATTTGTTTCAGCA
GATTATCTCAGATTTGAACTGA
GGTTTTAGCAGACATGTTTCGGGAGATGAAAGCTAAATGAAATCTTTTATGAAATAATGA

FIG. 7B
AGATATACATTATAAGGACAGCT
TCGGAGGTTTTTAAAGTGCCTATGGAAGAAAGTTACACGGGAGATGAGAAAGAGCAGCA
AAAGCCGTAAATTGGGCTATAAT
ATATGGCATAAAGCGATTATGGGCTTTTCGAGACCTTAAATATCAAGAAAGAAGC
AAAAGAGTACATAAATAATTATT
TTGAAAGATATAAAAGGAGTAAAGAGATTATATGGAAAAATAGTACGATTTGCAAAG
AAAATGGCTATGTGACTACAAATA
ATGAACAGAAGAGATATTTCCTGAAATAAACACTCAAGAAATTTTACTCAAAAGATCG
CAGGCGAAAGGTTAGCAATGAA
TGCTCCGATACAGGAAGTGCGGCTGATATAATAAATTGCAATGTTAAAGGTATA
CAACGATTTAAAAAATTAAGC
TTAAGTCTAAGCCTTATATTGCAAGTTTCATGACGAGCTTTGTAAGTGGAATCTTATAAGGA
TGAAGTAGATATCATAAAAAAG
ATACTTTAAGAAAATAAGGAAAATGTAGTGCAATTAAAAAGTTTCTCTGGTTGGTAAA
TTGGCGTGAGGCCCTAATTGGTT
TTTGGCCAAGTG

FIG. 7C