



(19)  
Bundesrepublik Deutschland  
Deutsches Patent- und Markenamt

(10) **DE 697 30 926 T2** 2006.02.23

(12)

## Übersetzung der europäischen Patentschrift

(97) **EP 0 823 479 B1**

(21) Deutsches Aktenzeichen: **697 30 926.6**

(96) Europäisches Aktenzeichen: **97 113 182.6**

(96) Europäischer Anmeldetag: **31.07.1997**

(97) Erstveröffentlichung durch das EPA: **11.02.1998**

(97) Veröffentlichungstag

der Patenterteilung beim EPA: **29.09.2004**

(47) Veröffentlichungstag im Patentblatt: **23.02.2006**

(51) Int Cl.<sup>8</sup>: **C12N 15/54** (2006.01)

**C12N 9/12** (2006.01)

**C12N 15/10** (2006.01)

**C12Q 1/68** (2006.01)

(30) Unionspriorität:

**23376 P 06.08.1996 US**

(73) Patentinhaber:

**F. Hoffmann-La Roche AG, Basel, CH**

(74) Vertreter:

**Vossius & Partner, 81675 München**

(84) Benannte Vertragsstaaten:

**AT, BE, CH, DE, DK, ES, FR, GB, IT, LI, NL, SE**

(72) Erfinder:

**Gelfand, David Harrow, Oakland, US; Kalman, Lisa**

**Vivian, San Francisco, US; Reichert, Fred**

**Lawrence, Oakland, US**

(54) Bezeichnung: **Modifizierte thermostabile DNA Polymerase**

Anmerkung: Innerhalb von neun Monaten nach der Bekanntmachung des Hinweises auf die Erteilung des europäischen Patents kann jedermann beim Europäischen Patentamt gegen das erteilte europäische Patent Einspruch einlegen. Der Einspruch ist schriftlich einzureichen und zu begründen. Er gilt erst als eingelegt, wenn die Einspruchsgebühr entrichtet worden ist (Art. 99 (1) Europäisches Patentübereinkommen).

Die Übersetzung ist gemäß Artikel II § 3 Abs. 1 IntPatÜG 1991 vom Patentinhaber eingereicht worden. Sie wurde vom Deutschen Patent- und Markenamt inhaltlich nicht geprüft.

## Beschreibung

### Gebiet der Erfindung

**[0001]** Die vorliegende Erfindung betrifft thermostabile DNA-Polymerasen, die eine erhöhte Wirksamkeit für den Einbau von Ribonucleosidtriphosphaten aufweisen. Die Erfindung stellt Verfahren und Mittel zur Isolierung solcher Polymerasen bereit. Die erfindungsgemäßen Enzyme sind für viele Anwendungen und insbesondere für Anwendungen der Nucleinsäuresequenzierung nützlich. Folglich stellt die Erfindung auch verbesserte Verfahren für die Nucleinsäure-Sequenzanalyse bereit.

### Hintergrund der Erfindung

**[0002]** Die DNA-Sequenzierung erfordert im Allgemeinen die Erzeugung von vier Populationen aus einzelsträngigen DNA-Fragmenten mit einem definierten Terminus und einem variablen Terminus. Der variable Terminus endet im Allgemeinen an spezifischen Nucleotidbasen (entweder Guanin (G), Adenin (A), Thymin (T) oder Cytosin (C)). Die vier unterschiedlichen Sätze von Fragmenten werden jeweils auf der Basis ihrer Länge getrennt. Bei einem solchen Verfahren wird ein hochauflösendes Polyacrylamidgel verwendet. Jede Bande auf einem solchen Gel entspricht einem spezifischen Nucleotid in der DNA-Sequenz, wodurch die Positionen in der Sequenz identifiziert werden.

**[0003]** Ein häufig angewendetes DNA-Sequenzierungsverfahren ist das Didesoxy- oder Kettenterminations-Sequenzierungsverfahren, das die enzymatische Synthese eines DNA-Stranges beinhaltet (Sanger et al., Proc. Natl. Acad. Sci. USA. 74 (1977), 5463). Im Allgemeinen werden vier verschiedene Synthesen durchgeführt, wobei bewirkt wird, daß jede Reaktion durch den Einbau eines geeigneten Kettenterminationsnucleotids wie eines Didesoxynucleotids an einer spezifischen Base (G, A, T oder C) abbricht. Die Reaktionsprodukte sind einfach zu interpretieren, da jede Bahn nur entweder G, A, T oder C entspricht.

**[0004]** Bei dem Didesoxy-Kettenterminationsverfahren wird ein kurzer einzelsträngiger Primer an eine einzelsträngige Matrize angelagert. Der Primer wird an seinem 3'-Ende durch den Einbau von Desoxynucleotiden (dNTPs) verlängert, bis ein Didesoxynucleotid (ddNTP) eingebaut wird. Wenn ein ddNTP eingebaut wird, endet die Verlängerung an dieser Base. Um die Genauigkeit der DNA-Replikation sicherzustellen, weisen DNA-Polymerasen jedoch eine sehr starke Vorliebe für den Einbau ihrer normalen Substrate, d. h. dNTPs, und wider dem Einbau von Nucleotidanaloga, bezeichnet als unkonventionelle Nucleotide, auf. Im Falle der DNA-Synthese werden Ribonucleotide (rNTPs) als unkonventionelle Nucleotide angesehen, da rNTPs ähnlich wie ddNTPs nicht das normale in vivo-Substrat einer DNA-Polymerase sind. In der Zelle verringert diese Eigenschaft den Einbau von abnormalen Basen wie Desoxyinosintri-phosphat (dITP) oder rNTPs in einen wachsenden DNA-Strang.

**[0005]** Zwei häufig angewendete automatisierte Sequenzierungsmethodiken sind die Farbstoff-Primer- und Farbstoff-Terminator-Sequenzierung. Diese Verfahren sind zur Verwendung mit fluoreszenzmarkierten Einheiten geeignet. Obwohl die Sequenzierung auch unter Verwendung von radioaktiv markierten Einheiten durchgeführt werden kann, wird die auf Fluoreszenz basierende Sequenzierung immer mehr bevorzugt. Kurz zusammengefaßt wird bei der Farbstoff-Primer-Sequenzierung ein fluoreszenzmarkierter Primer in Kombination mit unmarkierten ddNTPs verwendet. Das Verfahren macht vier Synthesereaktionen und bis zu vier Bahnen auf einem Gel für jede Matrize, die sequenziert werden soll, erforderlich (eine entsprechend jedem der basenspezifischen Terminationsprodukte). Nach der Primerverlängerung werden die Sequenzierungsreaktionsgemische, enthaltend Terminationsprodukte mit einem eingebauten Didesoxynucleotid, routinemäßig mittels Elektrophorese auf einem DNA-Sequenzierungsgel analysiert. Nach der Trennung mittels Elektrophorese werden die fluoreszenzmarkierten Produkte mit einem Laser auf der Unterseite des Gels angeregt, und die Fluoreszenz wird mit einem geeigneten Monitor nachgewiesen. Bei automatisierten Systemen tastet ein Detektor die Unterseite des Gels während der Elektrophorese ab, um nachzuweisen, welche Markierungseinheit verwendet worden ist, wenn die Reaktionsgemische durch die Gelmatrix hindurch laufen (Smith et al., Nature 321 (1986), 674–679). Bei einer Modifikation dieses Verfahrens werden vier Primer jeweils mit einem anderen Fluoreszenzmarker markiert. Nachdem die vier getrennten Sequenzierungsreaktionen beendet sind, werden die Reaktionsgemische kombiniert, und die kombinierten Reaktionsgemische werden einer Gelanalyse in einer einzigen Bahn unterzogen, wodurch die unterschiedlichen Fluoreszenzmarker (einer entsprechend jedem der vier verschiedenen basenspezifischen Terminationsprodukte) einzeln nachgewiesen werden.

**[0006]** Alternativ werden Farbstoff-Terminator-Sequenzierungsverfahren angewendet. Bei diesem Verfahren wird eine DNA-Polymerase verwendet, um dNTPs und fluoreszenzmarkierte ddNTPs an dem wachsenden

Ende eines DNA-Primers einzubauen (Lee et al., *Nucleic Acids Research* 20 (1992), 2471). Dieses Verfahren bietet den Vorteil, daß keine farbstoffmarkierten Primer synthetisiert werden müssen. Außerdem sind Farbstoff-Terminator-Reaktionen insofern günstiger, als alle vier Reaktionen in dem gleichen Röhrchen durchgeführt werden können. Modifizierte thermostabile DNA-Polymerasen mit einer reduzierten Diskrimination gegen ddNTPs sind beschrieben worden (vgl. Europäische Patentanmeldung mit der Veröffentlichungsnummer EP-A-655,506). Eine beispielhafte modifizierte thermostabile DNA-Polymerase ist die mutierte Form der DNA-Polymerase aus *T. aquaticus* mit einem Tyrosinrest an Position 667 (anstelle eines Phenylalaninrestes), d. h. eine sogenannte F667Y-mutierte Form der Taq-DNA-Polymerase. AmpliTaq® FS, hergestellt durch Hoffmann-La Roche und vertrieben durch Perkin Elmer, reduziert die Menge an dem ddNTP, die für eine wirksame Nucleinsäuresequenzierung eines Zielmoleküls erforderlich ist, um das Hundert- bis Tausendfache. AmpliTaq® FS ist eine mutierte Form der DNA-Polymerase aus *T. aquaticus*, welche die F667Y-Mutation und zusätzlich einen Asparaginsäurerest an Position 46 (anstelle eines Glycinrestes; G46D-Mutation) aufweist.

**[0007]** Es besteht ein Bedarf an thermostabilen DNA-Polymerasen, die alternative Nucleinsäure-Syntheseverfahren für eine genaue und kosteneffektive Nucleinsäure-DNA-Sequenzanalyse ermöglichen. Auf Fluoreszenz basierende Verfahren, die nicht die Verwendung von Didesoxynucleotiden erforderlich machen, wären wünschenswert. Die vorliegende Erfindung zielt auf diese Bedürfnisse ab.

#### Zusammenfassung der Erfindung

**[0008]** Die vorliegende Erfindung stellt matrizenabhängige thermostabile DNA-Polymeraseenzyme bereit, die das Aminosäuresequenzmotiv SerGlnIleXaaLeuArgXaa (SEQ ID NR: 1) umfassen, wobei "Xaa" an Position 4 dieser Sequenz ein beliebiger Aminosäurerest, aber kein Glutaminsäurerest (Glu) ist, und wobei "Xaa" an Position 7 dieser Sequenz ein Valinrest (Val) oder ein Isoleucinrest (Ile) ist. Dargestellt im Ein-Buchstaben-Code für Aminosäuren kann dieses Sequenzmotiv wiedergegeben werden als S Q I X L R V/I, wobei "X" an Position 4 dieser Sequenz ein beliebiger Aminosäurerest, aber kein Glutaminsäurerest ist. Thermostabile DNA-Polymeraseenzyme mit einer Aminosäuresequenz, umfassend dieses Sequenzmotiv, wobei "X" an Position 4 dieses Sequenzmotivs kein Glutaminsäurerest ist, weisen eine reduzierte Diskrimination gegen den Einbau von Ribonucleotiden im Vergleich zu bereits bekannten thermostabilen Polymerasen auf. In einem wachsenden DNA-Strang sind Ribonucleotide unkonventionelle Nucleotide. Folglich, gemäß einem ersten Aspekt, bauen die neuen erfindungsgemäßen Enzyme unkonventionelle Basenanaloga wie Ribonucleotide in einen wachsenden DNA-Strang mit einer um mehrere Größenordnungen höheren Wirksamkeit als bereits identifizierte thermostabile DNA-Syntheseenzyme ein. Gene, die diese Enzyme codieren, werden auch durch die vorliegende Erfindung bereitgestellt, sowie rekombinante Expressionsvektoren und solche Vektoren umfassende Wirtszellen. Mit solchen transformierten Wirtszellen können große Mengen an gereinigten thermostabilen Polymeraseenzymen bereitgestellt werden.

**[0009]** Durch die vorliegende Erfindung wird eine Region oder ein Sequenzmotiv innerhalb der Aminosäuresequenz von thermostabilen DNA-Polymerasen identifiziert, die/das die Wirksamkeit der Fähigkeit der Polymerase zum Einbau von Ribonucleotiden erhöht, während die Fähigkeit zum genauen Einbau von Desoxyribonucleotiden beibehalten wird. Veränderungen in dieser Region, z. B. ein oder mehrere Aminosäureaustausche (z. B. eingeführt durch ortsspezifische Mutagenese), stellen ein thermostabiles Polymeraseenzym bereit, das in der Lage ist, einen RNA- oder RNA-/DNA-Chimären- oder Hybrid-Strang an einer DNA-Matrize zu synthetisieren.

**[0010]** Gemäß einem anderen Aspekt stellt die Erfindung verbesserte Verfahren und Zusammensetzungen zur Bestimmung der Sequenz einer Zielnucleinsäure bereit, wodurch die Notwendigkeit für Kettenterminations-ddNTPs beseitigt wird. Durch die hierin bereitgestellten verbesserten Verfahren werden Ribonucleotide (rNTPs) in Primerverlängerungsprodukte eingebaut. Da die erfindungsgemäßen Enzyme fehlerlos und effizient entweder rNTPs oder dNTPs einbauen, können Sequenzierungsreaktionen Gemische aus beiden Nucleotiden verwenden. Nach der Primerverlängerung können die neu synthetisierten Oligonucleotidprodukte an den eingebauten rNTPs durch auf dem Fachgebiet bekannte Verfahren, z. B. durch Hydrolyse, gespalten werden, wodurch eine Population von Fragmenten bereitgestellt wird, die zur Fraktionierung und Sequenzanalyse durch herkömmliche Mittel wie eine Gelelektrophorese geeignet ist. Diese Verfahren verwenden die hierin bereitgestellten neuen thermostabilen Polymeraseenzyme. Folglich stellt die Erfindung gemäß diesem Aspekt thermostabile DNA-Polymeraseenzyme bereit, die dadurch gekennzeichnet sind, daß die Polymerase das entscheidende Motiv SerGlnIleXaaLeuArgXaa (SEQ ID NR: 1) umfaßt, wobei "Xaa" an Position 4 ein beliebiger Aminosäurerest sein kann, aber kein Glutaminsäurerest (Glu) ist, und wobei "Xaa" an Position 7 ein Valinrest (Val) oder ein Isoleucinrest (Ile) ist.

**[0011]** Gemäß einem anderen Aspekt der Erfindung stellen die hierin beschriebenen modifizierten Polymerasen ein Mittel zum Einbau von Ribonucleotiden oder Analoga bereit, die eine Hydroxylgruppe oder eine andere Substitution an der 2'-Position enthalten, die im Vergleich dazu nicht in konventionellen Desoxyribonucleotiden vorkommt. Diese Nucleotide können unterschiedlich markiert sein, wodurch Alternativen zur herkömmlichen Verwendung von Didesoxyribonucleotiden für Anwendungen der DNA-Sequenzierung bereitgestellt werden.

**[0012]** Die mutierten thermostabilen Polymeraseenzyme der Erfindung sind durch die Fähigkeit gekennzeichnet, daß sie unkonventionelle Nucleotide, besonders Ribonucleotide, wirksamer einbauen als die entsprechenden Wildtyp-Enzyme. In einer bevorzugten Ausführungsform der Erfindung kann das einzubauende unkonventionelle Nucleotid ein Kettenterminations-Basenanalogon wie 2'-Hydroxy-3'-desoxy-ATP (Cordycepintriphosphat), ein "Riboterminator"-Analogon von ATP, oder ein Nicht-Kettenterminationsnucleotid wie ein rNTP sein.

**[0013]** Gemäß einem anderen Aspekt der Erfindung werden mutierte thermostabile Polymeraseenzyme bereitgestellt, die durch die Fähigkeit gekennzeichnet sind, daß sie unkonventionelle Nucleotide, besonders Ribonucleotide, wirksamer einbauen als die entsprechenden Wildtyp-Enzyme. Folglich stellt die Erfindung gemäß diesem Aspekt rekombinante thermostabile DNA-Polymeraseenzyme bereit, die jeweils dadurch gekennzeichnet sind, daß (a) die Polymerase in ihrer nativen Form die Aminosäuresequenz SerGlnIleGluLeuArgXaa (SEQ ID NR: 2) umfaßt, wobei "Xaa" an Position 7 dieser Sequenz ein Valinrest (Val) oder ein Isoleucinrest (Ile) ist; (b) die Aminosäuresequenz in dem rekombinanten Enzym mutiert ist, vorzugsweise an Position 4 dieser Sequenz, so daß der Glutaminsäurerest an Position 4 ein anderer Aminosäurerest ist, vorzugsweise ein Glycinrest; und (c) das rekombinante Enzym eine reduzierte Diskrimination gegen den Einbau von Ribonucleotiden und Ribonucleotidanaloga im Vergleich zu der nativen Form des Enzyms aufweist.

**[0014]** Gemäß einem anderen Aspekt der Erfindung stellen die erfindungsgemäßen Polymerasen ein geeignetes Mittel zur Fragmentierung von Amplifikationsprodukten und Primerverlängerungsprodukten bereit, wobei solche fragmentierten Produkte bei auf Hybridisierung basierenden Methodiken und einer Vielzahl von Strategien zum Sequenznachweis nützlich sein können.

**[0015]** Die erfindungsgemäßen Enzyme und die sie codierenden Gene können verwendet werden, um Zusammensetzungen zur Verwendung in DNA-Sequenzierungsreaktionen bereitzustellen, die ein Gemisch aus konventionellen Nucleotiden und mindestens einem Ribonucleotid oder Ribonucleotidanalogen umfassen. In einer bevorzugten Ausführungsform der Erfindung ist das unkonventionelle Nucleotid ein Ribonucleotid und ist die Ribonucleotidkonzentration geringer als die Konzentration des entsprechenden Desoxyribonucleotids, d. h. das rNTP : dNTP-Verhältnis beträgt 1 : 1 oder weniger. Die erfindungsgemäßen Enzyme sind auch zur Kommerzialisierung in Kitanordnungen geeignet, wobei die Kits auch beliebige der folgenden zusätzlichen Elemente enthalten können, die für eine Nucleinsäure-Sequenzierungsreaktion notwendig sind, wie z. B. dNTPs, rNTPs, Puffer und/oder Primer.

#### Ausführliche Beschreibung der Erfindung

**[0016]** Die vorliegende Erfindung stellt neue und verbesserte modifizierte thermostabile DNA-Polymeraseenzyme, Zusammensetzungen und Kits, wie in dem beigefügten Satz von Patentansprüchen definiert, bereit. Die erfindungsgemäßen Enzyme bauen unkonventionelle Nucleosidtriphosphate wirksamer ein als die bereits bekannten Polymerasen oder die entsprechenden Wildtyp-Polymeraseenzyme, aus welchen diese neuen Polymerasen abgeleitet sind. DNA-Sequenzen, die diese modifizierten Enzyme codieren, Vektoren zur Expression der modifizierten Enzyme und mit solchen Vektoren transformierten Zellen werden ebenfalls bereitgestellt. Die erfindungsgemäßen Enzyme ermöglichen die Anwendung von neuen DNA-Sequenzierungsverfahren, die gegenüber den gemäß dem Stand der Technik bekannten DNA-Sequenzierungsverfahren vorteilhaft sind.

**[0017]** Um die Erfindung verständlicher zu machen, wird nachstehend eine Reihe von Begriffen definiert.

**[0018]** Der Begriff "konventionell" bei der Bezugnahme auf Nucleinsäurebasen, Nucleosidtriphosphate oder Nucleotide verweist auf solche, die natürlicherweise in dem beschriebenen Polynucleotid vorkommen (d. h. für DNA sind diese dATP, dGTP, dCTP und dTTP). Zusätzlich werden c7dGTP und dITP häufig anstelle von dGTP (obwohl mit einer geringeren Wirksamkeit eingebaut) in in vitro-DNA-Synthesereaktionen wie der Sequenzierung verwendet. Zusammengenommen können diese als Desoxyribonucleosidtriphosphate (dNTPs) bezeichnet werden.

