US 20140295500A1

(19) **United States**
(12) **Patent Application Publication**
Bauer et al.

(10) **Pub. No.: US 2014/0295500 A1**
(43) **Pub. Date: Oct. 2, 2014**

(54) **DNA POLYMERASES WITH IMPROVED ACTIVITY**

(71) Applicant: **Roche Molecular Systems, Inc.**, Pleasanton, CA (US)

(72) Inventors: **Keith Bauer**, San Rafael, CA (US); **Thomas W. Myers**, Dublin, CA (US); **Shawn Suko**, El Sobrante, CA (US)

(21) Appl. No.: **14/282,966**

(22) Filed: **May 20, 2014**

Related U.S. Application Data

(63) Continuation of application No. 13/706,107, filed on Dec. 5, 2012, now Pat. No. 8,759,063.

(60) Provisional application No. 61/568,375, filed on Dec. 8, 2011.

Publication Classification

(51) **Int. Cl.**
C12P 19/34 (2006.01)
C12N 9/12 (2006.01)

(52) **U.S. Cl.**
CPC ***C12P 19/34*** (2013.01); ***C12N 9/1252*** (2013.01)
USPC **435/91.2**; 435/194; 536/23.2

(57) **ABSTRACT**

Disclosed are DNA polymerases having increased reverse transcriptase efficiency relative to a corresponding, unmodified polymerase. The polymerases are useful in a variety of disclosed primer extension methods. Also disclosed are related compositions, including recombinant nucleic acids, vectors, and host cells, which are useful, e.g., for production of the DNA polymerases.

Figure 1

						*					
Z05	AHLSGDENL**I**	**RV**	**FQ**	**EG**	**K**	**DI**	**HTQ**	**TAS**	**W**	MFGVS	(SEQ ID NO:12)
Taq	AHLSGDENL**I**	**RV**	**FQ**	**EG**	**R**	**DI**	**HTE**	**TAS**	**W**	MFGVP	(SEQ ID NO:13)
Tfi	AHLSGDENL**I**	**RV**	**FR**	**EG**	**K**	**DI**	**HTE**	**TAA**	**W**	MFGVP	(SEQ ID NO:14)
Tfl	AHLSGDENL**I**	**RV**	**FQ**	**EG**	**R**	**DI**	**HTQ**	**TAS**	**W**	MFGVS	(SEQ ID NO:15)
Sps17	AHLSGDENL**I**	**RV**	**FR**	**EG**	**K**	**DI**	**HTE**	**TAA**	**W**	MFGVP	(SEQ ID NO:16)
Tth	AHLSGDENL**I**	**RV**	**FQ**	**EG**	**K**	**DI**	**HTQ**	**TAS**	**W**	MFGVP	(SEQ ID NO:17)
Tca	AHLSGDENL**I**	**RV**	**FQ**	**EG**	**K**	**DI**	**HTQ**	**TAS**	**W**	MFGVP	(SEQ ID NO:18)
Tma	AHLSGDENL**L**	**RA**	**FE**	**EG**	**I**	**DV**	**HTL**	**TAS**	**R**	IFNVK	(SEQ ID NO:19)
Tne	AHLSGDENL**V**	**KA**	**FE**	**EG**	**I**	**DV**	**HTL**	**TAS**	**R**	IYNVK	(SEQ ID NO:20)
Taf	AHVSKDENL**L**	**KA**	**FK**	**ED**	**L**	**DI**	**HTI**	**TAA**	**K**	IFGVS	(SEQ ID NO:21)
Dra	AHIADDPLM**Q**	**QA**	**FV**	**EG**	**A**	**DI**	**HRR**	**TAA**	**Q**	VLGLD	(SEQ ID NO:23)
Bst	AHIAEDDNL**I**	**EA**	**FR**	**RG**	**L**	**DI**	**HTK**	**TAM**	**D**	IFHVS	(SEQ ID NO:24)
Bca	AHIAEDDNL**M**	**EA**	**FR**	**RD**	**L**	**DI**	**HTK**	**TAM**	**D**	IFQVS	(SEQ ID NO:25)
	---------X_1	X_2X_3	FX_4	X_5X_6	X_7	DX_8	HTX_9	TAX_{10}	X_{11}	-----	(SEQ ID NO:26)