**[0019]** Der Begriff "Expressionssystem" verweist auf DNA-Sequenzen, enthaltend eine gewünschte codierende Sequenz und Kontrollsequenzen in funktioneller Verbindung, so daß die mit diesen Sequenzen transfor-

mierten Wirte in der Lage sind, die codierten Proteine herzustellen. Zum Bewirken der Transformation kann das Expressionssystem in einem Vektor eingeschlossen sein; jedoch kann die relevante DNA auch in das Wirtschromosom integriert werden.

**[0020]** Der Begriff "Gen" verweist auf eine DNA-Sequenz, die Kontroll- und codierende Sequenzen umfaßt, welche zur Herstellung eines gewinnbaren bioaktiven Polypeptids oder eines Vorläufers notwendig sind. Das Polypeptid kann durch eine Gensequenz vollständiger Länge oder durch einen beliebigen Teil der codierenden Sequenz codiert werden, so lange die enzymatische Aktivität beibehalten wird.

**[0021]** Der Begriff "Wirtszelle(n)" verweist auf sowohl einzellige Prokaryonten- als auch Eukaryonten-Organismen wie Bakterien, Hefe und Actinomyceten sowie einzelne Zellen aus höheren Pflanzen oder Tieren, wenn in einer Zellkultur gezüchtet.

**[0022]** So wie hierin verwendet, verweist der Begriff "DNA-Sequenzierungsreaktionsgemisch" auf ein Reaktionsgemisch, das Elemente umfaßt, die für eine DNA-Sequenzierungsreaktion notwendig sind. Folglich ist ein DNA-Sequenzierungsreaktionsgemisch zur Verwendung in einem DNA-Sequenzierungsverfahren zur Bestimmung der Nucleinsäuresequenz eines Zielmoleküls geeignet, obwohl das Reaktionsgemisch anfangs unvollständig sein kann, so daß die Initiation der Sequenzierungsreaktion durch den Anwender kontrolliert wird. Auf diese Weise kann die Reaktion initiiert werden, sobald ein letztes Element wie das Enzym zugegeben wird, um ein vollständiges DNA-Sequenzierungsreaktionsgemisch bereitzustellen. Typischerweise enthält eine DNA-Sequenzierungsreaktion einen Puffer, der für eine Polymerisierungsaktivität geeignet ist, Nucleosidtriphosphate und mindestens ein unkonventionelles Nucleotid. Das Reaktionsgemisch kann auch einen Primer, der zur Verlängerung an einem Zielmolekül durch ein Polymeraseenzym geeignet ist, eine Polymerase und eine Zielnucleinsäure enthalten. Im Allgemeinen ist entweder der Primer oder eines der Nucleotide mit einer nachweisbaren Einheit wie einem Fluoreszenzmarker markiert. Allgemein ist der Reaktionsansatz ein Gemisch, das vier konventionelle Nucleotide und mindestens ein unkonventionelles Nucleotid umfaßt. In einer bevorzugten Ausführungsform der Erfindung ist die Polymerase eine thermostabile DNA-Polymerase und ist das unkonventionelle Nucleotid ein Ribonucleotid.

**[0023]** Der Begriff "Oligonucleotid", so wie hierin verwendet, ist definiert als ein Molekül, das aus zwei oder mehreren Desoxyribonucleotiden oder Ribonucleotiden, vorzugsweise mehr als drei und üblicherweise mehr als zehn, besteht. Die genaue Größe eines Oligonucleotids hängt von vielen Faktoren einschließlich der endgültigen Funktion oder der Verwendung des Oligonucleotids ab.

**[0024]** Oligonucleotide können durch irgendein geeignetes Verfahren hergestellt werden, einschließlich zum Beispiel der Clonierung und Restriktion geeigneter Sequenzen und der direkten chemischen Synthese durch ein Verfahren wie das Phosphotriesterverfahren von Narang et al., Meth. Enzymol. 68 (1979), 90–99; das Phosphodiesterverfahren von Brown et al., Meth. Enzymol. 68 (1979), 109–151; das Diethylphosphoramiditverfahren von Beaucage et al., Tetrahedron Lett. 22 (1981), 1859–1862; das Triesterverfahren von Matteucci et al., J. Am. Chem. Soc. 103 (1981), 3185–3191; automatisierte Syntheseverfahren; oder das Festphasenverfahren von US-Patent Nr. 4,458,066.

**[0025]** Der Begriff "Primer", so wie hierin verwendet, verweist auf ein Oligonucleotid, ob natürlich oder synthetisch, das in der Lage ist, als eine Initiationsstelle der Synthese wirksam zu sein, wenn es unter Bedingungen gebracht wird, bei denen eine Primerverlängerung initiiert wird. Ein Primer ist vorzugsweise ein einzelsträngiges Oligodesoxyribonucleotid. Die geeignete Länge eines Primers hängt von der beabsichtigten Verwendung des Primers ab, aber liegt typischerweise im Bereich von 15 bis 35 Nucleotiden. Kurze Primermoleküle machen im Allgemeinen kühlere Temperaturen erforderlich, um ausreichend stabile Hybridkomplexe mit der Matrize zu bilden. Ein Primer muß nicht die genaue Sequenz der Matrize wiedergeben, aber muß ausreichend komplementär sein, um mit einer Matrize zu hybridisieren, damit die Primerverlängerung erfolgt.

**[0026]** Ein Primer kann, falls gewünscht, durch den Einbau eines Markers, der durch spektroskopische, photochemische, biochemische, immunochemische oder chemische Mittel nachweisbar ist, markiert sein. Zum Beispiel schließen nützliche Marker <sup>32</sup>P, Fluoreszenzfarbstoffe, elektronendichte Reagenzien, Enzyme (wie sie gewöhnlich in ELISAs verwendet werden), Biotin oder Haptene und Proteine, für welche Antisera oder monoclonale Antikörper erhältlich sind, ein.

**[0027]** Der Begriff "thermostabile Polymerase" verweist auf ein Enzym, das gegen Hitze stabil ist, hitzeresistent ist und eine ausreichende Aktivität beibehält, um aufeinanderfolgende Primerverlängerungsreaktionen hervorzurufen, wenn es den erhöhten Temperaturen für den Zeitraum ausgesetzt wird, der notwendig ist, um

die Denaturierung von doppelsträngigen Nucleinsäuren zu bewirken. So wie hierin verwendet, ist eine thermostabile Polymerase zur Verwendung in einer Temperaturwechselreaktion wie der Polymerasekettenreaktion (PCR) geeignet. Für eine thermostabile Polymerase verweist die enzymatische Aktivität auf die Katalyse der Kombination der Nucleotide in der geeigneten Weise, um Primerverlängerungsprodukte zu bilden, die zu einem Matrizen-Nucleinsäurestrang komplementär sind.

**[0028]** Die zur Nucleinsäure-Denaturierung notwendigen Erwärmungsbedingungen hängen z. B. von der Pufferkonzentration und der Zusammensetzung und Länge der denaturierten Nucleinsäuren ab, aber liegen typischerweise im Bereich von etwa 90°C bis etwa 105°C, vorzugsweise 90°C bis 100°C, für einen Zeitraum, vorwiegend abhängig von der Temperatur und der Länge der Nucleinsäure, von typischerweise wenigen Sekunden bis zu vier Minuten.

**[0029]** Der Begriff "unkonventionell" oder "modifiziert" bei der Bezugnahme auf eine Nucleinsäurebase, ein Nucleosidtriphosphat oder ein Nucleotid schließt Modifikationen, Derivate oder Analoga von konventionellen Basen oder Nucleotiden, die natürlicherweise in der DNA vorkommen, ein. Genauer, so wie hierin verwendet, sind unkonventionelle Nucleotide an der 2'-Position des Ribosezuckers im Vergleich zu konventionellen dNTPs modifiziert. Folglich, obwohl für RNA die in der Natur vorkommenden Nucleotide Ribonucleotide sind (d. h. ATP, GTP, CTP und UTP, zusammengenommen rNTPs), sind Ribonucleotide unkonventionelle Nucleotide, da diese Nucleotide eine Hydroxylgruppe an der 2'-Position des Zuckers aufweisen, die im Vergleich dazu in dNTPs, so wie hierin verwendet, nicht vorkommt. Ribonucleotidanaloga, die Substitutionen an der 2'-Position enthalten, wie 2'-Fluor- oder 2'-Amino-substituierte Analoga, sind im Schutzzumfang der Erfindung eingeschlossen. Zusätzlich können Ribonucleotidanaloga an der 3'-Position modifiziert sein, zum Beispiel durch Austausch der normalen Hydroxylgruppe durch ein Wasserstoffatom (3'-Desoxy), wodurch ein Ribonucleotidanaloga-Terminator bereitgestellt wird. Solche Nucleotide sind auch im Schutzzumfang des Begriffes "unkonventionelle Nucleotide" eingeschlossen.

**[0030]** Da DNA herkömmlicherweise aus dNTPs besteht, wäre der Einbau eines rNTP unkonventionell und somit wäre ein rNTP eine unkonventionelle Base. Folglich enthalten in einer bevorzugten Ausführungsform der Erfindung für DNA-Primerverlängerungsverfahren einschließlich DNA-Sequenzierungsverfahren die Nucleinsäureprodukte sowohl konventionelle als auch unkonventionelle Nucleotide und umfassen vorwiegend konventionelle Nucleotide, die dNTPs sind.

**[0031]** Unkonventionelle Basen können fluoreszenzmarkiert, zum Beispiel mit Fluorescein oder Rhodamin; nicht-fluoreszenzmarkiert, zum Beispiel mit Biotin; isotopmarkiert, zum Beispiel mit <sup>32</sup>P, <sup>33</sup>P oder <sup>35</sup>S; oder unmarkiert sein.

**[0032]** Um die Erfindung noch verständlicher zu machen, wird in der Beschreibung auf spezifische thermostabile DNA-Polymeraseenzyme verwiesen, um die Erfindung zu veranschaulichen; jedoch sollen diese Referenzen den Schutzzumfang der Erfindung nicht begrenzen. In einer bevorzugten Ausführungsform werden die erfindungsgemäßen thermostabilen Enzyme in einer Vielzahl von Nucleinsäure-Sequenzierungsverfahren verwendet, obwohl die hierin beschriebenen neuen thermostabilen Polymerasen für beliebige Zwecke, wobei eine solche Enzymaktivität notwendig oder erwünscht ist, verwendet werden können. Das Enzym kann auch in Amplifikationsreaktionen wie der PCR verwendet werden.

**[0033]** Die erfindungsgemäßen thermostabilen Polymerasen sind dadurch gekennzeichnet, daß sie jeweils das entscheidende Motiv SerGlnIleXaaLeuArgXaa (SEQ ID NR: 1) enthalten, wobei "Xaa" an Position 4 dieser Sequenz ein beliebiger Aminosäurerest, aber kein Glutaminsäurerest (Glu) ist, und wobei "Xaa" an Position 7 dieser Sequenz ein Valinrest (Val) oder ein Isoleucinrest (Ile) ist. Gene, die thermostabile Polymerasen codieren, welche einen Glutaminsäurerest an der Position 4 dieses Motivs aufweisen, können modifiziert werden, wie hierin beschrieben, um geeignet modifizierte Polymeraseenzyme bereitzustellen. Diese modifizierten thermostabilen Polymeraseenzyme sind dadurch gekennzeichnet, daß sie im Vergleich zu den entsprechenden nativen oder Wildtyp-Enzymen eine Modifikation in dem Aminosäuresequenzmotiv SerGlnIleGluLeuArgXaa (SEQ ID NR: 2) aufweisen, wobei "Xaa" an Position 7 dieser Sequenz ein Valinrest (Val) oder ein Isoleucinrest (Ile) ist, d. h. das Motiv ist durch einen Austausch des Glutaminsäurerestes an der Position 4 durch einen anderen Aminosäurerest modifiziert worden. Das entscheidende Motiv einer erfindungsgemäß bereitgestellten thermostabilen DNA-Polymerase ist nachstehend unter Verwendung des herkömmlichen Drei-Buchstaben-Aminosäurecodes (Lehninger, Biochemistry, New York, New York, Worth Publishers Inc., 1970, Seite 67) angegeben.

SEQ ID NR: 1 SerGlnIleXaaLeuArgXaa,

wobei "Xaa" an Position 4 ein beliebiger Aminosäurerest, aber kein Glutaminsäurerest (Glu) ist, und wobei

"Xaa" an Position 7 ein Valinrest (Val) oder ein Isoleucinrest (Ile) ist.

**[0034]** Sowohl codierende Gensequenzen als auch Proteine, enthaltend diese entscheidende Aminosäuresequenz, wobei "Xaa" an Position 4 kein Glutaminsäurerest (Glu) ist, stellen eine Polymerase mit einer verminderten Diskrimination gegen rNTPs bereit und sind im Schutzzumfang der Erfindung eingeschlossen. Innerhalb des entscheidenden Motivs können weitere Modifikationen im Hinblick auf andere Aminosäurereste in diesem entscheidenden Motiv, vorzugsweise im Hinblick auf einen Aminosäurerest, gewählt aus der Gruppe aus Glutamin (Gln oder Q), Leucin (Leu oder L) oder Arginin (Arg oder R), durchgeführt werden.

**[0035]** Die vorliegende Erfindung ist zur Herstellung von thermostabilen DNA-Polymeraseenzymen mit vorteilhaften Eigenschaften durch eine besondere Modifikation der eine thermostabile DNA-Polymerase codierenden Gensequenz geeignet. In einer bevorzugten Ausführungsform der Erfindung sind die Gensequenz und das codierte Enzym aus einer Spezies der Gattung *Thermus* abgeleitet, obwohl nicht-*Thermus*-Eubakterien im Schutzzumfang der Erfindung eingeschlossen sind, wie nachstehend ausführlich beschrieben wird. Analog können neue thermostabile DNA-Polymerasen im Hinblick auf die stark konservierte Natur des nun identifizierten entscheidenden Motivs beispielsweise basierend auf ihrer Homologie zu Taq-Polymerase identifiziert werden. Solche thermostabilen Polymerasen sind im Schutzzumfang der vorliegenden Erfindung eingeschlossen, so lange ihre Aminosäuresequenz das S Q I X L R V/I-Motiv umfaßt, worin X ein beliebiger Aminosäurerest, aber kein Glutaminsäurerest ist, und wobei die Aminosäuresequenz mindestens etwa 39%, vorzugsweise mindestens etwa 60%, mehr bevorzugt mindestens etwa 80%, Gesamthomologie (Sequenzidentität) im Vergleich zu der Aminosäuresequenz der nativen Taq-Polymerase zeigt. Die vollständige Sequenz der Taq-Polymerase ist in WO 89/06691 bereitgestellt und unter der Hinterlegungsnummer P90556 in der GENESEQ-Patent-Sequenzdatenbank oder unter der Hinterlegungsnummer M26480 in der EMBL-Sequenzdatenbank und unter der Hinterlegungsnummer A33530 in der PIR-Sequenzdatenbank zugänglich.

**[0036]** Beispielhafte erfindungsgemäße thermostabile DNA-Polymerasen sind rekombinante Derivate der nativen Polymerasen aus den nachstehend in Tabelle 1 aufgeführten Organismen. Tabelle 1 gibt die besondere Sequenz des entscheidenden Motivs und die Position des "X"-Restes für jede dieser nativen Polymerasen an. Da jede thermostabile DNA-Polymerase einzigartig ist, ist die Aminosäureposition des entscheidenden Motivs für jedes Enzym verschieden. Für die nachstehend aufgeführten Polymerasen ist der Aminosäurerest in der "X"-Position des entscheidenden S Q I X L R V/I-Motivs ein Glutaminsäurerest. Die bevorzugten erfindungsgemäßen Polymerasen weisen ein Molekulargewicht im Bereich von 85.000 bis 105.000, mehr bevorzugt zwischen 90.000 und 95.000, auf. Die Aminosäuresequenz dieser Polymerasen besteht aus etwa 750 bis 950 Aminosäureresten, vorzugsweise zwischen 800 und 850 Aminosäureresten. Die erfindungsgemäßen Polymerasen können auch aus etwa 540 oder mehr Aminosäuren bestehen und umfassen mindestens die Polymerasedomäne und einen Teil, welcher der 3'-5'-Exonucleasedomäne entspricht (die erhaltene Polymerase kann eine 3'-5'-Exonucleaseaktivität aufweisen oder nicht) und möglicherweise Teile der 5'-3'-Exonucleasedomäne, die im ersten Drittel der Aminosäuresequenz von vielen thermostabilen Polymeraseenzymen vollständiger Länge enthalten ist.

**[0037]** Für thermostabile DNA-Polymerasen, die nicht in Tabelle 1 aufgeführt sind, ist die Identifizierung des geeigneten Glutaminsäurerestes für die Modifikation einfach, sobald das entscheidende Motiv oder Konsensusmotiv in der Aminosäuresequenz identifiziert ist.

**[0038]** Unabhängig von der genauen Position innerhalb einer thermostabilen DNA-Polymerase dient der Austausch des Glutaminsäurerestes (Glu) durch einen anderen Aminosäurerest innerhalb des Sequenzmotivs SerGlnIleGluLeuArgXaa (SEQ ID NR: 2), wobei "Xaa" an Position 7 dieser Sequenz ein Valinrest (Val) oder ein Isoleucinrest (Ile) der Polymerasedomäne ist, dazu, thermostabile Polymerasen bereitzustellen, welche die Fähigkeit besitzen, unkonventionelle Nucleotide effizient einzubauen. In einer bevorzugten Ausführungsform ist der Glutaminsäurerest durch eine Aminosäure mit einem ungeladenen polaren R-Rest wie Glycin, Serin, Cystein oder Threonin oder durch eine Aminosäure mit einem kleinen unpolaren R-Rest wie z. B. Alanin ersetzt. In einer besonders bevorzugten Ausführungsform ist der Glutaminsäurerest durch einen Glycinrest (G) ersetzt. Aminosäure- und Nucleinsäuresequenz-Abgleichungsprogramme sind ohne weiteres von der Genetics Computer Group, 575 Science Drive, Madison, Wisconsin, erhältlich. Unter Angabe des hierin identifizierten besonderen Motivs dienen diese Programme einschließlich zum Beispiel "GAP", "BESTFIT" und "PILEUP" dazu, die Identifizierung der genauen Sequenzregion, die modifiziert werden soll, zu unterstützen.

**[0039]** Wie aus der nachstehenden Tabelle 1 erkennbar ist, gibt es im Wesentlichen zwei Formen des konservierten Sequenzmotivs SerGlnIleGluLeuArgXaa (SEQ ID NR: 2) innerhalb der Polymerasedomäne der thermostabilen DNA-Polymeraseenzyme aus thermophilen Organismen. Das Sequenzmotiv SerGlnIleGluLeu-

ArgVal (SEQ ID NR: 3) ist in den nativen thermostabilen Polymerasen aus *Thermus*-Spezies, wie z. B. aus *Thermus aquaticus*, *Thermus caldophilus*, *Thermus thermophilus*, *Thermus flavus* und aus *Thermus filiformis* sowie aus den *Thermus*-Spezies sps17 und Z05 vorhanden. Das Sequenzmotiv SerGlnIleGluLeuArgVal (SEQ ID NR: 3) ist auch in der Polymerasedomäne anderer thermostabiler DNA-Polymeraseenzyme, z. B. aus *Thermosipho africanus* und aus verschiedenen *Bacillus*-Stämmen wie *Bacillus caldotenax* und *Bacillus stearothermophilus*, vorhanden. Das Sequenzmotiv SerGlnIleGluLeuArgIle (SEQ ID NR: 4) kommt z. B. in den nativen thermostabilen Polymerasen aus *Thermotoga maritima*, *Thermotoga neapolitana* und *Anaerocellum thermophilum* vor.