Figure 2

A. **Sequence identities over the entire polymerase I enzyme (corresponding to amino acids 1-834 of Z05)**

Name	Z05	Taq	Tfi	Tfl	Sps17	Tth	Tca	Dra	Tma	Tne	Taf	Bst	Bca	
Z05		0.864	0.833	0.859	0.839	0.962	0.958	0.459	0.374	0.368	0.359	0.407	0.408	
Taq	0.864		0.831	0.854	0.836	0.872	0.864	0.468	0.382	0.368	0.351	0.397	0.397	
Tfi	0.833	0.831		0.82	0.991	0.829	0.824	0.45	0.371	0.375	0.353	0.405	0.397	
Tfl	0.859	0.854	0.82		0.824	0.853	0.848	0.462	0.381	0.374	0.356	0.397	0.398	
Sps17	0.839	0.836	0.991	0.824		0.835	0.83	0.452	0.375	0.377	0.355	0.407	0.399	
Tth	0.962	0.872	0.829	0.853	0.835		0.989	0.463	0.373	0.367	0.358	0.406	0.406	
Tca	0.958	0.864	0.824	0.848	0.83	0.989		0.46	0.371	0.365	0.356	0.404	0.404	
Dra	0.459	0.468	0.45	0.462	0.452	0.463	0.46		0.334	0.325	0.314	0.338	0.339	
Tma	0.374	0.382	0.371	0.381	0.375	0.373	0.371	0.334		0.854	0.567	0.37	0.377	
Tne	0.368	0.368	0.375	0.374	0.377	0.367	0.365	0.325	0.854		0.558	0.377	0.376	
Taf	0.359	0.351	0.353	0.356	0.355	0.358	0.356	0.314	0.567	0.558		0.356	0.364	
Bst	0.407	0.397	0.405	0.397	0.407	0.406	0.404	0.338	0.37	0.377	0.356		0.881	
Bca	0.408	0.397	0.397	0.398	0.399	0.406	0.404	0.339	0.377	0.376	0.364	0.881		

B. **Sequence identities over polymerase sub domain only (corresponding to amino acids 420-834 of Z05)**

Name	Z05	Taq	Tfi	Tfl	Sps17	Tth	Tca	Dra	Tma	Tne	Taf	Bst	Bca	
Z05		0.901	0.845	0.891	0.845	0.975	0.973	0.563	0.483	0.478	0.44	0.498	0.49	
Taq	0.901		0.879	0.901	0.877	0.906	0.901	0.561	0.488	0.473	0.44	0.503	0.495	
Tfi	0.845	0.879		0.857	0.997	0.853	0.853	0.566	0.495	0.49	0.449	0.512	0.49	
Tfl	0.891	0.901	0.857		0.855	0.889	0.889	0.571	0.492	0.48	0.444	0.494	0.485	
Sps17	0.845	0.877	0.997	0.855		0.853	0.853	0.566	0.495	0.49	0.449	0.512	0.49	
Tth	0.975	0.906	0.853	0.889	0.853		0.99	0.563	0.478	0.473	0.437	0.496	0.488	
Tca	0.973	0.901	0.853	0.889	0.853	0.99		0.563	0.478	0.473	0.437	0.496	0.488	
Dra	0.563	0.561	0.566	0.571	0.566	0.563	0.563		0.45	0.448	0.426	0.474	0.454	
Tma	0.483	0.488	0.495	0.492	0.495	0.478	0.478	0.45		0.883	0.622	0.474	0.475	
Tne	0.478	0.473	0.49	0.48	0.49	0.473	0.473	0.448	0.883		0.615	0.476	0.473	
Taf	0.44	0.44	0.449	0.444	0.449	0.437	0.437	0.426	0.622	0.615		0.46	0.473	
Bst	0.498	0.503	0.512	0.494	0.512	0.496	0.496	0.474	0.474	0.476	0.46		0.898	
Bca	0.49	0.495	0.49	0.485	0.49	0.488	0.488	0.454	0.475	0.473	0.473	0.898		