Tabelle 1

<u>Organismus</u>	<u>Aminosäure-Konsensusmotiv</u>	<u>Position des Glutaminsäurerestes</u>
	S Q I X L R V I	
<u><i>Thermus aquaticus</i> (Taq)</u>	S Q I E L R V	615
<u><i>Thermus caldophilus</i> (Tca)</u>	S Q I E L R V	617
<u><i>Thermus thermophilus</i> (Tth)</u>	S Q I E L R V	617
<u><i>Thermus flavus</i> (Tfl)</u>	S Q I E L R V	616
<u><i>Thermus filiformis</i> (Tfi)</u>	S Q I E L R V	613
<u><i>Thermus</i> -Spezies sps17</u>	S Q I E L R V	613
<u><i>Thermus</i> -Spezies Z05</u>	S Q I E L R V	617
<u><i>Thermotoga maritima</i> (Tma)</u>	S Q I E L R I	678
<u><i>Thermotoga neapolitana</i> (Tne)</u>	S Q I E L R I	678
<u><i>Thermosipho africanus</i> (Taf)</u>	S Q I E L R V	677
<u><i>Anaerocellum thermophilum</i> (Ath)</u>	S Q I E L R I	632
<u><i>Bacillus caldotenax</i> (Bca)</u>	S Q I E L R V	659
<u><i>Bacillus stearothermophilus</i> (Bst)</u>	S Q I E L R V	658, 661 oder 736*

\*Abhängig von der gewählten Aminosäuresequenz (vgl. nachstehend).

[0040] Die vollständigen Nucleinsäure- und Aminosäuresequenzen für jede der Taq-, Tth-, Z05-, sps17-, Tma- und Taf-Polymerasen sind in US-Patent Nr. 5,466,591 veröffentlicht worden und sind unter Bezugnahme herein eingeschlossen. Die Sequenzen für die DNA-Polymerase aus Tca, Tfl, Tne, Ath, Bca und Bst sind wie folgt veröffentlicht worden: Tca in der EMBL-Sequenzdatenbank unter der Hinterlegungsnummer U62584 (vgl. auch Kwon, Mol. Cells 7 (2) (1997), 264–271); Tfl bei Akhmetzjanov und Vakhitov, Nucleic Acids Research 20 (21) (1992), 5839; Tne in WO 97/09451 und WO 96/41014; Ath in der EMBL-Sequenzdatenbank unter der Hinterlegungsnummer X98575 (für Einzelheiten über den Ath-Stamm vgl. Rainey et al., J. Bacteriol. 175 (15) (1993), 4772–2779; Bst bei Uemori et al., J. Biochem. 113 (1993), 401–410, und unter der EMBL-Sequenzdatenbank-Hinterlegungsnummer U23149 (vgl. auch Phang et al., Gene 163 (1995), 65–68). Bst-Polymerase-Aminosäuresequenzen, umfassend ein E in dem entscheidenden Motiv an Position 658 werden auch durch die Japanische Patentveröffentlichung JP 05/304 964 A, die Europäische Patentveröffentlichung Nr. EP-A-699,760 und Aliotta et al., Genet. Anal. 12 (1996), 185–195, bereitgestellt; die Sequenz ist auch aus der EMBL-Sequenzdatenbank unter der Hinterlegungsnummer U33536 erhältlich. Die Sequenz, so wie in Gene 163 (1995), 65–68, veröffentlicht, enthält das "E" des entscheidenden Motivs an dem Rest Nummer 661. Bca bei Uemori et al., J. Biochem. 113 (1993), 401–410, und unter der EMBL-Sequenzdatenbank-Hinterlegungsnummer D12982. Die thermostabile DNA-Polymerase aus *Thermus filiformis* (vgl. FEMS Microbiol. Lett. 22 (1994), 149–153; auch erhältlich unter der ATCC-Hinterlegungsnummer 43280) kann unter Anwendung der in US-Patent Nr. 4,889,818 bereitgestellten Verfahren sowie basierend auf der in Tabelle 1 bereitgestellten Sequenzinformation gewonnen werden. Jede der vorstehenden Sequenzen und Veröffentlichungen ist unter Bezugnahme herein eingeschlossen. Die Sequenzhomologie (Sequenzidentität) zwischen der Aminosäuresequenz der nativen Form der Taq-Polymerase, so wie in WO 89/06691 bereitgestellt, und der vorstehend erwähnten Tfl-Polymerase beträgt 87,4%. Die entsprechenden Homologien betragen im Hinblick auf die Tth-Polymerase 87,4%, im Hinblick auf die Tca-Polymerase 86,6%, im Hinblick auf die Tca-Polymerase 86,6%, im Hinblick auf die Bst-Polymerase (Hinterlegungsnummer U23149) 42,0%, im Hinblick auf die Bca-Polymerase



42,6% und im Hinblick auf die Ath-Polymerase 39,7%.

**[0041]** Wie in Tabelle 1 gezeigt, ist das entscheidende Motiv bei den thermostabilen DNA-Polymerasen außerordentlich konserviert. Wenn "X" ein Glutaminsäurerest ist, stellt die Veränderung des Gens, das die Polymerase codiert, das erfindungsgemäße Enzym bereit, das ohne weiteres rNTPs einbaut, verglichen zum Beispiel mit der Taq-Polymerase, worin das entscheidende Motiv nicht modifiziert ist. Folglich betrifft die Erfindung eine Klasse von Enzymen, die auch zum Beispiel die thermostabile DNA-Polymerase und das entsprechende Gen sowie Expressionsvektoren aus *Thermus oshimai* (Williams et al., Int. J. Syst. Bacteriol. 46 (2) (1996), 403–408); *Thermus silvanus* und *Thermus chliarophilus* (Tenreiro et al., Int. J. Syst. Bacteriol. 45 (4) (1995), 633–639); *Thermus scotoductus* (Tenreiro et al., Res. Microbiol. 146 (4) (1995), 315–324); *Thermus brockianus* (Munster, Gen. Microbiol. 132 (1986), 1677) und *Thermus ruber* (Loginov et al., Int. J. Syst. Bacteriol. 34 (1984), 498–499; auch erhältlich unter der ATCC-Hinterlegungsnummer 35948) einschließt. Zusätzlich schließt die Erfindung zum Beispiel die modifizierten Formen der thermostabilen DNA-Polymerasen und das entsprechende Gen sowie Expressionsvektoren aus *Thermotoga elfii* (Ravot et al., Int. J. Syst. Bacteriol. 45 (1995), 312; auch erhältlich unter der DSM-Hinterlegungsnummer 9442) und *Thermotoga thermarum* (Windberger et al., Int. J. Syst. Bacteriol. 42 (1992), 327; auch erhältlich unter der DSM-Hinterlegungsnummer 5069) ein. Jede der vorstehenden Sequenzen und Veröffentlichungen sind unter Bezugnahme hierin eingeschlossen.

**[0042]** In einer bevorzugten Ausführungsform der Erfindung kommt das zu modifizierende entscheidende Motiv innerhalb der Aminosäuresequenz LeuAspTyrSerGlnIleGluLeuArgValLeuAlaHisLeuSer (SEQ ID NR: 5) vor. Folglich beinhaltet ein Aspekt der Erfindung die Erzeugung von thermostabilen DNA-Polymerase-Mutanten, die eine stark erhöhte Wirksamkeit für den Einbau von unkonventionellen Nucleotiden in einer matrizenabhängigen Weise zeigen. In einer besonders bevorzugten Ausführungsform umfaßt die Polymerasesequenz LeuAspTyrSerGlnIleGlyLeuArgValLeuAlaHisLeuSer (SEQ ID NR: 6). Solche thermostabilen DNA-Polymerasen sind für Verfahren wie die DNA-Sequenzierung, die DNA-gerichtete RNA-Synthese und die in vitro-Synthese einer rNTP substituierten DNA besonders geeignet.

**[0043]** Die Herstellung von thermostabilen DNA-Polymerasen mit einer erhöhten Effizienz des Einbaus von unkonventionellen Basen kann durch Verfahren wie die ortsspezifische Mutagenese erreicht werden. Vgl. zum Beispiel Sambrook et al., *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor, 1989, zweite Auflage, Kapitel 15.51, "Oligonucleotide-Mediated Mutagenesis", welches unter Bezugnahme hierin eingeschlossen ist. Zum Beispiel hat eine Mutation von "A" zu einem "G" in der zweiten Position des Codons, das den Glutaminsäurerest bei Rest 615 in der *Thermus aquaticus* (Taq)-DNA-Polymerase-Gensequenz codiert (vgl. SEQ ID NR: 7), eine mehr als das 500-fache Steigerung der Effizienz des Einbaus von unkonventionellen Nucleotiden, wie hierin definiert, zur Folge, während die Fähigkeit des Enzyms beibehalten wird, eine PCR in Anwesenheit von konventionellen Nucleotiden, d. h. dNTPs, zu vermitteln. Bei der Taq-DNA-Polymerase resultiert diese besondere Mutation in einem Aminosäureaustausch von E (Glutaminsäure) zu G (Glycin). Obwohl dieser besondere Aminosäureaustausch die Fähigkeit des Enzyms zum Einbau von unkonventionellen Nucleotiden wesentlich verändert, wird erwartet, daß der Austausch des Glutaminsäurerestes durch irgendwelche anderen Aminosäurereste wie z. B. durch einen Serin-, Cystein-, Threonin-, Alanin-, Valin- oder Leucinrest die gleiche Wirkung hat. Andere Aminosäuresubstitutionen, die E615 ersetzen, sind daher im Schutzzumfang der Erfindung eingeschlossen, obgleich E615G eine bevorzugte Ausführungsform darstellt. Folglich ist ein wichtiger Aspekt der Erfindung, daß der vierte Aminosäurerest in dem Motiv von SEQ ID NR: 1 kein Glutaminsäurerest ist.

**[0044]** Die ortsspezifische Mutagenese kann auch durch eine ortsspezifische Primergerichtete Mutagenese erreicht werden. Diese Technik ist nun Standard auf dem Fachgebiet und wird unter Verwendung eines synthetischen Oligonucleotidprimers durchgeführt, der zu einer einzelsträngigen Phagen-DNA, die durch Mutagenese verändert werden soll, bis auf eine begrenzte Zahl von Fehlpaarungen, welche die gewünschte Mutation darstellen, komplementär ist. Kurz zusammengefaßt wird das synthetische Oligonucleotid als ein Primer verwendet, um die Synthese eines Stranges zu steuern, der zu dem Plasmid oder Phagen komplementär ist, und die erhaltene doppelsträngige DNA wird in ein Phagen unterstützendes Wirtsbakterium transformiert. Die erhaltenen Bakterien können zum Beispiel durch eine DNA-Sequenzanalyse oder Sondenhybridisierung untersucht werden, um solche Plaques zu identifizieren, die die gewünschte mutierte Gensequenz tragen. Alternativ können "rekombinante PCR"-Verfahren angewendet werden, die in PCR Protocols, San Diego, Academic Press, Innis et al., Hrsg., 1990, Kapitel 22 mit dem Titel "Recombinant PCR", von Higuchi, Seiten 177–183, beschrieben sind.

**[0045]** Wie in Tabelle 1 gezeigt, ist der Glutaminsäurerest innerhalb des entscheidenden Motivs der Taq-Polymerase in anderen thermostabilen DNA-Polymerasen konserviert, aber kann an einer anderen, aber nahen Position, in der Aminosäuresequenz angeordnet sein. Eine Mutation des konservierten Glutaminsäurerestes

innerhalb von SEQ ID NR: 2 der thermostabilen DNA-Polymerasen aus *Thermus*-Spezies und der verwandten DNA-Polymerasen aus *Thermotoga*-, *Thermosiphon*- und *Anaerocellum*-Spezies zeigt eine ähnliche erhöhende Wirkung auf die Fähigkeit der Polymerase zum effizienten Einbau von unkonventionellen Nucleotiden im Vergleich zu der Taq-Polymerase, welche die SEQ ID NR: 2 umfaßt. Mutationen des Glutaminsäurerestes innerhalb des entscheidenden Motivs in anderen thermostabilen DNA-Polymerasen können unter Verwendung der zur ortsspezifischen Mutagenese angewendeten Prinzipien und Techniken erreicht werden. Es existieren mehrere Sequenzeingaben für die DNA-Polymerase aus *Bacillus stearothermophilus* in den GeneBank- oder SwissProt/PIR-Datenbanken. Diese Sequenzen sind eng miteinander verwandt, aber unterscheiden sich etwas voneinander, wobei aber jede die identische entscheidende Motivsequenz SerGlnIleGluLeuArgXaa (SEQ ID NR: 2) enthält, wobei "Xaa" an Position 7 dieser Sequenz ein Valinrest (Val) oder ein Isoleucinrest (Ile) ist, obgleich an verschiedenen Positionen in der Sequenz.

**[0046]** Basierend auf der öffentlich zugänglichen Aminosäure- und Nucleinsäuresequenzinformation für thermostabile DNA-Polymerasen, wie hierin beschrieben, ist es auch möglich, durch herkömmliche Rekombinationsmethodiken chimäre Polymerasen zu konstruieren, die aus Domänen bestehen, welche aus verschiedenen thermostabilen DNA-Polymerasen stammen. Die US-Patente Nrn. 5,466,591 und 5,374,553 beschreiben Verfahren zum Austausch der verschiedenen funktionellen Segmente von thermostabilen Polymerasen wie der 5'-3'-Exonucleasedomäne, der 3'-5'-Exonucleasedomäne und der Polymerasedomäne, um neue Enzyme bereitzustellen. Die bevorzugten chimären thermostabilen Polymeraseenzyme umfassen eine 5'-3'-Exonucleasedomäne, eine 3'-5'-Exonucleasedomäne und eine Polymerasedomäne, wobei eine Domäne aus einer anderen Polymerase abgeleitet ist und wobei die Polymerasedomäne die entscheidende Motivsequenz SerGlnIleXaaLeuArgXaa (SEQ ID NR: 1) umfaßt, wobei "Xaa" an Position 4 dieser Sequenz ein beliebiger Aminosäurerest, aber kein Glutaminsäurerest (Glu) ist, und wobei "Xaa" an Position 7 dieser Sequenz ein Valinrest (Val) oder ein Isoleucinrest (Ile) ist. Beispiele für solche chimären Moleküle sind chimäre Taq/Tma-Enzyme, die zusammengesetzt sind, wie in Tabelle 2 angegeben ist. Wie in dieser Tabelle gezeigt, enthält die Polymerasedomäne dieser chimären Taq/Tma-Enzyme die Mutation in dem vorstehend spezifizierten entscheidenden Motiv.

Tabelle 2

	5' - 3' - Exonuclease - domäne	3' - 5' - Exonuclease - domäne	Polymerasedomäne
Taq	AA 1-289	AA 290-422	AA 423-832
Tma	AA 1-291	AA 292-484	AA 485-893
Taq/Tma	AA 1-289 (Taq)	AA 292-484 (Tma)	AA 423-832 (Taq) mit E615G - Mutation
Taq/Tma	AA 1-289 (Taq)	AA 292-484 (Tma)	AA 485-893 (Tma) mit E678G - Mutation

**[0047]** Plasmid pC1 ist gemäß dem Budapester Vertrag bei der ATCC am 17. Juli 1996 unter der Hinterlegungsnummer 98107 hinterlegt worden. Das Plasmid pC1 enthält ein Gen, das eine thermostabile DNA-Polymerase codiert, die in dem Codon mutiert ist, welches den Glutaminsäurerest an Position 615 der Aminosäuresequenz der nativen Taq-Polymerase codiert, wodurch eine mutierte Form der Taq-Polymerase mit einem Glycinrest an Position 615 erhalten wird (die E615G mutierte Taq-Polymerase weist die Sequenz von SEQ ID NR: 8 auf). Diese Hinterlegung stellt ein anderes Mittel für die Bereitstellung von thermostabilen DNA-Polymerasen mit einer erhöhten Wirksamkeit für den Einbau von unkonventionellen Nucleotidanaloga bereit. Beispiel I veranschaulicht die Verwendung umgebender Restriktionsstellen, die zur Subclonierung der E615G-Mutation geeignet sind, um andere thermostabile DNA-Polymeraseenzyme zu erzeugen. Da die vollständige Gensequenz für zahlreiche thermostabile DNA-Polymerasen bekannt ist, sind andere Mittel zur Einführung einer Mutation in das E615 codierende Codon, wie durch Restriktionsspaltung und Fragmentaustausch oder durch ortsspezifische in vitro-Mutagenese, basierend auf der Sequenzinformation über das hierin bereitgestellte entscheidende Motiv für Fachleute ohne weiteres zugänglich.

**[0048]** Das durch ortsspezifische Mutagenese hergestellte modifizierte Gen oder Genfragment kann aus dem Plasmid oder Phagen durch herkömmliche Mittel gewonnen und in einen Expressionsvektor für die spätere Züchtung und Reinigung des erhaltenen Enzyms ligiert werden. Zahlreiche Clonierungs- und Expressionsvektoren einschließlich Säuger- und Bakteriensysteme sind zur Durchführung der Erfindung geeignet und sind

zum Beispiel bei Sambrook et al., *Molecular Cloning: A Laboratory Manual*, zweite Auflage, Cold Spring Harbor, 1989, beschrieben. Der Einfachheit halber wird die vorliegende Erfindung unter Verwendung des aus Lambda abgeleiteten PL-Promotors (Shimatake et al., *Nature* 292 (1981), 128) veranschaulicht. Die Verwendung dieses Promotors ist in den US-Patenten Nrn. 4,711,845 und 5,079,352 genauer beschrieben.

**[0049]** Die erfindungsgemäßen thermostabilen DNA-Polymerasen werden im Allgemeinen aus Mikroorganismen wie z. B. *E. coli* gereinigt, die mit einem Expressionsvektor transformiert worden sind, der mit einem Gen funktionell verbunden ist, das eine Wildtyp- oder eine modifizierte thermostabile DNA-Polymerase codiert. Ein Beispiel für einen geeigneten Wirtsmikroorganismus ist der *E. coli*-Stamm DG116, beschrieben durch Lawyer et al., *PCR Methods and Applications* 2 (1993), 275–287, wobei der Stamm auch von der American Type Culture Collection unter der Hinterlegungsnummer ATCC 53601 erhältlich ist. Verfahren zur Reinigung der thermostabilen Polymerase sind auch zum Beispiel bei Lawyer et al., *PCR Methods and Applications* 2 (1993), 275–287, beschrieben.

**[0050]** Den Fachleuten ist bekannt, daß die vorstehenden thermostabilen DNA-Polymerasen mit einer erhöhten Effizienz für den Einbau von unkonventionellen Nucleotiden am einfachsten unter Anwendung von Verfahren der DNA-Rekombinationstechnik hergestellt werden. Falls der Wunsch besteht, eines der erfindungsgemäßen Enzyme oder ein Derivat oder Homolog solcher Enzyme herzustellen, erfordert die Herstellung einer rekombinanten Form des Enzyms typischerweise die Konstruktion eines Expressionsvektors, die Transformation einer Wirtszelle mit dem Vektor und die Züchtung der transformierten Wirtszelle unter Bedingungen, so daß eine Expression erfolgt. Mittel zur Herstellung von Expressionsvektoren, zur Transformation und zur Züchtung von transformierten Wirtszellen sind auf dem Fachgebiet gut bekannt und sind zum Beispiel bei Sambrook et al., 1989, a. a. O., ausführlich beschrieben.