Figure 3

A. Sequence identities over the entire polymerase I enzyme (corresponding to amino acids 1-834 of Z05)

Name	**Z05**	**Tth**	**Tfi**	**Tfl**	**Tca**	**Taq**	**Sps17**
Z05		0.962	0.833	0.859	0.958	0.864	0.839
Tth	0.962		0.829	0.853	0.989	0.872	0.835
Tfi	0.833	0.829		0.82	0.824	0.831	0.991
Tfl	0.859	0.853	0.82		0.848	0.854	0.824
Tca	0.958	0.989	0.824	0.848		0.864	0.83
Taq	0.864	0.872	0.831	0.854	0.864		0.836
Sps17	0.839	0.835	0.991	0.824	0.83	0.836	

B. Sequence identities over polymerase sub domain only (corresponding to amino acids 420-834 of Z05)

Name	**Z05**	**Tth**	**Tfi**	**Tfl**	**Tca**	**Taq**	**Sps17**
Z05		0.975	0.845	0.891	0.973	0.901	0.845
Tth	0.975		0.853	0.889	0.99	0.906	0.853
Tfi	0.845	0.853		0.857	0.853	0.879	0.997
Tfl	0.891	0.889	0.857		0.889	0.901	0.855
Tca	0.973	0.99	0.853	0.889		0.901	0.853
Taq	0.901	0.906	0.879	0.901	0.901		0.877
Sps17	0.845	0.853	0.997	0.855	0.853	0.877	

DNA POLYMERASES WITH IMPROVED ACTIVITY

CROSS-REFERENCES TO RELATED APPLICATIONS

[0001] The present application is a continuation of U.S. patent application Ser. No. 13/706,107, filed Dec. 5, 2012, which claims benefit of priority to U.S. Provisional Patent Application No. 61/568,375, filed Dec. 8, 2011, each of which is incorporated herein by reference in its entirety.

REFERENCE TO A "SEQUENCE LISTING," A TABLE, OR A COMPUTER PROGRAM LISTING APPENDIX SUBMITTED AS AN ASCII TEXT FILE

[0002] The Sequence Listing written in file-140-1.TXT, created on Dec. 5, 2012, 131,072 bytes, machine format IBM-PC, MS-Windows operating system, is hereby incorporated by reference in its entirety for all purposes.

FIELD OF THE INVENTION

[0003] The present invention provides DNA polymerases with improved activities, including increased reverse transcriptase efficiency, mismatch tolerance, extension rate and/or tolerance of reverse transcriptase (RT) and polymerase inhibitors, as well as use of such polymerases in various applications, including nucleic acid polynucleotide extension and amplification.

BACKGROUND OF THE INVENTION

[0004] DNA polymerases are responsible for the replication and maintenance of the genome, a role that is central to accurately transmitting genetic information from generation to generation. DNA polymerases function in cells as the enzymes responsible for the synthesis of DNA. They polymerize deoxyribonucleoside triphosphates in the presence of a metal activator, such as Mg^{2+}, in an order dictated by the DNA template or polynucleotide template that is copied. In vivo, DNA polymerases participate in a spectrum of DNA synthetic processes including DNA replication, DNA repair, recombination, and gene amplification. During each DNA synthetic process, the DNA template is copied once or at most a few times to produce identical replicas. In contrast, in vitro, DNA replication can be repeated many times such as, for example, during polymerase chain reaction (see, e.g., U.S. Pat. No. 4,683,202 to Mullis).

[0005] In the initial studies with polymerase chain reaction (PCR), the DNA polymerase was added at the start of each round of DNA replication (see U.S. Pat. No. 4,683,202, supra). Subsequently, it was determined that thermostable DNA polymerases could be obtained from bacteria that grow at elevated temperatures, and that these enzymes need to be added only once (see U.S. Pat. No. 4,889,818 to Gelfand and U.S. Pat. No. 4,965,188 to Mullis). At the elevated temperatures used during PCR, these enzymes are not irreversibly inactivated. As a result, one can carry out repetitive cycles of polymerase chain reactions without adding fresh enzymes at the start of each synthetic addition process. DNA polymerases, particularly thermostable polymerases, are the key to a large number of techniques in recombinant DNA studies and in medical diagnosis of disease. For diagnostic applications in particular, a target nucleic acid sequence may be only a small portion of the DNA or RNA in question, so it may be difficult to detect the presence of a target nucleic acid sequence without amplification.

[0006] The overall folding pattern of DNA polymerases resembles the human right hand and contains three distinct subdomains of palm, fingers, and thumb. (See Beese et al., *Science* 260:352-355, 1993); Patel et al., *Biochemistry* 34:5351-5363, 1995). While the structure of the fingers and thumb subdomains vary greatly between polymerases that differ in size and in cellular functions, the catalytic palm subdomains are all superimposable. For example, motif A, which interacts with the incoming dNTP and stabilizes the transition state during chemical catalysis, is superimposable with a mean deviation of about one Å amongst mammalian pol α and prokaryotic pol I family DNA polymerases (Wang et al., *Cell* 89:1087-1099, 1997). Motif A begins structurally at an antiparallel β-strand containing predominantly hydrophobic residues and continues to an α-helix. The primary amino acid sequence of DNA polymerase active sites is exceptionally conserved. In the case of motif A, for example, the sequence DYSQIELR (SEQ ID NO:22) is retained in polymerases from organisms separated by many millions years of evolution, including, e.g., *Thermus aquaticus, Chlamydia trachomatis*, and *Escherichia coli.*

[0007] In addition to being well-conserved, the active site of DNA polymerases has also been shown to be relatively mutable, capable of accommodating certain amino acid substitutions without reducing DNA polymerase activity significantly. (See, e.g., U.S. Pat. No. 6,602,695 to Patel et al.). Such mutant DNA polymerases can offer various selective advantages in, e.g., diagnostic and research applications comprising nucleic acid synthesis reactions.

[0008] There are at least two steps in the enzymatic process of DNA polymerization; 1) the incorporation of the incoming nucleotide and 2) the extension of the newly incorporated nucleotide. The overall faithfulness or "fidelity" of the DNA polymerase is generally thought of as a conglomerate of these two enzymatic activities, but the steps are distinct. A DNA polymerase may misincorporate the incoming nucleotide, but if it is not efficiently extended the extension rate will be severely decreased and overall product formation would be minimal. Alternatively, it is possible to have a DNA polymerase misincorporate the incoming nucleotide and readily misextend the newly formed mismatch. In this case, the overall extension rate would be high, but the overall fidelity would be low. An example of this type of enzyme would be ES112 DNA polymerase (E683R Z05 DNA polymerase; see U.S. Pat. No. 7,179,590, entitled "High temperature reverse transcription using mutant DNA polymerases" filed Mar. 30, 2001 by Smith et al., which is incorporated by reference) when using Mn^{2+} as the divalent metal ion activator. The enzyme has a very high efficiency because unlike typical DNA polymerases that tend to hesitate/stall when a mismatch is encountered, the ES112 DNA polymerase readily extends the mismatch. The phenotype displayed in ES112 is more pronounced during the RT step, presumably because of structural effects of the RNA/DNA heteroduplex vs. the DNA/DNA homoduplex. A second example would be if the DNA polymerase does not readily misincorporate (may be even less likely to misincorporate), but does have increased capacity to misextend a mismatch. In this case, the fidelity is not significantly altered for the overall product. In general, this type of enzyme is more favorable for extension reactions than the characteristics of ES 112 in Mn^{2+} because the fidelity of the product is improved. However this attribute can be utilized to allow the misextension of a mismatched oligonucleotide primer such as when an oligonucleotide primer of a single sequence is hybridized to a target that has sequence heterogeneity (e.g., viral targets), but the normal or lower misincorporation rate allows for completion of DNA synthesis beyond the original oligonucleotide primer. An example of this type of DNA polymerase is Z05 D580G DNA polymerase. (see U.S. Patent Publication No. 2009/0148891 entitled "DNA Polymerases and Related Methods" filed Oct. 17, 2007 by Bauer et. al., which is incorporated by reference). This type of activity is referred to as "mismatch tolerant" because it is more tolerant to mismatches in the oligonucleotide primer. While the examples above have discussed primer extension type reactions, the activity can be more significant in reactions such as RT-PCR and PCR where primer extension is reoccurring frequently. Data suggests that while enzymes such as Z05 D580G are more "tolerant" to mismatches, they also have enhanced ability to extend oligonucleotide primers containing modified bases (eg., t-butyl benzyl modified bases) or in the presence of DNA binding dyes such as SYBR Green I (see U.S. Patent Publication No. 2009/028053 entitled "Improved DNA Polymerases and Related Methods" filed Apr. 16, 2009 by Bauer et al., which is incorporated by reference).