**[0051]** Die vorliegende Erfindung stellt thermostabile DNA-Polymerasen bereit, die zur Verwendung in Verbindung mit Ribonucleosidtriphosphaten für zahlreiche Anwendungen einschließlich Nucleinsäureamplifikations-, Nachweis- und DNA-Sequenzierungsverfahren geeignet sind. Die Verwendung von Ribonucleotiden bei der Sequenzierung vermeidet die hohen Kosten von Kettenterminationsanaloga wie ddNTPs und erleichtert, was wichtig ist, die Herstellung von neuen Amplifikationsprodukten, die nicht nur für die DNA-Sequenzanalyse, sondern auch für andere Analyseformen wie die Elektrophorese oder Hybridisierung geeignet sind, ohne daß nachträgliche DNA-Sequenzierungsreaktionen durchgeführt werden müssen.

**[0052]** Es ist gezeigt worden, daß Pyrophosphatase die Sequenzierungsergebnisse unter Verwendung von sowohl mesophilen Polymerasen als auch thermostabilen DNA-Polymerasen verbessert, indem sie das Ausmaß einer Pyrophosphorolyse verringert, wenn sich die Verlängerungsprodukte anreichern. In der Tat erfordern Verfahren, die vor einer Zyklus-Sequenzierung durchgeführt werden, daß das zusätzliche Enzym in der Sequenzierungsreaktion eingeschlossen ist. Ein sehr nützlicher und vorteilhafter Aspekt der vorliegenden Erfindung ist jedoch, daß die Pyrophosphatase nicht zur DNA-Sequenzierung erforderlich ist. Folglich vermeidet die Verwendung der hierin bereitgestellten neuen Enzyme die Notwendigkeit für die zusätzlichen Kosten der Zugabe eines zweiten Enzyms zu dem Sequenzierungsreaktionsgemisch.

**[0053]** Unter Verwendung der erfindungsgemäßen Enzyme werden die Amplifikations- und Sequenzierungsreaktionen kombiniert, was Zeit und Material spart, sowie die gesamte Analyse vereinfacht. Diese und andere Vorteile sind hauptsächlich dadurch erhältlich, daß der Einbau von sowohl konventionellen Nucleotiden als auch Ribonucleotiden und Ribonucleotidanaloga in ein Primerverlängerungsprodukt einen chimären RNA/DNA-Strang bereitstellt, wobei die RNA für eine Hydrolyse empfindlich ist. Die Behandlung beeinflusst nicht das DNA-Gerüst und stellt eine Population von Nucleinsäurefragmenten bereit, die jeweils an der Position enden, wo ein Ribonucleotid anstelle des entsprechenden dNTP eingebaut wurde. Die Hydrolyse wird ohne weiteres durch verschiedene Mittel einschließlich, aber nicht begrenzt auf, Alkali (z. B. durch Behandlung mit NaOH, z. B. in einer Endkonzentration von 0,2 M, wie nachstehend in Beispiel VI gezeigt), Hitze oder eine enzymatische Behandlung mit einer RNase (Vogel et al., Hrsg., *Informational Macromolecular*, New York, Academic Press, 1963, Kapitel von Berg et al., mit dem Titel "The Synthesis of Mixed Polynucleotides Containing Ribo- and Deoxyribonucleotide by Purified Preparation of DNA Polymerase from *E. coli*", Seiten 467–483) erreicht.

**[0054]** In einer bevorzugten Ausführungsform stellt die vorliegende Erfindung neue und verbesserte Zusammensetzungen bereit, die für DNA-Sequenzierungsverfahren besonders nützlich sind. Die hierin beschriebenen neuen Enzyme sind bei Nucleinsäure-Sequenzierungsverfahren unter Verwendung von entweder Farbstoff-Terminatoren oder Farbstoff-Primern sowie bei anderen Sequenzierungsverfahren vorteilhaft. Wie bereits beschrieben, erfordern Kettenterminationsverfahren im Allgemeinen eine matrizenabhängige Primerverlänge-

zung in Anwesenheit von Ketteterminationsnucleotiden, woraus eine Verteilung von Teilfragmenten resultiert, die nachfolgend nach Größe getrennt werden. Eine Standard-Didesoxysequenzierung verwendet Didesoxynucleosidtriphosphate zur Kettentermination und eine DNA-Polymerase wie das Klenow-Fragment von *E. coli* Poll (vgl. Sanger et al., a. a. O.).

**[0055]** Folglich erfordert das basische Didesoxy-Sequenzierungsverfahren (i) das Anlagern eines Oligonucleotidprimers an eine Matrize; (ii) das Verlängern des Primers mit einer DNA-Polymerase in vier getrennten Reaktionen, wobei jede ein markiertes Nucleotid oder einen markierten Primer, ein Gemisch aus unmarkierten dNTPs und ein Ketteterminations-ddNTP enthält; (iii) das Auftrennen der vier Sätze von Reaktionsprodukten zum Beispiel mit Hilfe einer hochauflösenden denaturierenden Polyacrylamid/Harnstoff-Gelelektrophorese, Kapillartrennung oder durch ein anderes Mittel zur Auftrennung; und (iv) das Erzeugen einer autoradiographischen Abbildung des Gels, das untersucht werden kann, um die Sequenz abzuleiten. Alternativ können Massenspektroskopieverfahren oder auf Hybridisierung basierende Verfahren unter Verwendung von fluoreszenzmarkierten Primern oder Nucleotiden angewendet werden, um eine DNA-Sequenzinformation abzuleiten.

**[0056]** Die Verfügbarkeit von thermoresistenten Polymerasen wie der Taq-Polymerase hat zu verbesserten Sequenzierungsverfahren (vgl. US-Patent Nr. 5,075,216) und Modifikationen hiervon, bezeichnet als "Zyklus-Sequenzierung", geführt. Bei der Zyklus-Sequenzierung werden Zyklen der Erhitzung und Kühlung wiederholt, um zu ermöglichen, daß zahlreiche Verlängerungsprodukte von jedem Zielmolekül erzeugt werden (Murray, *Nucleic Acids Research* 17 (1989), 8889). Diese asymmetrische Amplifikation von Zielsequenzen, die zu der Matrizensequenz komplementär sind, in Anwesenheit von Didesoxy-Ketteterminatoren ergibt eine Familie von Verlängerungsprodukten aller möglichen Längen.

**[0057]** Nach der Denaturierung des Verlängerungsreaktionsprodukts der DNA-Matrize finden mehrere Zyklen einer Primeranlagerung und Primerverlängerung in Anwesenheit von Didesoxy-Terminatoren statt. Thermostabile DNA-Polymerasen weisen bei der Zyklus-Sequenzierung mehrere Vorteile auf; sie tolerieren die stringenter Anlagerungstemperaturen, die für eine spezifische Hybridisierung des Primers mit den Nucleinsäure-Zielmolekülen erforderlich sind, und sie tolerieren auch die vielfachen Zyklen einer Denaturierung bei einer hohen Temperatur, die in jedem Zyklus erfolgt, d. h. 90–95°C. Aus diesem Grund sind verschiedene Formen der AmpliTaq®-DNA-Polymerase in Kits für eine Taq-Zyklus-Sequenzierung, vertrieben durch Perkin Elmer, Norwalk, CT, eingeschlossen worden.

**[0058]** Trotzdem stellt die Eigenschaft der Taq-DNA-Polymerase, eine Diskrimination gegen den Einbau von unkonventionellen Nucleotiden wie ddNTPs zu zeigen, ein Problem dar, wenn sie zur Zyklus-Sequenzierung verwendet wird, bei der ddNTPs oder fluoreszenzmarkierte ddNTPs als Ketteterminatoren eingebaut werden müssen. Vor der vorliegenden Erfindung machte die DNA-Sequenzierung mit thermostabilen DNA-Polymerasen allgemein ein Gemisch aus Ketteterminationsnucleotiden, im allgemeinen Didesoxynucleotiden, in hohen Konzentrationen erforderlich, um sicherzustellen, daß eine Population von Verlängerungsprodukten erzeugt werden würde, die alle möglichen Fragmentlängen über eine Strecke von mehreren hundert Basen repräsentiert. Um diese Kostenfrage zu lösen, verwendeten die Protokolle häufig sehr geringe Konzentrationen der konventionellen dNTPs, was die Reaktionen ineffizient machte. Diese Reaktionsgemische mit einer geringen dNTP-Konzentration und einer hohen ddNTP-Konzentration erzeugen eine Umgebung, in der die thermostabile Polymerase durch die Nucleotidsubstrate im Wesentlichen verbraucht wird.

**[0059]** Trotz des Aufkommens von modifizierten Enzymen wie der AmpliTaq®-DNA-Polymerase FS, die ermöglichen, daß die Konzentration der dNTPs über die optimaleren Konzentrationen hinaus erhöht werden kann, sind die früheren Enzyme immer noch auf die Anwesenheit der kostspieligen ddNTPs zur DNA-Sequenzierung angewiesen. Im Gegensatz dazu, stellt die vorliegende Erfindung Enzyme bereit, die nicht nur ermöglichen, daß die Konzentration der dNTPs erhöht werden kann, sondern auch die Verwendung der kostspieligen ddNTPs vermeiden, indem statt dessen rNTPs für den Einbau in den wachsenden Strang verwendet werden. Die Fähigkeit der neuen Enzyme, effizient eine teilweise Ribonucleotidsubstitution zu bewirken, erleichtert die Erzeugung von leiterförmigen DNA-Sequenzierungsmustern in Abwesenheit einer separaten Reaktion für den Einbau eines Terminationsnucleotids.

**[0060]** Die Wahl von unkonventionellen Nucleotidanaloga, die zur Verwendung in DNA-Sequenzierungsverfahren geeignet sind, wurde durch die Fähigkeit der thermostabilen DNA-Polymerase bestimmt, diese Analoga einzubauen. Unglücklicherweise sind diese Nucleotidanaloga sehr teuer. Zum Beispiel sind die Kosten für ddNTPs etwa 25-mal höher als die Kosten für entweder rNTPs oder dNTPs. Da die vorherigen thermostabilen DNA-Polymerasen nicht in der Lage waren, effizient rNTPs in einer Matrizen gerichteten Weise in einen wachsenden DNA-Strang einzubauen, waren solche Ribonucleotide, die ohne weiteres erhältlich und preisgünstig

sind, keine Alternative zur Verwendung bei der DNA-Sequenzierung in Verbindung mit einer thermostabilen DNA-Polymerase. Die vorliegende Erfindung beseitigt die Notwendigkeit für ddNTPs in DNA-Sequenzierungsreaktionen. Folglich stellt die Erfindung gemäß einem Aspekt Verfahren zur DNA-Sequenzierungsanalyse bereit, die wesentlich preisgünstiger als frühere Kettenterminationsverfahren sind.

**[0061]** Die Anwesenheit von Mangan in einer Primerverlängerungsreaktion kann die Fähigkeit einer Polymerase beeinflussen, das richtig basengepaarte Nucleotid genau einzubauen. Mangan kann verwendet werden, um eine falsche Basenpaarung zu erzwingen oder um die Diskrimination gegen die Insertion eines Nucleotidanalogs zu vermindern. Mangan ist durch Forscher verwendet worden, um eine Mutagenese in DNA-Replikations- oder Amplifikationsverfahren zu induzieren. Folglich kann Mangan die Genauigkeit einer Polymerisierungsreaktion sowie die Ausbeute einer Reaktion beeinflussen. Die erhaltene Sequenz kann Fehler enthalten, oder die aus einem DNA-Sequenzierungsverfahren erhaltene Information kann unklar sein. Die vorliegenden Verfahren machen es nicht erforderlich, daß Mangan als ein zweiwertiges Kation in dem Sequenzierungsreaktionsgemisch eingeschlossen ist, um die Polymerase zum Einbau eines unkonventionellen Nucleotids zu zwingen. Im Gegensatz zu den vorherigen DNA-Polymerasen identifiziert die vorliegende Erfindung das entscheidende Motiv innerhalb der Polymerasedomäne zur Kontrolle der Fähigkeit des Enzyms, zwischen 2'-substituierten und unsubstituierten Nucleotiden zu unterscheiden, ohne daß Mangan erforderlich ist.

**[0062]** Die erfindungsgemäßen Enzyme benötigen keine hohen Konzentrationen der unkonventionellen Basenanaloga zur Sequenzierung. Vor der vorliegenden Erfindung lagen die unkonventionellen Basenanaloga und die entsprechenden konventionellen Basen im Allgemeinen in einem Verhältnis (z. B. ddATP : ATP) im Bereich von etwa 1,3 : 1 bis 24 : 1 für Kettenterminations-DNA-Sequenzierungsverfahren vor (vgl. auch US-Patent Nr. 5,075,216 von Innis et al.). Im Vergleich dazu ermöglichen die erfindungsgemäß bereitgestellten thermostabilen Polymerasen, daß das Verhältnis der unkonventionellen Basenanaloga zu den konventionellen Basen um ein Hundertfaches bis mehrere tausendmal verringert werden kann. Ein rNTP : dNTP-Verhältnis von 1 : 1 oder weniger in Verbindung mit den hierin bereitgestellten neuen Enzymen ist für eine DNA-Sequenzanalyse ausreichend. In einer bevorzugten Ausführungsform der Erfindung wird das rNTP : dNTP-Verhältnis auf weniger als 1 : 8 verringert. Das Verhältnis des 2'-substituierten Nucleotids zu dem entsprechenden natürlichen dNTP kann abhängig von der besonderen Versuchsanordnung und der gewünschten Länge der Fragmente so wenig wie 1 : 80 oder 1 : 200 betragen.

**[0063]** Folglich, da die vorliegenden Enzyme ohne weiteres unkonventionelle Nucleotide wie 2'-substituierte Nucleotide einbauen, ist es nicht erforderlich, den Einbau des rNTP durch die Verwendung einer hohen Konzentration des rNTP und einer begrenzenden Konzentration des entsprechenden dNTP zu erzwingen. Demgemäß ermöglichen die vorliegenden Verfahren die Verwendung von optimalen Konzentrationen der dNTPs in Kombination mit geringen Mengen an rNTPs.

**[0064]** Wenn erfindungsgemäße modifizierte Polymeraseenzyme in einem geeigneten Sequenzierungsverfahren wie z. B. einer Farbstoff-Primer-Sequenzierung verwendet werden, werden gute DNA-Sequenzierungsergebnisse mit einer dNTP-Konzentration im Bereich von 50–500 µM jedes dNTP erhalten. Vorzugsweise liegt die dNTP-Konzentration zwischen 100–300 µM. In diesen Bereichen kann das entsprechende rNTP in etwa der gleichen Konzentration wie das dNTP oder in einer geringeren Konzentration anwesend sein. Vorzugsweise ist das rNTP in etwa 0,1 µM–100 µM anwesend, am meisten bevorzugt liegt das rNTP in etwa 2,5 µM bis 25 µM vor.

**[0065]** Die zur Verwendung in Verbindung mit den vorliegenden modifizierten Enzymen geeignete Konzentration der rNTPs kann ohne weiteres durch Fachleute mittels Titrations- und Optimierungsexperimenten bestimmt werden. Die erforderliche Menge eines rNTP oder Analogons wird durch die Art des Experiments beeinflusst und kann durch die Größe und Reinheit des Zielmoleküls sowie die Wahl des Puffers und der besonderen Art des Enzyms bestimmt werden.

**[0066]** Das Verhältnis von rNTP : dNTP bestimmt die Häufigkeit, mit der rNTPs in das wachsende Oligonucleotid eingebaut werden. Da die Hydrolyse an jedem eingebauten rNTP abläuft, kann das Verhältnis von rNTP : dNTP angepaßt werden, um dem Anwender die Möglichkeit zu geben, die Größe der erhaltenen Fragmente flexibel zu erhöhen oder zu verringern.

**[0067]** Es ist hinreichend bekannt, daß die DNA ein aus dNTPs synthetisiertes Polymer ist. Jedes Desoxynucleosidtriphosphat umfaßt einen Ribosezucker, der eine Hydroxylgruppe an der 3'-Position und ein Wasserstoffatom an der 2'-Position enthält. Ribonucleotide enthalten auch eine Hydroxylgruppe in dem 3'-Teil des Zuckers. Jedoch unterscheiden sich rNTPs von dNTPs an der 2'-Position des Zuckers, wobei eine zweite Hydro-

xylgruppe das Wasserstoffatom ersetzt. Im vorliegenden Zusammenhang veranschaulichen rNTPs die Fähigkeit der erfindungsgemäßen Enzyme zum genauen Einbau von 2'-substituierten Nucleotiden. Jedoch sind die erfindungsgemäßen Verbindungen nicht auf die Verwendung von unkonventionellen Nucleotiden begrenzt, welche Ribonucleotide sind. Eine Modifikation der Sequenz der thermostabilen Polymerase in der hierin identifizierten entscheidenden Domäne ermöglicht den Matrizen gerichteten Einbau von anderen 2'-substituierten Nucleotiden wie 2'-Hydroxyl-3'-desoxy-Nucleotiden und substituierten 2'-Fluor- oder -Aminonucleotiden.

**[0068]** Wie in den Beispielen hierin beschrieben, wird der Einbau von 3'-Desoxy-2'-Hydroxy-ATP, hierin als Cordycepintriphosphat bezeichnet, durch die Anwesenheit einer zweiten Mutation in der thermostabilen Polymerase, welche die Diskrimination gegen den Einbau eines Nucleotids reduziert, das eine Desoxygruppe an der 3'-Position der Ribose enthält, erleichtert. Solche Enzyme sind zum Beispiel bereits in EP-A-65506 beschrieben worden. ATCC-Hinterlegungsnummer 69820, hinterlegt gemäß dem Budapester Vertrag am 10. Mai 1995, stellt das Gen bereit, das eine modifizierte thermostabile DNA-Polymerase aus *Thermus aquaticus* codiert, die eine reduzierte Diskrimination gegen den Einbau von Analoga wie ddNTPs aufweist. Didesoxynucleotide weisen im Vergleich zu konventionellen dNTPs eine substituierte 3'-Position auf. Folglich stellt die Doppelmutation, veranschaulicht hierin durch eine E615G, F667Y-Taq-Polymerase-Mutante, in Verbindung mit der vorliegenden Erfindung ein Mittel zur Verwendung von Nucleotidanaloga bereit, die an den 3'- und 2'-Positionen der Ribose im Vergleich zu dNTPs substituiert sind (vgl. Beispiele III und V).

**[0069]** Eine besondere Anwendung der Erfindung ist ein rNTP-Sequenzierungsverfahren, wobei der Sequenzierungsprimer mit einem unterscheidbaren fluoreszierenden oder radioaktiven Marker nachweisbar markiert ist. Anders als bei ddNTPs hat der Einbau eines unmodifizierten rNTP kein Kettenterminationsereignis zur Folge. Die DNA-Sequenzierungsreaktion, umfassend sowohl rNTPs als auch dNTPs in Kombination mit einem erfindungsgemäßen Enzym, ergibt ein Gemisch aus statistisch substituierten Primerverlängerungsprodukten, die für eine Spaltung an der 3'-5'-Phosphodiesterbindung zwischen einem Ribo- und einem benachbarten Desoxyribonucleotid empfindlich sind. Nach der Primerverlängerung, zum Beispiel bei einer PCR-Amplifikation oder einer Zyklus-Sequenzierung, und vor der Auftrennung der Primerverlängerungsprodukte, zum Beispiel mittels Gelelektrophorese, wird das Reaktionsgemisch mit entweder Alkali, Hitze, einer Ribonuclease oder einem anderen Mittel zur Hydrolyse der Verlängerungsprodukte bei jedem Ribonucleotid behandelt. Für jedes markierte Primerverlängerungsprodukt ist nur das am meisten 5'-endende Fragment, welches das unmittelbare Verlängerungsprodukt des markierten Primers ist, auf einem Sequenzierungsgel nachweisbar. Für ein bestimmtes Zielmolekül ergibt die Analyse des erhaltenen Sequenzierungsgels ein leiterförmiges Sequenzierungsmuster, d. h. eine Reihe von identifizierbaren Signalen in den G-, A-, T- und C-Bahnen, die der Nucleinsäuresequenz des Zielmoleküls entsprechen. Das erhaltene leiterförmige Sequenzierungsmuster stellt die gleiche Information bereit, ob das Verfahren ddNTPs durch herkömmliche Mittel oder rNTPs und die hierin beschriebenen neuen thermostabilen Polymerasen verwendet. Folglich sind durch die Anwendung der vorliegenden Erfindung nicht mehr länger teure ddNTPs für die DNA-Sequenzierung erforderlich (vgl. Beispiel VI).

**[0070]** In einem anderen Sequenzierungsverfahren werden Kettenterminations-Ribonucleotide verwendet. In dieser Ausführungsform der Erfindung werden 2'-Hydroxy-3'-desoxy-Nucleotide wie Cordycepintriphosphat als Terminatoren verwendet. Diese rNTP-Analoga können fluoreszenzmarkiert sein und zur DNA-Sequenzierung verwendet werden. Lee et al. (a. a. O.) haben die Verwendung von Farbstoff-Terminator-ddNTPs beschrieben. EP-A-655,506 und US-Serien 08/448,223, eingereicht am 23. Mai 1995, beschreiben modifizierte Enzyme zur Verwendung in Verbindung mit ddNTPs. Eine thermostabile DNA-Polymerase, umfassend sowohl die in der AmpliTaq®-DNA-Polymerase FS vorhandene Modifikation (vgl. vorstehend) als auch die in SEQ ID NR: 1 spezifizierten, wobei X kein Glutaminsäurerest (E) ist, wie hierin beschrieben, kann zum effizienten Einbau der markierten rNTP-Analoga in einer Kettenterminations-Sequenzierungsreaktion verwendet werden. Dieses Verfahren kann automatisiert werden und macht nicht die Synthese von farbstoffmarkierten Primern erforderlich. Außerdem, da Farbstoff-Terminator-Reaktionen ermöglichen, daß alle vier Reaktionen in dem gleichen Röhrchen durchgeführt werden können, sind sie praktischer als Farbstoff-Primer-Verfahren. Die 2'-Hydroxy-3'-desoxy-Nucleotide können aus im Handel erhältlichen 3'-Nucleotiden (3'-dA, 3'-dC, 3'-dG und 3'-dT, z. B. erhältlich von Sigma Chemical Corporation, St. Louis, MO) und unter Addition eines 5'-Triphosphats, wie bei Ludwig, *Biophosphates and Their Synthesis Structure, Metabolism and Activity*, Hrsg., Bruzik und Stec, Amsterdam, Elsevier Science Publishers, 1987, Seiten 201–204, hergestellt werden.

**[0071]** Zusätzlich zur Verwendung der erfindungsgemäßen Enzyme in neuen Sequenzierungsverfahren sind die hierin beschriebenen modifizierten Enzyme in einer Reihe von molekularbiologischen Anwendungen nützlich. In einer Ausführungsform wird das modifizierte Enzym in einem Amplifikationsreaktionsgemisch verwendet, das sowohl konventionelle als auch unkonventionelle Nucleotide, zum Beispiel dNTPs, und mindestens ein nachweisbar markiertes rNTP, wobei die Markierungen zum Beispiel Fluoreszenzmarker oder Radioisoto-

pe einschließen, umfaßt. Die Matrizen gerichtete Synthese eines komplementären Strangs ergibt ein DNA-Produkt, das an verschiedenen Positionen über seine gesamte Länge Ribonucleosidmonophosphate enthält. Hitze und/oder eine Alkalibehandlung hydrolysieren das Nucleinsäureverlängerungsprodukt an jedem Ribonucleotid. Auf diese Weise wird eine Familie von DNA-Fragmenten bereitgestellt, wobei jedes Fragment eine markierte Einheit an seinem 3'-Ende enthält. Die Größe der erhaltenen Nucleinsäurefragmente kann durch das Anpassen des Verhältnisses und der Menge des in der Reaktion eingeschlossenen rNTP modifiziert werden.

**[0072]** Die Amplifikation eines Zielmoleküls unter Verwendung von rNTPs und der vorliegenden Enzyme stellt zahlreiche Vorteile abhängig von der besonderen Anwendung bereit. Bei dem vorstehend beschriebenen Verfahren unter Verwendung eines markierten rNTP sind in der erhaltenen Familie die Fragmente alle mit der gleichen Intensität markiert: einem Marker pro Oligonucleotidfragment. Verfahren wie der Nucleinsäurenachweis unter Verwendung eines Oligonucleotidsonden-Arrays, das auf einem Silicon-Chip fixiert ist, erfordert normalerweise, daß das amplifizierte Zielmolekül statistisch innerhalb eines festgelegten reproduzierbaren Größenbereiches fragmentiert wird, um die Bildung von Sekundärstrukturen zur Kontrolle der Hybridisierungskinetik zu begrenzen. Außerdem kann zum Nachweis der Hybridisierung an ein Array aus Tausenden von Sonden auf einem Chip bevorzugt werden, daß die Nucleinsäurefragmente mit der gleichen Intensität markiert sind. Die vorliegende Erfindung stellt ein Mittel zur Herstellung von Familien von Fragmenten bereit, welche diesen Standard erfüllen, und erleichtert auf diese Weise die Verwendung von anderen Nachweisarrangements wie der auf einem Chip basierenden Verfahren, die zum Beispiel bei Cronin et al., Human Mutation 7 (1996), 244–255, beschrieben sind.

**[0073]** In einer anderen Ausführungsform stellt die Verwendung eines markierten Primers und eines unmarkierten Primers in einer Amplifikationsreaktion, die eine erfindungsgemäße thermostabile Polymerase und sowohl rNTPs als auch dNTPs umfaßt, ein Mittel für die gleichzeitige Durchführung von Amplifikations- und Sequenzierungsreaktionen bereit. Dieses Verfahren macht erforderlich, daß vier getrennte Amplifikationsreaktionen, eine für jedes rNTP, durchgeführt werden. Folglich, da das erfindungsgemäße Enzym zum Beispiel zur Amplifikation eines Zielmoleküls geeignet ist, zum Beispiel durch eine PCR oder ein anderes Mittel zur Amplifikation, kann das erhaltene Produkt, falls es vorhanden ist, zum Beispiel durch herkömmliche Verfahren wie eine Gelelektrophorese oder Sondenhybridisierung unter Verwendung eines Teils des Reaktionsprodukts nachgewiesen werden. Diese Nachweisverfahren haben nicht die Hydrolyse der eingebauten Ribonucleotide zur Folge, und die chimären RNA/DNA-Stränge verhalten sich, wie für ein herkömmliches Nucleinsäure-Amplifikationsprodukt erwartet wird. Falls ein gewünschtes Produkt nachgewiesen wird, kann ein verbleibender Teil des gleichen Reaktionsgemisches mit Alkali behandelt und mittels Gelelektrophorese zur Bestimmung der Nucleinsäuresequenz analysiert werden. Auf diese Weise ist nach dem Nachweis des Produkts keine nachfolgende Sequenzierungsreaktion notwendig. Dieses vereinfachte Verfahren spart Zeit und Material und ergibt eine erhöhte Genauigkeit durch das Weglassen von Schritten: das nachgewiesene Produkt ist das sequenzierte Produkt.

**[0074]** Ein ähnliches Verfahren mit vier markierten rNTPs und einem biotinylierten Primer könnte auch angewendet werden. Nach der Amplifikation wird das Produkt mit einem Alkali gespalten, und die Primer assoziierten Produkte werden durch eine Reaktion mit Streptavidin beschichteten Kügelchen entfernt. Die abgefangenen Produkte werden anschließend auf einem Sequenzierungsgel analysiert. Diese Modifikation ermöglicht, daß die Sequenzierungsreaktion in einem Röhrchen durchgeführt werden kann, wodurch die Notwendigkeit für vier getrennte Amplifikationen vermieden wird.

**[0075]** Gemäß einem anderen Aspekt der Erfindung sind die hierin beschriebenen Enzyme zur Herstellung von RNA aus einer DNA-Matrize oder zur Erzeugung einer substituierten DNA für eine Alkali vermittelte Sterilisation ohne die Verwendung von herkömmlichen Sterilisierungsmitteln wie Uracil-N-Glycosylase (UNG), wie in der Internationalen Patentveröffentlichung Nr. WO 92/01814 beschrieben, nützlich.

**[0076]** In einer beispielhaften Ausführungsform enthält die thermostabile Polymerase auch eine Mutation in der 5'-3'-Exonuclease-domäne, welche dazu dient, diese Exonucleaseaktivität stark zu verringern. Modifizierte Formen der Taq-Polymerase sind in US-Patent Nr. 5,466,591 beschrieben. In einer Ausführungsform dieser Erfindung ist das Codon, welches den Glycinrest (G) an der Aminosäureposition 46 codiert, durch ein Codon ersetzt worden, das einen Asparaginsäurerest (D) codiert. Das erhaltene Enzym weist infolge der verringerten 5'-3'-Exonucleaseaktivität eine größere Nützlichkeit in Zyklus-Sequenzierungsreaktionen auf und ist eine bevorzugte Grundlage zur Verwendung in der vorliegenden Erfindung. Die Aminosäuresequenz der Polymerase-domäne und die Aktivität der Polymerase werden durch die Anwesenheit der (G46D)-Mutante im Vergleich zu dem Wildtyp-Enzym nicht beeinflusst.

**[0077]** In einer kommerziellen Ausführungsform der Erfindung stellen Kits für die Nucleinsäuresequenzierung, umfassend eine erfindungsgemäße thermostabile Polymerase, eine kommerzielle Ausführungsform der Erfindung dar. Solche Kits schließen typischerweise zusätzliche Reagenzien für die DNA-Sequenzierung ein, wie z. B. rNTPs, dNTPs und geeignete Puffer. Wenn die rNTPs unmarkiert sind, kann auch ein markierter Primer eingeschlossen sein.

**[0078]** Die folgenden Beispiele werden nur zur Veranschaulichung angegeben und sollen den Schutzbereich der beanspruchten Erfindung in keiner Weise begrenzen.

#### Beispiel I

Expression eines modifizierten Gens der Taq-Polymerase mit einer reduzierten Diskrimination gegen unkonventionelle Nucleotide

**[0079]** Der C-terminale Aminosäureanteil der Taq-DNA-Polymerase codiert das aktive Zentrum der Polymerase (Lawyer et al., PCR Methods and Applications 2 (1993), 275–287, hierin unter Bezugnahme eingeschlossen). Ein DNA-Fragment, das diese Region enthält, wurde aus dem vollständigen Taq-Gen isoliert und mittels PCR-Amplifikation in Gegenwart von Mangan durch Mutagenese verändert (Leung et al., Technique 1 (1) (1989), 11–15). Für dieses Beispiel wurden alle Restriktionsenzyme von New England Biolabs, Beverly, MA, bezogen. Die durch Mutagenese veränderten Fragmente wurden mit PstI und BglII gespalten und in ein Taq-Expressionsplasmid, hierin das Plasmid pLK102, cloniert, das mit PstI und BglII gespalten worden war. Das Plasmid pLK102 ist eine modifizierte Form des Taq-Expressionsplasmids pSYC1578 (Lawyer et al., a. a. O.). Das HincII/EcoRV-Fragment, angeordnet 3' zu der Region, die die Polymerase codiert, wurde deletiert, um das Plasmid pLK101 zu erzeugen. Ein 898 Basenpaare umfassendes PstI-BglII-Fragment wurde anschließend aus pLK101 deletiert und durch einen kurzen PstI-EcoRV-BglII-Oligonucleotiddoppelstrang ersetzt, um das Plasmid pLK102 zu erzeugen. Diese Deletion entfernt somit 900 Basenpaare aus dem 3'-Ende des Taq-DNA-pol-Gens und ersetzt diese Region durch ein kurzes DNA-Stück.

**[0080]** Die erhaltenen Expressionsplasmide wurden in den E. coli-Stamm N1624 transformiert (beschrieben durch Gottesman, J. Mol. Biol. 77 (1973), 531; auch erhältlich vom E. coli Genetic Stock Center an der Yale-Universität unter der Stamm-Nr. CGSC #5066), und die erhaltenen Transformanten wurden auf die Fähigkeit zum effizienten Einbau von rNTPs im Vergleich zu dem Wildtyp-Enzym durchmustert. Unter Anwendung dieses Verfahrens wurde die Mutante C1 identifiziert, welche die Fähigkeit zu einem effizienterem Einbau von rNTPs besitzt.

**[0081]** Um zu bestimmen, welcher Teil des Taq-Polymerase-Gens für den veränderten Phänotyp verantwortlich war, wurde das durch Mutagenese veränderte Taq-Expressionsplasmid, bezeichnet als pC1, das aus der Mutante C1 isoliert worden war, mit verschiedenen Restriktionsenzymen gespalten, und die erhaltenen Restriktionsfragmente wurden in das Wildtyp-Taq-DNA-Polymerase-Gen von pLK101 subcloniert, wodurch die nicht durch Mutagenese veränderten Restriktionsfragmente ersetzt wurden. Die Analyse der erhaltenen Subclone zeigte, daß die für den Phänotyp verantwortliche Mutation innerhalb eines 265 Basenpaare umfassenden NheI-BamHI-Restriktionsfragments enthalten war.

**[0082]** In dieser Region von pC1 wurde eine DNA-Sequenzanalyse unter Verwendung des ABI PRISM<sup>®</sup> Farbstoff-Terminator-Zyklus-Sequenzierungs-Hauptkits mit AmpliTaq<sup>®</sup>-DNA-Polymerase FS von Applied Biosystems, Foster City, CA, und dem Applied Biosystems-DNA-Sequenzierungssystem Modell 373A durchgeführt. Die Sequenzanalyse identifizierte zwei Missense-Mutationen in dem Taq-Polymerase-Gen zwischen den NheI- und BamHI-Stellen. Eine Mutation an Aminosäureposition 615 bewirkte, daß ein Glutaminsäurerest (E) durch einen Glycinrest (G) ersetzt wird, und eine andere Mutation an Position 653 ersetzte einen Alaninrest (A) durch einen Threoninrest (T). Die Nummerierung beginnt an dem Codon, das den ersten Methioninrest des reifen Proteins codiert, so wie in US-Patent Nr. 5,079,352. Die E615G-Mutation wurde durch einen GAG-zu-GGG-Austausch in Codon 615 verursacht. Die A653T-Mutation wurde durch einen GCC-zu-ACC-Austausch in Codon 653 verursacht. Das Plasmid C1 in dem E. coli-Wirtsstamm N1624 wurde gemäß dem Budapester Vertrag bei der ATCC am 17. Juli 1996 unter der Hinterlegungsnummer 98107 hinterlegt.

**[0083]** Die zwei Punktmutationen wurden durch Subclonierung jeweils separat in ein Wildtyp-Taq-Polymerase-Gen mittels rekombinanter PCR (Innis et al., Hrsg., PCR Protocols, San Diego, Academic Press, 1990, Kapitel 22, mit dem Titel "Recombinant PCR", Higuchi, Seiten 177–183) getrennt analysiert. Die erhaltenen Expressionsprodukte wurden analysiert, um zu bestimmen, ob E615G oder A653T oder beide Mutationen für den



Phänotyp des Ribonucleotid-Einbaus verantwortlich waren. Die Ergebnisse dieses Experiments zeigten, daß die E615G-Mutation allein für den mutierten Phänotyp verantwortlich war.

**[0084]** Zur weiteren Analyse und Quantifizierung der Effizienz des Einbaus von Nucleotidanaloga wurde das 265 Basenpaare umfassende BamHI-NheI-PCR-Fragment, enthaltend E615G, in einen Taq-Expressionsvektor, pRDA3-2, cloniert. Der Expressionsvektor pRDA3-2 enthält das vollständige Taq-Gen in funktioneller Verbindung mit dem PL-Promotor aus dem Phagen Lambda. Die Exonucleasedomäne des Taq-Gens in diesem Vektor enthält eine Punktmutation in dem Codon, das einen Glycinrest codiert, und zwar an dem Aminosäurerest 46, welche die 5'-3'-Exonucleaseaktivität verringert. Jedoch ist die Gensequenz innerhalb der Polymerasedomäne des Expressionsvektors pRDA3-2 mit der Wildtyp-Taq-Gensequenz identisch. Das Plasmid pRDA3-2 ist in US-Patent Nr. 5,466,591, hierin unter Bezugnahme eingeschlossen, worin das Plasmid als "Clon 3-2" bezeichnet wird, ausführlich beschrieben. Das Plasmid pRDA3-2 wurde mit BamHI und NheI gespalten, und das 265 Basenpaare umfassende PCR-Fragment wurde durch herkömmliche Mittel in den Vektor ligiert.

**[0085]** Das erhaltene Plasmid, pLK108, wurde in den E. coli-Stamm DG116 transformiert (Lawyer et al., 1993, a. a. O., auch erhältlich von der American Type Culture Collection unter der ATCC-Nr. 53606). Das Plasmid pLK108 codiert eine thermostabile DNA-Polymerase, welche hierin als G46D, E615G-Taq bezeichnet wird. Eine Mutante, G46D, E615G, F667Y-Taq, wurde durch Kombination der E615G- und F667Y-Mutationen durch eine rekombinante PCR in ein BamHI-NheI-Fragment erzeugt. Dieses Fragment wurde in Plasmid pRDA3-2 cloniert, um das Plasmid pLK109 zu erzeugen. Das exprimierte Protein der thermostabilen DNA-Polymerase aus den Plasmiden pLK108 und pLK109 wurde gemäß dem durch Lawyer et al., 1993, a. a. O., beschriebenen Verfahren gereinigt, obwohl die Chromatographieschritte ausgelassen wurden. Die Sequenz der Insertionen wurde mittels DNA-Sequenzanalyse bestätigt. Eine zusätzliche Mutation in der Sequenz wurde in der pLK108-Insertion nachgewiesen; jedoch veränderte diese Mutation die Aminosäuresequenz des Proteins nicht.

**[0086]** Nach einer teilweisen Reinigung wurde die Aktivität des modifizierten Enzyms durch den bei Lawyer et al., J. Biol. Chem. 264 (1989), 6427–6437, hierin unter Bezugnahme eingeschlossen, beschriebenen Aktivitätstest bestimmt. Die Aktivität des modifizierten Enzyms wurde wie folgt berechnet: eine Enzymeinheit entspricht 10 nMol des Produkts, synthetisiert in 30 Minuten. Die Aktivität der DNA-Polymerase des Wildtyp-Enzyms ist linear proportional zu einer Enzymkonzentration von bis zu 80–100 pMol eingebautem dCMP (verdünntes Enzym in 0,12–0,15 Einheiten pro Reaktion). Die Aktivität der E615G-, G46D- und E615G, F667Y, G46D-Mutanten ist linear proportional zu Enzymkonzentrationen von bis zu 0,25–3 pMol eingebautem dCMP (verdünntes Enzym in  $6 \times 10^{-4}$  bis  $5 \times 10^{-3}$  Einheiten pro Reaktion). Diese Enzymzubereitung wurde in den in den Beispielen III–V beschriebenen Einbau- und Sequenzierungsreaktionen verwendet. Für die Beispiele II und VI wurde das Enzym gereinigt, wie bei Lawyer et al. (a. a. O.) beschrieben ist.