[0009] Reverse transcription polymerase chain reaction (RT-PCR) is a technique used in many applications to detect/ and or quantify RNA targets by amplification. In order to amplify RNA targets by PCR, it is necessary to first reverse transcribe the RNA template into cDNA. Typically, RT-PCR assays rely on a non-thermostable reverse transcriptase (RNA dependent DNA polymerase), derived from a mesophilic organism, for the initial cDNA synthesis step (RT). An additional thermostable DNA polymerase is required for amplification of cDNA to tolerate elevated temperatures required for nucleic acid denaturation in PCR. There are several potential benefits of using thermoactive or thermostable DNA polymerases engineered to perform more efficient reverse transcription for RT-PCR assays. Increased reverse transcriptase activity coupled with the ability to use higher reverse transcription incubation temperatures, that allow for relaxing of RNA template secondary structure, can result in overall higher cDNA synthesis efficiency and assay sensitivity. Higher temperature incubation could also increase specificity by reducing false priming in the reverse transcription step. Enzymes with improved reverse transcription efficiency can simplify assay design by allowing for reduced RT incubation times and/or enzyme concentration. When using dUTP and UNG, nonspecific extension products containing dUMP that are formed during nonstringent set-up conditions are degraded by UNG and cannot be utilized either as primers or as templates. When using a non-thermostable reverse transcriptase (RNA dependent DNA polymerase) derived from a mesophilic organism, it is not possible to utilize the dUTP and UNG methodologies. (Myers, T. W. et al., Amplification of RNA: High Temperature Reverse Transcription and DNA Amplification with *Thermus thermophilus* DNA Polymerase, in *PCR Strategies*, Innis, M. A., Gelfand, D. H., and Sninsky, J. J., Eds., Academic Press, San Diego, Calif., 58-68, (1995)). However, the use of a thermoactive or thermostable DNA polymerase of the invention for the reverse transcription step enables the reaction to be completely compatible with the utilization of the dUTP/uracil N-glycosylase (UNG) carry-over prevention system (Longo et al., Use of Uracil DNA Glycosylase to Control Carry-over Contamination in Polymerase Chain Reactions. *Gene* 93:125-128, (1990). In addition to providing carry-over contamination control, the use of dUTP and UNG provides a "hot-start" to reduce nonspecific amplification (Innis and Gelfand 1999).

BRIEF SUMMARY OF THE INVENTION

[0010] Provided herein are DNA polymerases having improved activities, including increased reverse transcriptase efficiency, mismatch tolerance, extension rate and/or tolerance of RT and polymerase inhibitors, relative to a corresponding, unmodified control polymerase, and methods of making and using such DNA polymerases. In some embodiments, the improved DNA polymerase has increased reverse transcriptase efficiency as compared with a control DNA polymerase. In some embodiments, the improved DNA polymerase has the same or substantially similar DNA-dependent polymerase activity as compared with a control DNA polymerase. Thus, in some embodiments, the improved DNA polymerase comprises an amino acid sequence that is substantially identical (e.g., at least about 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% identical) to SEQ ID NO:1, wherein the amino acid of the DNA polymerase corresponding to position 640 of SEQ ID NO:1 is any amino acid other than I or V. In some embodiments, the control DNA polymerase has the same amino acid sequence as the DNA polymerase except that the amino acid of the control DNA polymerase corresponding to position 640 of SEQ ID NO:1 is I or V. For example, in some embodiments, the amino acid at the position corresponding to position 640 of SEQ ID NO:1 of the improved polymerase is selected from G, A, R, F, W, P, S, T, C, Y, N, Q, D, E, K, L, M, or H. In some embodiments, the amino acid at the position corresponding to position 640 of SEQ ID NO:1 of the improved polymerase is F.