## Beispiel II

### Test zum Vergleich der Effizienz des Einbaus

**[0087]** Die relative Fähigkeit von G46D- und G46D, E615G-Taq zum Einbau von rNTPs wurde bestimmt, indem die Menge an  $[\alpha\text{-}^{32}\text{P}]\text{rNTP}$  gemessen wurde, die jedes Enzym bei einer begrenzenden Enzymkonzentration in eine aktivierte Lachssperma-DNA-Matrize einbauen konnte. Um den Einbau von rATP zu messen, wurde ein Reaktionsgemisch so hergestellt, daß die Endkonzentrationen in einer 50 µl-Reaktion wie folgt waren: 12,5 µg aktivierte Lachssperma-DNA, hergestellt, wie nachstehend beschrieben; jeweils 200 µM dCTP, dGTP und dTTP (Perkin Elmer, Norwalk, CT); 100 µM  $[\alpha\text{-}^{32}\text{P}]\text{rATP}$ ; 1 mM  $\beta$ -Mercaptoethanol; 25 mM N-Tris[hydroxymethyl]methyl-3-amino-propansulfonsäure (TAPS), pH 9,5, 20°C; 50 mM KCl; und 2,25 mM  $\text{MgCl}_2$ .

**[0088]** Ähnliche Testgemische wurden hergestellt, um den Einbau von rCTP, rGTP und rUTP zu messen. In jedem Fall wurde das rNTP radiomarkiert und war in 100 µM vorhanden, und die drei restlichen dNTPs (dATP, dGTP und dTTP für rCTP; dATP, dCTP und dTTP für rGTP; und dATP, dCTP und dGTP für rUTP) waren jeweils in 200 µM vorhanden. Als Standard wurde auch der Einbau des entsprechenden  $[\alpha\text{-}^{32}\text{P}]\text{dNTP}$  durch jedes Enzym gemessen. Das Testgemisch für diese Tests entsprach demjenigen für den vorstehenden rNTP-Einbautest, außer daß jedes  $[\alpha\text{-}^{32}\text{P}]\text{rNTP}$  durch 100 µM des entsprechenden  $[\alpha\text{-}^{32}\text{P}]\text{dNTP}$  ersetzt wurde. Rohe Lachssperma-DNA, 1 g/l, von Worthington Biochemical (Freehold, NJ) wurde durch Inkubation in 10 mM Tris-HCl, pH 7,2, 5 mM  $\text{MgCl}_2$ , bei 2°C–8°C für 96 Stunden aktiviert. EDTA und NaCl wurden anschließend zu 12,5 mM bzw. 0,1 M zugegeben. Die DNA wurde dann mit Phenol/Chloroform extrahiert und anschließend mit Ethanol gefällt und in 10 mM Tris, 1 mM EDTA, pH 7,5, resuspendiert. Die aktivierte DNA-Präparation wurde dann ge-

gen den gleichen Puffer dialysiert.

**[0089]** 45 µl jedes Reaktionsgemisches wurden in fünf 0,5 ml-Röhrchen (z. B. Eppendorf) für jeden der 5'-markierten Nucleotidvorläufer aliquotiert. So wurden G46D-Taq und G46D, E615G-Taq jeweils in doppelter Ausfertigung getestet, wobei ein Röhrchen für eine negative Kontrolle zurückbehalten wurde. Die Polymerisierungsreaktion in zwei Röhrchen jedes Testgemisches wurde mit 5 µl von entweder G46D-Taq-Polymerase (0,02 Einheiten) oder G46D, E615G-Taq (0,002 Einheiten) initiiert. Als Kontrolle für das Hintergrundniveau wurden 5 µl Enzymverdünnungspuffer anstelle des Enzyms zu der negativen Kontrollreaktion zugegeben.

**[0090]** Jede Reaktion wurde mit Hilfe eines Vortexgeräts kurz gemischt und für 10 Minuten bei 75°C inkubiert. Die Reaktionen wurden durch Zugabe von 10 µl 60 mM EDTA gestoppt und auf Eis gelagert. Für jede Probe wurden 50 µl-Aliquots der 60 µl-Reaktion mit 1 ml 2 mM EDTA, 50 µg/ml gescherter Lachssperma-DNA verdünnt. Die DNA wurde mit TCA durch Standardverfahren gefällt und auf GF/C-Filterscheiben (Whatman, Kent, England) gesammelt. Die Menge des eingebauten [ $\alpha$ -<sup>32</sup>P] markierten Nucleotids oder Ribonucleotids wurde mittels Flüssigkeitsszintillationsspektroskopie quantifiziert, und anschließend wurde die Anzahl der eingebauten pMole berechnet. Die Anzahl der pMole für jedes rNTP, die durch jedes Enzym eingebaut wurde, wurde auf die Anzahl der pMole des entsprechenden [ $\alpha$ -<sup>32</sup>P]dNTP, die durch jedes Enzym eingebaut wurde, normiert. Die erhaltenen Daten sind nachstehend gezeigt.

Einbauverhältnis von rNTP zu dNTP für G46D- und G46D, E615G-Taq

Enzym	Eingebaute pMole (Prozent)							
	dATP	rATP	dCTP	rCTP	dGTP	rGTP	dTTP	rUTP
G46D	27.74	0.052	34.6	0.76	36.94	0.133	28.79	0
	(100%)	(0.18%)	(100%)	(0.22%)	(100%)	(0.36%)	(100%)	(0)
G46D, E615G	0.67	1.41	2.82	5.33	3.27	5.96	0.688	0.545
	(100%)	(210%)	(100%)	(189%)	(100%)	(181%)	(100%)	(79%)

**[0091]** Diese Ergebnisse zeigen, daß G46D, E615G eine größere Effizienz für den Einbau von Ribonucleotiden aufweist, die mehr als 500-fach effizienter ist als für G46D (z. B. für rGTP: 181:0,36 = 502-fach; für rCTP: 189:0,22 = 859-fach; und für rATP: 210:0,18 = 1166-fach effizienter). Folglich stellte eine Missense-Mutation in dem Polymerase-Gen in Codon 615 einen neuen Phänotyp bereit: eine thermostabile DNA-Polymerase mit der Fähigkeit zum effizienten Einbau von Ribonucleotiden zusätzlich zu Desoxyribonucleotiden.

### Beispiel III

Test zum Vergleich der Effizienz des Einbaus von 3'-Desoxy-ATP (Cordycepin)

**[0092]** Die relative Fähigkeit von G46D-; G46D, E615G-; G46D, E615G, F667Y-; und G46D, F667Y-Taq zum Einbau von 3'-Desoxyadenosin-5'-triphosphat (Cordycepintriphosphat) wurde bestimmt, indem die Menge an [ $\alpha$ -<sup>32</sup>P]Cordycepintriphosphat gemessen wurde, die jedes Enzym bei einer begrenzenden Enzymkonzentration in eine aktivierte Lachssperma-DNA-Matrize einbauen konnte. Um den Einbau von [ $\alpha$ -<sup>32</sup>P]Cordycepintriphosphat zu messen, wurde der Test so aufgebaut, daß die Endkonzentrationen in einer 50 µl-Reaktion wie folgt waren: 12,5 µg aktivierte Lachssperma-DNA; jeweils 200 µM dCTP, dGTP und dTTP; 50 µM dATP (Perkin Elmer); 50 µM [ $\alpha$ -<sup>32</sup>P]-3'-dATP/3'-dATP (New England Nuclear, Sigma); 1 mM  $\beta$ -Mercaptoethanol; 25 mM N-Tris[hydroxymethyl]methyl-3-amino-propansulfonsäure (TAPS), pH 9,5, 20°C; 55 mM KCl; und 2,25 mM MgCl<sub>2</sub>.

**[0093]** 45 µl jedes Reaktionsgemisches wurden in neun 0,5 ml-Röhrchen aliquotiert, so daß jede Reaktion mit entweder G46D-; G46D, E615G-; G46D, E615G, F667Y-; oder G46D, F667Y-Taq in doppelter Ausfertigung durchgeführt wird, wobei ein Röhrchen für eine Kontrolle ohne Enzym zurückbehalten wurde. Die Polymerisierungsreaktion in zwei Röhrchen des Testgemisches wurde mit 5 µl (0,058 Einheiten) der G46D-Taq-Polymerase initiiert. Dieselbe Reaktion wurde für G46D, E615-Taq (0,0025 Einheiten), G46D, E615G, F667Y-Taq (0,0034 Einheiten) oder G46D, F66Y-Taq (0,083 Einheiten) durchgeführt. Als Kontrolle für das Hintergrundniveau wurde die Reaktion in dem verbleibenden Röhrchen mit Enzymverdünnungspuffer anstelle des Enzyms initiiert.

**[0094]** Jede Reaktion wurde mit Hilfe eines Vortexgeräts kurz gemischt und für 10 Minuten bei 75°C inkubiert. Die Reaktionen wurden durch Zugabe von 10 µl 60 mM EDTA gestoppt und auf Eis gelagert. Für jede Probe

wurden 50 µl-Aliquots der 60 µl-Reaktion mit 1 ml 2 mM EDTA, 50 µg/ml gescherter Lachssperma-DNA verdünnt. Die DNA wurde mit TCA durch Standardverfahren gefällt und auf GF/C-Filterscheiben (Whatman, Kent, England) gesammelt. Die Menge des eingebauten [ $\alpha$ - $^{32}$ P] markierten Nucleotids wurde mittels Flüssigkeitsszintillationsspektroskopie quantifiziert, und anschließend wurde die Anzahl der eingebauten pMole berechnet. Die Anzahl der pMole an [ $\alpha$ - $^{32}$ P]Cordycepintriphosphat, die durch jedes Enzym eingebaut wurde, wurde durch die in dem Test verwendete Anzahl der Einheiten jedes Enzyms dividiert, um die eingebauten pMole pro Enzymeinheit zu erhalten. Ein Tabelle dieser Daten ist nachstehend gezeigt.

Einbau von [ $\alpha$ - $^{32}$ P]Cordycepin durch G46D-; G46D, E615G-; G46D, E615G, F667Y-; und G46D, F667Y-Taq

Enzym	Eingebaute pMole pro Enzymeinheit
G46D	0.221
G46D, E615G	1.56
G46D, E615G, F667Y	893.6
G46D, F667Y	0.74

**[0095]** Diese Ergebnisse zeigen, daß sowohl die E615G- als auch die F667Y-Mutation für den effizienten Einbau des Cordycepinmoleküls in die DNA erforderlich sind.

#### Beispiel IV

DNA-Sequenzierung mit alkalischer Spaltung unter Verwendung der G46D, E615G-Taq-DNA-Polymerase

**[0096]** Dieses Beispiel veranschaulicht die Anwendung der erfindungsgemäßen modifizierten Polymerase bei der Sequenzierung mit alkalischer Spaltung unter Verwendung einer teilweise mit rNTP substituierten DNA. Das Verhältnis von rNTP zu dNTP in den Reaktionsgemischen lag zwischen 1 : 80 und 1 : 8. Die Primerverlängerungsreaktionen wurden in einem Puffer durchgeführt, der aus 50 mM Bicin (N,N-Bis-(2-hydroxyethyl)glycin; pH 8,3), 25 mM KOAc und 2,5 mM MgCl<sub>2</sub> bestand. Vier Einzelreaktionen, eine für jedes der vier rNTPs, wurden durchgeführt. Jede Reaktion (50 µl) enthielt jeweils 200 µM dATP, dCTP, dGTP und dTTP (Perkin Elmer) und 0,09 pMol einer einzelsträngigen M13mp18-DNA-Matrize (Perkin Elmer), angelagert an 5'-[ $^{32}$ P]-markiertes DG48 (Lawyer et al., PCR Methods and Applications 2 (1993), 275–287). Die Reaktionen enthielten auch 2,5 µM, 2,5 µM, 2,5 µM oder 25 µM rATP, rCTP, rGTP bzw. rUTP.

**[0097]** Jede der vier Reaktionen wurde durch Zugabe von 7 Einheiten G46D, E615G-Taq-DNA-Polymerase initiiert und bei 75°C für 10 Minuten inkubiert. Die Reaktionen wurden durch Zugabe von 10 µl 60 mM EDTA gestoppt und auf Eis gestellt. 20 µl jeder Reaktion wurden zu 80 µl 50 µM Bicin (pH 8,3), 25 mM KOAc und 2,5 mM MgCl<sub>2</sub> zugegeben. Spaltprodukte wurden durch Zugabe von 7 µl 1 N NaOH und Inkubation für 15 Minuten bei 98°C erzeugt. Die Reaktionen wurden durch Zugabe von 7 µl 1 N HCl neutralisiert. Jede Reaktion wurde durch die Zugabe von 312 µl 95% Ethanol und 10 µl 3 M Natriumacetat (pH 4,8) ausgefällt. Die Reaktionen wurden in einer Mikrozentrifuge für 15 Minuten zentrifugiert, um den Niederschlag zu sammeln, der Überstand wurde entfernt, die Pellets wurden mit 500 µl 70% Ethanol gewaschen und getrocknet. Jedes Pellet wurde in 5 µl 0,5 × Stopppuffer (erhältlich von Perkin Elmer, Norwalk, CT; enthält 95% Formamid, 20 mM EDTA und 0,05% Bromphenolblau) resuspendiert, bei 98°C für 3 Minuten erhitzt und direkt auf ein 6% Polyacrylamid/8 M Harnstoff-DNA-Sequenzierungsgel, das vorher einer Elektrophorese unterworfen worden war, geladen und einer Elektrophorese unterzogen. Das Gel wurde getrocknet und einem Röntgenfilm ausgesetzt. Der erhaltene Film zeigte ein deutliches leiterförmiges Sequenzmuster, das über 100 Basen der richtigen Sequenz bereitstellte.

#### Beispiel V

DNA-Sequenzierung unter Verwendung der G46D E615G F667Y-Taq-DNA-Polymerase und von 3'-Desoxy-Nucleotidtriphosphaten

**[0098]** Dieses Beispiel veranschaulicht die Anwendung der modifizierten Polymerase G46D, E615G, F667Y-Taq bei der DNA-Sequenzierung unter Verwendung von 3'-Desoxy-Nucleotidtriphosphaten. Dieses Experiment wurde unter Verwendung von 3'-Desoxy-ATP durchgeführt; es könnte jedoch auch auf die Verwendung mit anderen 3'-Desoxy-Nucleotiden übertragen werden. Die Primerverlängerungsreaktionen wurden in einem Puffer durchgeführt, der aus 50 µM Bicin (pH 8,3), 25 µM KOAc und 2,5 mM MgCl<sub>2</sub> bestand. Jede Reaktion (50 µl) enthielt jeweils 200 µM dATP, dCTP, dGTP und dTTP (Perkin Elmer) und 0,09 pMol einer einzelsträngigen M13mp18-DNA-Matrize (Perkin Elmer), angelagert an 5'-[ $^{32}$ P] markiertes DG48 (Lawyer et al., PCR

Methods and Applications 2 (1993), 275–287). Die Reaktionen enthielten auch 0  $\mu\text{M}$ , 0,1  $\mu\text{M}$ , 0,25  $\mu\text{M}$ , 0,5  $\mu\text{M}$ , 1  $\mu\text{M}$  oder 5  $\mu\text{M}$  3'-Desoxy-ATP.

**[0099]** Jede der Reaktionen wurde durch Zugabe von 7 Einheiten G46D, E615G, F667Y-Taq-DNA-Polymerase initiiert und bei 75°C für 10 Minuten inkubiert. Die Reaktionen wurden durch Zugabe von 10  $\mu\text{l}$  60 mM EDTA gestoppt und auf Eis gestellt. 30  $\mu\text{l}$  jeder Reaktion wurden mit Ethanol gefällt und in Stopppuffer resuspendiert, bei 98°C für 3 Minuten erhitzt und direkt auf ein 6% Polyacrylamid/8 M Harnstoff-DNA-Sequenzierungsgel, das vorher einer Elektrophorese unterworfen worden war, geladen und einer Elektrophorese unterzogen. Das Gel wurde getrocknet und einem Röntgenfilm ausgesetzt. Die Bahnen, welche die in Anwesenheit von Cordycepin durchgeführten Reaktionen enthielten, zeigten deutlich erkennbare leiterförmige Terminationsmuster. Die Bahnen, welche das meiste Cordycepin enthielten, d. h. 5  $\mu\text{M}$ , zeigten ein leiterförmiges Terminationsmuster, wobei die Bahnen im Durchschnitt kürzer waren als in den Bahnen, in denen die Cordycepin-Konzentrationen niedriger waren. Die Bahn, welche die in Abwesenheit von Cordycepin durchgeführte Reaktion enthielt, zeigte meistens das Produkt vollständiger Länge und kein leiterförmiges Terminationsmuster. Diese Ergebnisse zeigen, daß das mutierte Enzym zum Einbau von Cordycepin fähig ist und daß der Einbau dieses Moleküls in ein Primerverlängerungsprodukt eine Termination bewirkt. Dieses Verfahren könnte auch angewendet werden, um mit 3'-Desoxy-CTP, 3'-Desoxy-GTP und 3'-Desoxy-UTP ein leiterförmiges DNA-Sequenzierungsmuster zu erzeugen.

### Beispiel VI

#### Farbstoff-Primer-PCR-Sequenzierung mit der G46D, E615G-Taq-DNA-Polymerase

**[0100]** Dieses Beispiel zeigt die Anwendung der erfindungsgemäßen modifizierten Polymerase bei der Farbstoff-Primer-Sequenzierung unter Verwendung von Ribonucleosidtriphosphaten (rNTPs) in einer PCR und bei einem Verhältnis von rNTP : dNTP von nicht mehr als 1 : 30. Vier Einzelreaktionen, eine für jedes der rNTPs, wurden durchgeführt. Die PCR-Sequenzierungsreaktionen wurden in einem Puffer durchgeführt, der aus 25 mM Tris-HCl (pH 9), 5,0 mM  $\text{MgCl}_2$  und 10% Glycerin (Vol./Vol.) bestand. Jede Reaktion enthielt auch jeweils 500  $\mu\text{M}$  dATP, dCTP, dGTP und dTTP (Perkin Elmer),  $5 \times 10^6$  Kopien/ $\mu\text{l}$  pBSM13+-Plasmid (Stratagene)-Matrize, linearisiert mit XmnI-Restriktionsendonuclease, und 0,05 Einheiten/ $\mu\text{l}$  G46D, E615G-Taq-DNA-Polymerase. Ribo-ATP-Reaktionen (10  $\mu\text{l}$ ) enthielten 2,5  $\mu\text{M}$  ATP (Pharmacia Biotech), 0,1  $\mu\text{M}$  JOE-M13-Farbstoff-Rückwärtsprimer (Perkin Elmer) und 0,1  $\mu\text{M}$  Primer ASC46 (5'-CGCCATTCGCCATTCAG). Ribo-CTP-Reaktionen (10  $\mu\text{l}$ ) enthielten 2,5  $\mu\text{M}$  CTP (Pharmacia Biotech), 0,1  $\mu\text{M}$  FAM-M13-Farbstoff-Rückwärtsprimer (Perkin Elmer) und 0,1  $\mu\text{M}$  Primer ASC46. Ribo-GTP-Reaktionen (20  $\mu\text{l}$ ) enthielten 2,5  $\mu\text{M}$  GTP (Pharmacia Biotech), 0,1  $\mu\text{M}$  TAMRA-M13-Farbstoff-Rückwärtsprimer (Perkin Elmer) und 0,1  $\mu\text{M}$  Primer ASC46. Ribo-UTP-Reaktionen (20  $\mu\text{l}$ ) enthielten 16  $\mu\text{M}$  UTP (Pharmacia Biotech), 0,1  $\mu\text{M}$  ROX-M13-Farbstoff-Rückwärtsprimer (Perkin Elmer) und 0,1  $\mu\text{M}$  Primer ASC46.

**[0101]** Jede der vier Reaktionen wurde in eine vorerhitzte (75°C) Perkin Elmer-GeneAmp® PCR-System 9600-Thermocycler-Vorrichtung eingebracht und 30 Zyklen von 95°C für 10 Sekunden, 55°C für 10 Sekunden, 1 Minute erwärmen auf 65°C und 65°C für 5 Minuten unterzogen. Die rATP- und rCTP-Reaktionen erzeugten jeweils  $6 \times 10^{11}$  Kopien eines farbstoffmarkierten amplifizierten Produkts aus 300 Basenpaaren, und die rGTP- und UTP-Reaktionen erzeugten jeweils  $1,2 \times 10^{12}$  Kopien eines farbstoffmarkierten amplifizierten Produkts aus 300 Basenpaaren.

**[0102]** Zur Bestimmung der DNA-Sequenz der amplifizierten PCR-Produkte ohne das Erfordernis einer separaten enzymatischen DNA-Sequenzierungsreaktion wurden die Reaktionen wie folgt vereinigt, mit einer Base und Hitze behandelt, neutralisiert und ausgefällt. Jeweils 4  $\mu\text{l}$  der ATP- und CTP-Reaktionen und jeweils 8  $\mu\text{l}$  der GTP- und UTP-Reaktionen wurden vereinigt. 2  $\mu\text{l}$  0,25 M EDTA (pH 8,0) (10 mM Endkonzentration), 10  $\mu\text{l}$  1 M NaOH (200 mM Endkonzentration) und 14  $\mu\text{l}$   $\text{H}_2\text{O}$  wurden zu der vereinigten Reaktion zugegeben, die dann bei 95°C für 5 Minuten in einer GeneAmp® PCR-System 9600-Thermocycler-Vorrichtung inkubiert und mit 10  $\mu\text{l}$  1 M HCl neutralisiert wurde. Die vereinigte Reaktion wurde anschließend durch die Zugabe von 150  $\mu\text{l}$  95% Ethanol, gefolgt durch eine Inkubation bei 4°C für 15 Minuten, ausgefällt. Die Reaktion wurde dann in einer Mikrozentrifuge für 15 Minuten bei 4°C zentrifugiert, um den Niederschlag zu sammeln, und der Überstand wurde durch Absaugen entfernt. Das Pellet wurde mit 300  $\mu\text{l}$  70% Ethanol gewaschen und in einer Mikrozentrifuge für 15 Minuten zentrifugiert, der Überstand wurde durch Absaugen entfernt und das Pellet wurde getrocknet. Das Pellet wurde in 6  $\mu\text{l}$  Formamid, 50 mg/ml Blue-Dextran (in 25 mM EDTA) 5 : 1 (Vol./Vol.) resuspendiert und bei 90°C für 3 Minuten erhitzt. 1,5  $\mu\text{l}$  des resuspendierten Pellets wurden direkt auf ein 5% Long Ranger (FMC BioProducts), 6 M Harnstoff-Sequenzierungsgel geladen, das vorher einer Elektrophorese unterworfen worden war. Das Pellet wurde dann einer Elektrophorese unterzogen und in einem Perkin El-

mer-ABI Prism™ 377-DNA-Sequenzanalysengerät gemäß den Anweisungen des Herstellers analysiert. Die automatische Basenbezeichnung durch die Perkin Elmer-ABI-Prism™-Sequenzierungsanalyse-Software ergab eine Genauigkeit von mehr als 99% für die DNA-Sequenzbestimmung des mittels PCR amplifizierten Produkts aus 300 Basenpaaren.

SEQUENZPROTOKOLL

(1) ALLGEMEINE INFORMATION:

(i) ANMELDER:

(A) NAME: F. Hoffmann-La Roche Ltd.

(B) STRASSE: Grenzacherstrasse 124

(C) STADT: Basel

(D) LAND: BS

(E) STAAT: Schweiz

(F) POSTLEITZAHL: CH-4070

(G) TELEFON: (0) 61 688 24 03

(H) TELEFAX: (0) 61 688 13 95

(I) TELEX: 962292/965512 hlr ch

(ii) TITEL DER ERFINDUNG: Modifizierte thermostabile DNA-Polymerase

(iii) ANZAHL DER SEQUENZEN: 8

(iv) COMPUTERLESBARE FORM:

(A) MEDIUMTYP: Diskette

(B) COMPUTER: IBM PC-kompatibel

(C) BETRIEBSSYSTEM: PC-DOS/MS-DOS

(D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

(vi) DATEN DER FRÜHEREN ANMELDUNG:

(A) ANMELDUNGSNUMMER: US 60/023,376

(B) EINREICHUNGSDATUM: 6. August 1996

(2) INFORMATION FÜR SEQ ID NR: 1:

(i) SEQUENZEIGENSCHAFTEN:

(A) LÄNGE: 7 Aminosäuren

(B) TYP: Aminosäure

(D) TOPOLOGIE: Linear

(ii) MOLEKÜLTYP: Peptid

(ix) MERKMAL:

(A) BEZEICHNUNG/SCHLÜSSEL: Peptid

(B) LAGE: 4

(D) ANDERE INFORMATION: /Marker = Xaa

/Anmerkung = "wobei Xaa ein beliebiger Aminosäurerest, aber nicht Glu ist"

(ix) MERKMAL:

(A) BEZEICHNUNG/SCHLÜSSEL: Peptid

(B) LAGE: 7

(D) ANDERE INFORMATION: /Marker = Xaa

/Anmerkung = "wobei Xaa Ile oder Val ist"

(xi) SEQUENZBESCHREIBUNG: SEQ ID NR: 1:

Ser Gln Ile Xaa Leu Arg Xaa  
1 5

(2) INFORMATION FÜR SEQ ID NR: 2:

(i) SEQUENZEIGENSCHAFTEN:

(A) LÄNGE: 7 Aminosäuren

(B) TYP: Aminosäure

(D) TOPOLOGIE: Linear

(ii) MOLEKÜLTYP: Peptid

(ix) MERKMAL:

(A) BEZEICHNUNG/SCHLÜSSEL: Peptid

(B) LAGE: 7

(D) ANDERE INFORMATION: /Marker = Xaa

/Anmerkung = "wobei Xaa Ile oder Val ist"

(xi) SEQUENZBESCHREIBUNG: SEQ ID NR: 2:

Ser Gln Ile Glu Leu Arg Xaa  
1 5

(2) INFORMATION FÜR SEQ ID NR: 3:

(i) SEQUENZEIGENSCHAFTEN:

(A) LÄNGE: 7 Aminosäuren

(B) TYP: Aminosäure

(D) TOPOLOGIE: Linear

(ii) MOLEKÜLTYP: Peptid

(xi) SEQUENZBESCHREIBUNG: SEQ ID NR: 3:

Ser Gln Ile Glu Leu Arg Val  
1 5

(2) INFORMATION FÜR SEQ ID NR: 4:

(i) SEQUENZEIGENSCHAFTEN:

(A) LÄNGE: 7 Aminosäuren

(B) TYP: Aminosäure

(D) TOPOLOGIE: Linear

(ii) MOLEKÜLTYP: Peptid

(xi) SEQUENZBESCHREIBUNG: SEQ ID NR: 4:

Ser Gln Ile Glu Leu Arg Ile  
1 5

(2) INFORMATION FÜR SEQ ID NR: 5:

(i) SEQUENZEIGENSCHAFTEN:

(A) LÄNGE: 15 Aminosäuren

(B) TYP: Aminosäure

(D) TOPOLOGIE: Linear

(ii) MOLEKÜLTYP: Peptid

(xi) SEQUENZBESCHREIBUNG: SEQ ID NR: 5:

Leu Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala His Leu Ser  
1 5 10 15

(2) INFORMATION FÜR SEQ ID NR: 6:

(i) SEQUENZEIGENSCHAFTEN:

(A) LÄNGE: 15 Aminosäuren

(B) TYP: Aminosäure

(D) TOPOLOGIE: Linear

(ii) MOLEKÜLTYP: Peptid

(xi) SEQUENZBESCHREIBUNG: SEQ ID NR: 6:

Leu Asp Tyr Ser Gln Ile Gly Leu Arg Val Leu Ala His Leu Ser  
1 5 10 15



(2) INFORMATION FÜR SEQ ID NR: 7:

(i) SEQUENZEIGENSCHAFTEN:

(A) LÄNGE: 2626 Basenpaare

(B) TYP: Nucleinsäure

(C) STRANGFORM: Doppelstrang

(D) TOPOLOGIE: Linear

(ii) MOLEKÜLTYP: DNA (genomisch)

(iii) HYPOTHETISCH: Nein

(iv) ANTISENSE: Nein

(vi) URSPRÜNGLICHE HERKUNFT:

(A) ORGANISMUS: *Thermus aquaticus*

(ix) MERKMAL:

(A) BEZEICHNUNG/SCHLÜSSEL: CDS

(B) LAGE: 121 .. 2616

(xi) SEQUENZBESCHREIBUNG: SEQ ID NR: 7:

```
AAGCTCAGAT CTACCTGCCT GAGGGCGTCC GGTTCAGCT GGCCTTCCC GAGGGGGAGA      60
GGGAGGCGTT TCTAAAAGCC CTTCAGGACG CTACCCGGGG GCGGGTGGTG GAAGGGTAAC      120
```

DE 697 30 926 T2 2006.02.23

ATG AGG GGG ATG CTG CCC CTC TTT GAG CCC AAG GGC CGG GTC CTC CTG	168
Met Arg Gly Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu Leu	
1 5 10 15	
GTG GAC GGC CAC CAC CTG GCC TAC CGC ACC TTC CAC GCC CTG AAG GGC	216
Val Asp Gly His His Leu Ala Tyr Arg Thr Phe His Ala Leu Lys Gly	
20 25 30	
CTC ACC ACC AGC CGG GGG GAG CCG GTG CAG GCG GTC TAC GGC TTC GCC	264
Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe Ala	
35 40 45	
AAG AGC CTC CTC AAG GCC CTC AAG GAG GAC GGG GAC GCG GTG ATC GTG	312
Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Asp Ala Val Ile Val	
50 55 60	
GTC TTT GAC GCC AAG GCC CCC TCC TTC CGC CAC GAG GCC TAC GGG GGG	360
Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Gly Gly	
65 70 75 80	
TAC AAG GCG GGC CGG GCC CCC ACG CCG GAG GAC TTT CCC CGG CAA CTC	408
Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln Leu	
85 90 95	
GCC CTC ATC AAG GAG CTG GTG GAC CTC CTG GGG CTG GCG CGC CTC GAG	456
Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Leu Ala Arg Leu Glu	
100 105 110	
GTC CCG GGC TAC GAG GCG GAC GAC GTC CTG GCC AGC CTG GCC AAG AAG	504
Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Ser Leu Ala Lys Lys	
115 120 125	
GCG GAA AAG GAG GGC TAC GAG GTC CGC ATC CTC ACC GCC GAC AAA GAC	552
Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Lys Asp	
130 135 140	
CTT TAC CAG CTC CTT TCC GAC CGC ATC CAC GTC CTC CAC CCC GAG GGG	600
Leu Tyr Gln Leu Leu Ser Asp Arg Ile His Val Leu His Pro Glu Gly	
145 150 155 160	
TAC CTC ATC ACC CCG GCC TGG CTT TGG GAA AAG TAC GGC CTG AGG CCC	648
Tyr Leu Ile Thr Pro Ala Trp Leu Trp Glu Lys Tyr Gly Leu Arg Pro	
165 170 175	
GAC CAG TGG GCC GAC TAC CGG GCC CTG ACC GGG GAC GAG TCC GAC AAC	696
Asp Gln Trp Ala Asp Tyr Arg Ala Leu Thr Gly Asp Glu Ser Asp Asn	
180 185 190	
CTT CCC GGG GTC AAG GGC ATC GGG GAG AAG ACG GCG AGG AAG CTT CTG	744
Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Arg Lys Leu Leu	
195 200 205	
GAG GAG TGG GGG AGC CTG GAA GCC CTC CTC AAG AAC CTG GAC CGG CTG	792
Glu Glu Trp Gly Ser Leu Glu Ala Leu Leu Lys Asn Leu Asp Arg Leu	
210 215 220	
AAG CCC GCC ATC CGG GAG AAG ATC CTG GCC CAC ATG GAC GAT CTG AAG	840
Lys Pro Ala Ile Arg Glu Lys Ile Leu Ala His Met Asp Asp Leu Lys	
225 230 235 240	

CTC	TCC	TGG	GAC	CTG	GCC	AAG	GTG	CGC	ACC	GAC	CTG	CCC	CTG	GAG	GTG	888
Leu	Ser	Trp	Asp	Leu	Ala	Lys	Val	Arg	Thr	Asp	Leu	Pro	Leu	Glu	Val	
				245					250					255		
GAC	TTC	GCC	AAA	AGG	CGG	GAG	CCC	GAC	CGG	GAG	AGG	CTT	AGG	GCC	TTT	936
Asp	Phe	Ala	Lys	Arg	Arg	Glu	Pro	Asp	Arg	Glu	Arg	Leu	Arg	Ala	Phe	
			260					265					270			
CTG	GAG	AGG	CTT	GAG	TTT	GGC	AGC	CTC	CTC	CAC	GAG	TTC	GGC	CTT	CTG	984
Leu	Glu	Arg	Leu	Glu	Phe	Gly	Ser	Leu	Leu	His	Glu	Phe	Gly	Leu	Leu	
		275					280					285				
GAA	AGC	CCC	AAG	GCC	CTG	GAG	GAG	GCC	CCC	TGG	CCC	CCG	CCG	GAA	GGG	1032
Glu	Ser	Pro	Lys	Ala	Leu	Glu	Glu	Ala	Pro	Trp	Pro	Pro	Pro	Glu	Gly	
	290					295					300					
GCC	TTC	GTG	GGC	TTT	GTG	CTT	TCC	CGC	AAG	GAG	CCC	ATG	TGG	GCC	GAT	1080
Ala	Phe	Val	Gly	Phe	Val	Leu	Ser	Arg	Lys	Glu	Pro	Met	Trp	Ala	Asp	
305					310					315					320	
CTT	CTG	GCC	CTG	GCC	GCC	GCC	AGG	GGG	GGC	CGG	GTC	CAC	CGG	GCC	CCC	1128
Leu	Leu	Ala	Leu	Ala	Ala	Ala	Arg	Gly	Gly	Arg	Val	His	Arg	Ala	Pro	
				325					330					335		
GAG	CCT	TAT	AAA	GCC	CTC	AGG	GAC	CTG	AAG	GAG	GCG	CGG	GGG	CTT	CTC	1176
Glu	Pro	Tyr	Lys	Ala	Leu	Arg	Asp	Leu	Lys	Glu	Ala	Arg	Gly	Leu	Leu	
			340					345					350			
GCC	AAA	GAC	CTG	AGC	GTT	CTG	GCC	CTG	AGG	GAA	GGC	CTT	GGC	CTC	CCG	1224
Ala	Lys	Asp	Leu	Ser	Val	Leu	Ala	Leu	Arg	Glu	Gly	Leu	Gly	Leu	Pro	
		355					360					365				
CCC	GGC	GAC	GAC	CCC	ATG	CTC	CTC	GCC	TAC	CTC	CTG	GAC	CCT	TCC	AAC	1272
Pro	Gly	Asp	Asp	Pro	Met	Leu	Leu	Ala	Tyr	Leu	Leu	Asp	Pro	Ser	Asn	
	370					375					380					
ACC	ACC	CCC	GAG	GGG	GTG	GCC	CGG	CGC	TAC	GGC	GGG	GAG	TGG	ACG	GAG	1320
Thr	Thr	Pro	Glu	Gly	Val	Ala	Arg	Arg	Tyr	Gly	Gly	Glu	Trp	Thr	Glu	
385					390					395					400	
GAG	GCG	GGG	GAG	CGG	GCC	GCC	CTT	TCC	GAG	AGG	CTC	TTC	GCC	AAC	CTG	1368
Glu	Ala	Gly	Glu	Arg	Ala	Ala	Leu	Ser	Glu	Arg	Leu	Phe	Ala	Asn	Leu	
				405					410					415		
TGG	GGG	AGG	CTT	GAG	GGG	GAG	GAG	AGG	CTC	CTT	TGG	CTT	TAC	CGG	GAG	1416
Trp	Gly	Arg	Leu	Glu	Gly	Glu	Glu	Arg	Leu	Leu	Trp	Leu	Tyr	Arg	Glu	
			420					425					430			
GTG	GAG	AGG	CCC	CTT	TCC	GCT	GTC	CTG	GCC	CAC	ATG	GAG	GCC	ACG	GGG	1464
Val	Glu	Arg	Pro	Leu	Ser	Ala	Val	Leu	Ala	His	Met	Glu	Ala	Thr	Gly	
		435					440					445				
GTG	CGC	CTG	GAC	GTG	GCC	TAT	CTC	AGG	GCC	TTG	TCC	CTG	GAG	GTG	GCC	1512
Val	Arg	Leu	Asp	Val	Ala	Tyr	Leu	Arg	Ala	Leu	Ser	Leu	Glu	Val	Ala	
	450					455					460					

GAG Glu 465	GAG Glu 465	ATC Ile	GCC Ala	CGC Arg	CTC Leu 470	GAG Glu	GCC Ala	GAG Glu	GTC Val	TTC Phe 475	CGC Arg	CTG Leu	GCC Ala	GGC Gly 480	CAC His	1560
CCC Pro	TTC Phe	AAC Asn	CTC Leu	AAC Asn 485	TCC Ser	CGG Arg	GAC Asp	CAG Gln	CTG Leu 490	GAA Glu	AGG Arg	GTC Val	CTC Leu	TTT Phe 495	GAC Asp	1608
GAG Glu	CTA Leu	GGG Gly	CTT Leu 500	CCC Pro	GCC Ala	ATC Ile	GGC Gly	AAG Lys 505	ACG Thr	GAG Glu	AAG Lys	ACC Thr	GGC Gly 510	AAG Lys	CGC Arg	1656
TCC Ser	ACC Thr	AGC Ser 515	GCC Ala	GCC Ala	GTC Val	CTG Leu	GAG Glu 520	GCC Ala	CTC Leu	CGC Arg	GAG Glu	GCC Ala 525	CAC His	CCC Pro	ATC Ile	1704
GTG Val	GAG Glu 530	AAG Lys	ATC Ile	CTG Leu	CAG Gln	TAC Tyr 535	CGG Arg	GAG Glu	CTC Leu	ACC Thr	AAG Lys 540	CTG Leu	AAG Lys	AGC Ser	ACC Thr	1752
TAC Tyr 545	ATT Ile	GAC Asp	CCC Pro	TTG Leu	CCG Pro 550	GAC Asp	CTC Leu	ATC Ile	CAC His	CCC Pro 555	AGG Arg	ACG Thr	GGC Gly	CGC Arg	CTC Leu 560	1800
CAC His	ACC Thr	CGC Arg	TTC Phe	AAC Asn 565	CAG Gln	ACG Thr	GCC Ala	ACG Thr	GCC Ala	ACG Thr	GGC Gly	AGG Arg	CTA Leu	AGT Ser	AGC Ser	1848
TCC Ser	GAT Asp	CCC Pro	AAC Asn 580	CTC Leu	CAG Gln	AAC Asn	ATC Ile	CCC Pro 585	GTC Val	CGC Arg	ACC Thr	CCG Pro	CTT Leu 590	GGG Gly	CAG Gln	1896
AGG Arg	ATC Ile	CGC Arg 595	CGG Arg	GCC Ala	TTC Phe	ATC Ile	GCC Ala 600	GAG Glu	GAG Glu	GGG Gly	TGG Trp	CTA Leu 605	TTG Leu	GTG Val	GCC Ala	1944
CTG Leu 610	GAC Asp	TAT Tyr	AGC Ser	CAG Gln	ATA Ile	GGG Gly 615	CTC Leu	AGG Arg	GTG Val	CTG Leu	GCC Ala 620	CAC His	CTC Leu	TCC Ser	GGC Gly	1992
GAC Asp 625	GAG Glu	AAC Asn	CTG Leu	ATC Ile	CGG Arg 630	GTC Val	TTC Phe	CAG Gln	GAG Glu	GGG Gly 635	CGG Arg	GAC Asp	ATC Ile	CAC His	ACG Thr 640	2040
GAG Glu	ACC Thr	GCC Ala	AGC Ser	TGG Trp 645	ATG Met	TTC Phe	GGC Gly	GTC Val	CCC Pro 650	CGG Arg	GAG Glu	GCC Ala	GTG Val	GAC Asp 655	CCC Pro	2088
CTG Leu	ATG Met	CGC Arg	CGG Arg 660	GCG Ala	GCC Ala	AAG Lys	ACC Thr	ATC Ile	AAC Asn	TTC Phe 665	GGG Gly	GTC Val	CTC Leu 670	TAC Tyr	GGC Gly	2136
ATG Met	TCG Ser	GCC Ala 675	CAC His	CGC Arg	CTC Leu	TCC Ser	CAG Gln 680	GAG Glu	CTA Leu	GCC Ala	ATC Ile	CCT Pro 685	TAC Tyr	GAG Glu	GAG Glu	2184
GCC Ala 690	CAG Gln	GCC Ala	TTC Phe	ATT Ile	GAG Glu	CGC Arg 695	TAC Tyr	TTT Phe	CAG Gln	AGC Ser	TTC Phe 700	CCC Pro	AAG Lys	GTG Val	CGG Arg	2232