[0011] In some embodiments, the improved DNA polymerase further comprises an amino acid sequence that is substantially identical (e.g., at least about 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% identical) to SEQ ID NO:1, wherein the amino acid of the DNA polymerase corresponding to position 580 of SEQ ID NO:1 is any amino acid other than D or E. In some embodiments, the amino acid of the DNA polymerase corresponding to position 580 of SEQ ID NO:1 is any amino acid other than D. In some embodiments, the amino acid of the DNA polymerase corresponding to position 580 of SEQ ID NO:1 is selected from the group consisting of L, G, T, Q, A, S, N, R, and K. In some embodiments, the amino acid of the DNA polymerase corresponding to position 580 of SEQ ID NO:1 is G.

[0012] In some embodiments, the improved DNA polymerase further comprises an amino acid sequence that is substantially identical (e.g., at least about 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% identical) to SEQ ID NO:1, wherein the amino acid of the DNA polymerase corresponding to position 709 of SEQ ID NO:1 is any amino acid other than I. In some embodiments, the amino acid of the DNA polymerase corresponding to position 709 of SEQ ID NO:1 is selected from the group consisting of K, R, S, G, and A. In some embodiments, the amino acid of the DNA polymerase corresponding to position 709 of SEQ ID NO:1 is K.

[0013] In some embodiments, the improved DNA polymerase comprises an amino acid sequence that is substantially identical (e.g., at least about 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% identical) to SEQ ID NO:1, wherein the amino acid of the DNA polymerase corresponding to position 640 of SEQ ID NO:1 is any amino acid other than I, the amino acid corresponding to position 580 of SEQ ID NO:1 is any amino acid other than D or E, and the amino acid corresponding to position 709 of SEQ ID NO:1 is any amino acid other than I. Thus, in some embodiments, the improved DNA polymerase comprises an amino acid sequence that is substantially identical (e.g., at least about 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% identical) to SEQ ID NO:1, wherein the amino acid of the DNA polymerase corresponding to position 640 of SEQ ID NO:1 is F, the amino acid corresponding to position 580 of SEQ ID NO:1 is G, and the amino acid corresponding to position 709 of SEQ ID NO:1 is K.

[0014] In some embodiments, the improved DNA polymerase has increased reverse transcriptase efficiency without a substantial decrease in DNA-dependent polymerase activity compared with a control DNA polymerase, wherein the amino acid of the DNA polymerase corresponding to position 640 of SEQ ID NO:1 is any amino acid other than I, and the amino acid corresponding to position 709 of SEQ ID NO:1 is any amino acid other than I, and wherein the control DNA polymerase has the same amino acid sequence as the DNA polymerase except that the amino acid of the control DNA polymerase corresponding to position 640 of SEQ ID NO:1 is I and the amino acid corresponding to position 709 of SEQ ID NO:1 is I. Thus, in some embodiments, the amino acid of the DNA polymerase corresponding to position 640 of SEQ ID NO:1 is F, and the amino acid corresponding to position 709 of SEQ ID NO:1 is K. In some embodiments, the improved DNA polymerase further comprises an amino acid substitution at the amino acid corresponding to position 580 of SEQ ID NO:1. Thus, in some embodiments, the amino acid of the DNA polymerase corresponding to position 640 of SEQ ID NO:1 is any amino acid other than I, the amino acid corresponding to position 709 of SEQ ID NO:1 is any amino acid other than I, and the amino acid corresponding to position 580 of SEQ ID NO:1 is any amino acid other than D or E. In some embodiments, the amino acid of the DNA polymerase corresponding to position 640 of SEQ ID NO:1 is F, the amino acid corresponding to position 709 of SEQ ID NO:1 is K, and the amino acid corresponding to position 580 of SEQ ID NO:1 is G.