GCC TGG ATT GAG AAG ACC CTG GAG GAG GGC AGG AGG CGG GGG TAC GTG	2280
Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Arg Arg Gly Tyr Val	
705 710 715 720	
GAG ACC CTC TTC GGC CGC CGC CGC TAC GTG CCA GAC CTA GAG GCC CGG	2328
Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Glu Ala Arg	
725 730 735	
GTG AAG AGC GTG CGG GAG GCG GCC GAG CGC ATG GCC TTC AAC ATG CCC	2376
Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met Pro	
740 745 750	
GTC CAG GGC ACC GCC GCC GAC CTC ATG AAG CTG GCT ATG GTG AAG CTC	2424
Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys Leu	
755 760 765	
TTC CCC AGG CTG GAG GAA ATG GGG GCC AGG ATG CTC CTT CAG GTC CAC	2472
Phe Pro Arg Leu Glu Glu Met Gly Ala Arg Met Leu Leu Gln Val His	
770 775 780	
GAC GAG CTG GTC CTC GAG GCC CCA AAA GAG AGG GCG GAG GCC GTG GCC	2520
Asp Glu Leu Val Leu Glu Ala Pro Lys Glu Arg Ala Glu Ala Val Ala	
785 790 795 800	
CGG CTG GCC AAG GAG GTC ATG GAG GGG GTG TAT CCC CTG GCC GTG CCC	2568
Arg Leu Ala Lys Glu Val Met Glu Gly Val Tyr Pro Leu Ala Val Pro	
805 810 815	
CTG GAG GTG GAG GTG GGG ATA GGG GAG GAC TGG CTC TCC GCC AAG GAG	2616
Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys Glu	
820 825 830	
TGATACCACC	2626

(2) INFORMATION FÜR SEQ ID NR: 8:

(i) SEQUENZEIGENSCHAFTEN:

(A) LÄNGE: 832 Aminosäuren

(B) TYP: Aminosäure

(D) TOPOLOGIE: Linear

(ii) MOLEKÜLTYP: Protein

(xi) SEQUENZBESCHREIBUNG: SEQ ID NR: 8:

Met	Arg	Gly	Met	Leu	Pro	Leu	Phe	Glu	Pro	Lys	Gly	Arg	Val	Leu	Leu
1				5					10					15	
Val	Asp	Gly	His	His	Leu	Ala	Tyr	Arg	Thr	Phe	His	Ala	Leu	Lys	Gly
			20					25					30		
Leu	Thr	Thr	Ser	Arg	Gly	Glu	Pro	Val	Gln	Ala	Val	Tyr	Gly	Phe	Ala
		35					40					45			
Lys	Ser	Leu	Leu	Lys	Ala	Leu	Lys	Glu	Asp	Gly	Asp	Ala	Val	Ile	Val
	50					55					60				

Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Gly Gly  
 65 70 75 80  
 Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln Leu  
 85 90 95  
 Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Leu Ala Arg Leu Glu  
 100 105 110  
 Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Ser Leu Ala Lys Lys  
 115 120 125  
 Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Lys Asp  
 130 135 140  
 Leu Tyr Gln Leu Leu Ser Asp Arg Ile His Val Leu His Pro Glu Gly  
 145 150 155 160  
 Tyr Leu Ile Thr Pro Ala Trp Leu Trp Glu Lys Tyr Gly Leu Arg Pro  
 165 170 175  
 Asp Gln Trp Ala Asp Tyr Arg Ala Leu Thr Gly Asp Glu Ser Asp Asn  
 180 185 190  
 Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Arg Lys Leu Leu  
 195 200 205  
 Glu Glu Trp Gly Ser Leu Glu Ala Leu Leu Lys Asn Leu Asp Arg Leu  
 210 215 220  
 Lys Pro Ala Ile Arg Glu Lys Ile Leu Ala His Met Asp Asp Leu Lys  
 225 230 235 240  
 Leu Ser Trp Asp Leu Ala Lys Val Arg Thr Asp Leu Pro Leu Glu Val  
 245 250 255  
 Asp Phe Ala Lys Arg Arg Glu Pro Asp Arg Glu Arg Leu Arg Ala Phe  
 260 265 270  
 Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly Leu Leu  
 275 280 285  
 Glu Ser Pro Lys Ala Leu Glu Glu Ala Pro Trp Pro Pro Pro Glu Gly  
 290 295 300  
 Ala Phe Val Gly Phe Val Leu Ser Arg Lys Glu Pro Met Trp Ala Asp  
 305 310 315 320  
 Leu Leu Ala Leu Ala Ala Ala Arg Gly Gly Arg Val His Arg Ala Pro  
 325 330 335  
 Glu Pro Tyr Lys Ala Leu Arg Asp Leu Lys Glu Ala Arg Gly Leu Leu  
 340 345 350  
 Ala Lys Asp Leu Ser Val Leu Ala Leu Arg Glu Gly Leu Gly Leu Pro  
 355 360 365  
 Pro Gly Asp Asp Pro Met Leu Leu Ala Tyr Leu Leu Asp Pro Ser Asn  
 370 375 380

Thr Thr Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly Glu Trp Thr Glu  
 385 390 395 400  
 Glu Ala Gly Glu Arg Ala Ala Leu Ser Glu Arg Leu Phe Ala Asn Leu  
 405 410 415  
 Trp Gly Arg Leu Glu Gly Glu Glu Arg Leu Leu Trp Leu Tyr Arg Glu  
 420 425 430  
 Val Glu Arg Pro Leu Ser Ala Val Leu Ala His Met Glu Ala Thr Gly  
 435 440 445  
 Val Arg Leu Asp Val Ala Tyr Leu Arg Ala Leu Ser Leu Glu Val Ala  
 450 455 460  
 Glu Glu Ile Ala Arg Leu Glu Ala Glu Val Phe Arg Leu Ala Gly His  
 465 470 475 480  
 Pro Phe Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg Val Leu Phe Asp  
 485 490 495  
 Glu Leu Gly Leu Pro Ala Ile Gly Lys Thr Glu Lys Thr Gly Lys Arg  
 500 505 510  
 Ser Thr Ser Ala Ala Val Leu Glu Ala Leu Arg Glu Ala His Pro Ile  
 515 520 525  
 Val Glu Lys Ile Leu Gln Tyr Arg Glu Leu Thr Lys Leu Lys Ser Thr  
 530 535 540  
 Tyr Ile Asp Pro Leu Pro Asp Leu Ile His Pro Arg Thr Gly Arg Leu  
 545 550 555 560  
 His Thr Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly Arg Leu Ser Ser  
 565 570 575  
 Ser Asp Pro Asn Leu Gln Asn Ile Pro Val Arg Thr Pro Leu Gly Gln  
 580 585 590  
 Arg Ile Arg Arg Ala Phe Ile Ala Glu Glu Gly Trp Leu Leu Val Ala  
 595 600 605  
 Leu Asp Tyr Ser Gln Ile Gly Leu Arg Val Leu Ala His Leu Ser Gly  
 610 615 620  
 Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly Arg Asp Ile His Thr  
 625 630 635 640  
 Glu Thr Ala Ser Trp Met Phe Gly Val Pro Arg Glu Ala Val Asp Pro  
 645 650 655  
 Leu Met Arg Arg Ala Ala Lys Thr Ile Asn Phe Gly Val Leu Tyr Gly  
 660 665 670  
 Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile Pro Tyr Glu Glu  
 675 680 685  
 Ala Gln Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys Val Arg  
 690 695 700

Ala	Trp	Ile	Glu	Lys	Thr	Leu	Glu	Glu	Gly	Arg	Arg	Arg	Gly	Tyr	Val
705					710					715					720
Glu	Thr	Leu	Phe	Gly	Arg	Arg	Arg	Tyr	Val	Pro	Asp	Leu	Glu	Ala	Arg
				725					730					735	
Val	Lys	Ser	Val	Arg	Glu	Ala	Ala	Glu	Arg	Met	Ala	Phe	Asn	Met	Pro
			740					745					750		
Val	Gln	Gly	Thr	Ala	Ala	Asp	Leu	Met	Lys	Leu	Ala	Met	Val	Lys	Leu
		755					760					765			
Phe	Pro	Arg	Leu	Glu	Glu	Met	Gly	Ala	Arg	Met	Leu	Leu	Gln	Val	His
	770					775					780				
Asp	Glu	Leu	Val	Leu	Glu	Ala	Pro	Lys	Glu	Arg	Ala	Glu	Ala	Val	Ala
785					790					795					800
Arg	Leu	Ala	Lys	Glu	Val	Met	Glu	Gly	Val	Tyr	Pro	Leu	Ala	Val	Pro
				805					810					815	
Leu	Glu	Val	Glu	Val	Gly	Ile	Gly	Glu	Asp	Trp	Leu	Ser	Ala	Lys	Glu
			820					825					830		

### Patentansprüche

1. Thermostabiles DNA-Polymeraseenzym, welches die Aminosäuresequenz SerGlnIleXaaLeuArgXaa (SEQ ID NR: 1) umfasst, wobei „Xaa“ an Position 4 der Sequenz ein beliebiger Aminosäurerest ist, aber kein Glutaminsäurerest (Glu), und wobei „Xaa“ an Position 7 der Sequenz ein Valinrest (Val) oder ein Isoleucinrest (Ile) ist, **dadurch gekennzeichnet**, dass diese eine reduzierte Diskrimination gegen den Einbau eines unkonventionellen Nucleotids aufweist, im Vergleich zu einer natürlich vorkommenden thermostabilen DNA-Polymerase, wobei diese reduzierte Diskrimination dadurch gekennzeichnet ist, dass die Fähigkeit der Polymerase zum Einbau eines unkonventionellen Nucleotids relativ zur Fähigkeit der entsprechenden nativen Form der Polymerase zur Inkorporation dieses unkonventionellen Nucleotids mindestens 20-fach gesteigert ist.

2. Thermostabiles DNA-Polymeraseenzym, das dadurch gekennzeichnet ist, dass es ein rekombinantes Derivat einer natürlich vorkommenden thermostabilen DNA-Polymerase ist, wobei diese natürlich vorkommende thermostabile DNA-Polymerase das Aminosäuresequenzmotiv SerGlnIleGluLeuArgXaa (SEQ ID NR: 2) umfasst, wobei „Xaa“ an Position 7 der Sequenz ein Valinrest (Val) oder ein Isoleucinrest (Ile) ist; wobei das rekombinante Derivat modifiziert wurde, um eine Aminosäure anders als Glutaminsäure (Glu) an Position 4 des Sequenzmotivs zu erhalten; und wobei dieses rekombinante Derivat reduzierte Diskrimination gegen den Einbau eines unkonventionellen Nucleotids aufweist im Vergleich zu der natürlich vorkommenden thermostabilen DNA-Polymerase, wobei diese reduzierte Diskriminierung dadurch gekennzeichnet ist, dass die Fähigkeit dieser Polymerase zum Einbau eines unkonventionellen Nucleotids relativ zur Fähigkeit der entsprechenden nativen Form der Polymerase zum Einbau dieses unkonventionellen Nucleotids mindestens 20-fach gesteigert ist.

3. Thermostabiles DNA-Polymeraseenzym nach Anspruch 1 oder 2, dadurch gekennzeichnet, dass die Aminosäure in Position 4 Glycin (Gly) ist.

4. Thermostabile DNA-Polymerase nach einem der Ansprüche 1 bis 3, dadurch gekennzeichnet, dass diese Polymerase eine Aktivität besitzt zur Verwendung in einer DNA-Sequenzierungsreaktion, wobei die Verwendung von Didesoxynucleotiden vermieden wird, wobei diese Reaktion ein unkonventionelles Nucleotid umfasst, welches vorzugsweise ein Ribonucleosid-Triphosphat ist, und ein entsprechendes konventionelles Nucleotid in einem Verhältnis von 1 : 1 oder weniger.

5. Thermostabile DNA-Polymerase nach einem der Ansprüche 1 bis 3, dadurch gekennzeichnet, dass diese Polymerase eine Aktivität zur Verwendung in DNA-Sequenzierungsreaktionen hat, wobei die Verwendung von Didesoxynucleotiden vermieden wird, wobei diese Reaktion ein unkonventionelles Nucleotid umfasst, welches ein Ribonucleosid-Triphosphat ist, in einer Konzentration von weniger als etwa 100 µM und ein entsprechendes konventionelles Nucleotid, welches in einer Konzentration von mehr als etwa 100 µM vorliegt.



6. Thermostabiles DNA-Polymeraseenzym nach einem der Ansprüche 1 bis 5, welches ein rekombinantes Derivat eines natürlich vorkommenden thermostabilen DNA-Polymeraseenzym ist von einem Organismus ausgewählt aus der Gruppe bestehend aus *Thermus aquaticus*, *Thermus caldophilus*, *Thermus chliarophilus*, *Thermus filiformis*, *Thermus flavus*, *Thermus oshimai*, *Thermus ruber*, *Thermus scotoductus*, *Thermus silvanus*, *Thermus species Z05*, *Thermus species sps17*, *Thermus thermophilus*, *Thermotoga maritima*, *Thermotoga neopolitana*, *Thermosiphon africanus*, *Anaerocellum thermophilum*, *Bacillus caldodenax* und *Bacillus stearothermophilus*.

7. Thermostabiles DNA-Polymeraseenzym nach einem der Ansprüche 2 bis 5, welches ein rekombinantes Derivat einer natürlich vorkommenden thermostabilen *Thermus species* DNA-Polymerase ist, vorzugsweise eine Taq-DNA-Polymerase oder eine homologe Polymerase davon, mehr bevorzugt eine thermostabile DNA-Polymerase umfassend die Aminosäuresequenz LeuAspTyrSerGlnIleGluLeuArgValLeuAlaHisLeuSer (SEQ ID NR: 5), am meisten bevorzugt die thermostabile DNA-Polymerase mit der Aminosäuresequenz von SEQ ID NR: 8 und modifizierte Formen davon, wie die mutanten Formen G46D und/oder F667Y.

8. Thermostabiles DNA-Polymeraseenzym nach einem der Ansprüche 1 bis 5, welches mindestens etwa 39%, vorzugsweise mindestens etwa 60%, mehr bevorzugt mindestens etwa 80% Sequenzhomologie zu der Aminosäuresequenz der Taq-DNA-Polymerase (SEQ ID NR: 7) hat.

9. Nucleinsäuresequenz, die ein thermostabiles DNA-Polymeraseenzym nach einem der Ansprüche 1 bis 8 codiert.

10. Vektor umfassend eine Nucleinsäuresequenz, die ein thermostabiles DNA-Polymeraseenzym nach einem der Ansprüche 1 bis 8 codiert.

11. Wirtszelle umfassend eine Nucleinsäuresequenz, die ein thermostabiles DNA-Polymeraseenzym nach einem der Ansprüche 1 bis 8 codiert.

12. Verfahren zur Herstellung eines thermostabilen DNA-Polymeraseenzym, umfassend:  
 (a) Züchten einer Wirtszelle nach Anspruch 11 unter Bedingungen welche die Expression des thermostabilen DNA-Polymeraseenzym begünstigen; und  
 (b) Isolieren des thermostabilen DNA-Polymeraseenzym aus der Wirtszelle oder dem Kulturmedium.

13. Thermostabiles DNA-Polymeraseenzym, hergestellt gemäß einem Verfahren nach Anspruch 12.

14. Verwendung eines thermostabilen DNA-Polymeraseenzym nach einem der Ansprüche 1 bis 8 in einer Nucleinsäureamplifikation oder Sequenzierungsreaktion.

15. Zusammensetzung zur Verwendung in einer DNA-Sequenzierungsreaktion umfassend: eine Nucleinsäurematrize, einen Oligonucleotidprimer, der komplementär zu der Matrize ist, eine thermostabile DNA-Polymerase nach einem der Ansprüche 1 bis 8, ein Gemisch von konventionellen dNTPs und mindestens ein unkonventionelles Nucleotid, wobei das Verhältnis von diesem unkonventionellen Nucleotid zu dem entsprechenden konventionellen Nucleotid 1 : 1 oder niedriger ist.

16. Zusammensetzung nach Anspruch 15, wobei das unkonventionelle Nucleotid ein Ribonucleotid ist, wobei dieses Ribonucleotid vorzugsweise enthalten ist in einer Konzentration von weniger als etwa 100  $\mu\text{M}$  und das entsprechende konventionelle Nucleotid enthalten ist in einer Konzentration von mehr als etwa 100  $\mu\text{M}$ .

17. Zusammensetzung nach Anspruch 16, darüber hinaus gekennzeichnet dadurch, dass das unkonventionelle Nucleotid unmarkiert ist.

18. Verfahren zur Sequenzierung eines Nucleinsäureziels, wobei das Verfahren die folgenden Schritte umfasst:

- (a) Bereitstellen eines unkonventionellen Nucleotids und eines entsprechenden konventionellen Nucleotids in einer DNA-Sequenzierungsreaktion, wobei dieses unkonventionelle und die entsprechenden konventionellen Nucleotide in einem Verhältnis von weniger als etwa 1 : 1 anwesend sind;
- (b) Behandeln der Reaktion aus Schritt (a) in Anwesenheit einer thermostabilen DNA-Polymerase nach einem der Ansprüche 1 bis 8 unter Bedingungen zur Primerverlängerung zur Bereitstellung von Primerverlängerungsprodukten, die das unkonventionelle Nucleotid umfassen;
- (c) Behandeln der Primerverlängerungsprodukte aus Schritt (b) unter Bedingungen zur Hydrolyse der Primer-

verlängerungsprodukte;

(d) Auftrennen der Reaktionsprodukte aus Schritt (c); und

(e) Bestimmen der Sequenz des Nucleinsäureziels.

19. Verfahren zur Sequenzierung nach Anspruch 18, wobei das unkonventionelle Nucleotid ein Ribonucleotid ist, welches vorzugsweise in einer Konzentration von etwa 0,1  $\mu\text{M}$ –100  $\mu\text{M}$  anwesend ist.

20. Verfahren zur Sequenzierung nach Anspruch 18, wobei das entsprechende konventionelle Nucleotid in einer Konzentration von etwa 50  $\mu\text{M}$ –500  $\mu\text{M}$  anwesend ist.

21. Kit zur Sequenzierung einer Nucleinsäure umfassend eine thermostabile DNA-Polymerase nach einem der Ansprüche 1 bis 8 und gegebenenfalls weitere Reagenzien zur Verwendung in einem solchen Sequenzierungsverfahren, wie zum Beispiel einen oder mehrere Oligonucleotidprimer, ein Gemisch aus konventionellen dNTPs, und mindestens ein unkonventionelles Nucleotid, wobei das Verhältnis des unkonventionellen Nucleotids zu dem entsprechenden konventionellen Nucleotid vorzugsweise geringer ist als 1.

Es folgt kein Blatt Zeichnungen