[0015] Various DNA polymerases are amenable to mutation according to the present invention. Particularly suitable are thermostable polymerases, including wild-type or naturally occurring thermostable polymerases from various species of thermophilic bacteria, as well as synthetic thermostable polymerases derived from such wild-type or naturally occurring enzymes by amino acid substitution, insertion, or deletion, or other modification. Exemplary unmodified forms of polymerase include, e.g., CS5, CS6 or Z05 DNA polymerase, or a functional DNA polymerase having at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% amino acid sequence identity thereto. Other unmodified polymerases include, e.g., DNA polymerases from any of the following species of thermophilic bacteria (or a functional DNA polymerase having at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% amino acid sequence identity to such a polymerase): *Thermotoga maritima; Thermus aquaticus; Thermus thermophilus; Thermus flavus; Thermus filiformis; Thermus* sp. sps17; *Thermus* sp. Z05; *Thermotoga neopolitana; Thermosipho africanus; Thermus caldophilus, Deinococcus radiodurans, Bacillus stearothermophilus* or *Bacillus caldotenax*. Suitable polymerases also include those having reverse transcriptase (RT) activity and/or the ability to incorporate unconventional nucleotides, such as ribonucleotides or other 2'-modified nucleotides.

[0016] While thermostable DNA polymerases possessing efficient reverse transcription activity are particularly suited for performing RT-PCR, especially single enzyme RT-PCR, thermoactive, but not thermostable DNA polymerases possessing efficient reverse transcription activity also are amenable to mutation according to the present invention. For example, the attributes of increased reverse transcriptase efficiency, mismatch tolerance, extension rate, and/or tolerance of RT inhibitors are useful for the RT step in an RT-PCR and this step does not need to be performed at temperatures that would inactivate a thermoactive but not thermostable DNA polymerase. Following the RT step, a thermostable DNA polymerase could either be added or it could already be included in the reaction mixture to perform the PCR amplification step. For example, the improved DNA polymerase described herein can be combined with a second thermostable DNA polymerase prior to the RT step in a buffer suitable for extension and amplification of RNA and DNA templates, as described in the Examples. Examples of suitable thermostable DNA polymerases are described in U.S. Pat. No. 4,889,818 to Gelfand et al., and U.S. Pat. Nos. 5,773,258 and 5,677,152 to Birch et al., which are expressly incorporated by reference herein in their entirety. In some embodiments, the second thermostable DNA polymerase is AmpliTaq® DNA polymerase (Deoxy-nucleoside triphosphate: DNA Deoxynucleotidyltransferase, E.C.2.7.7.7). In some embodiments, the second thermostable DNA polymerase is a reversibly inactivated thermostable polymerase, as described below. In one embodiment, the reversibly inactivated thermostable polymerase is AmpliTaq Gold® DNA polymerase (Roche Applied Science, Indianapolis, Ind., USA). This second methodology would especially benefit by using a chemically modified thermostable DNA polymerase (or other HotStart technology to inactivate the thermostable DNA polymerase) so that it would not be fully active during the RT step. An example of a thermoactive but not thermostable DNA polymerase possessing efficient reverse transcription activity is the DNA polymerase from *Carboxydothermus hydrogenoformans* (Chy; SEQ ID NO:39). See, e.g., U.S. Pat. Nos. 6,468,775 and 6,399,320.

[0017] In some embodiments, the DNA polymerase has at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% amino acid sequence identity to a polymerase selected from the group consisting of:

[0018] (a) a *Thermus* sp. Z05 DNA polymerase (Z05) (SEQ ID NO:1);

[0019] (b) a *Thermus aquaticus* DNA polymerase (Taq) (SEQ ID NO:2);

[0020] (c) a *Thermus filiformis* DNA polymerase (Tfi) (SEQ ID NO:3);

[0021] (d) a *Thermus flavus* DNA polymerase (Tfl) (SEQ ID NO:4);

[0022] (e) a *Thermus* sp. sps17 DNA polymerase (Sps17) (SEQ ID NO:5);

[0023] (f) a *Thermus thermophilus* DNA polymerase (Tth) (SEQ ID NO:6); and

[0024] (g) a *Thermus caldophilus* DNA polymerase (Tca) (SEQ ID NO:7)

[0025] (h) *Carboxydothermus hydrogenoformans* DNA polymerase (Chy) (SEQ ID NO:39)

[0026] In some embodiments, the DNA polymerase is a *Thermotoga* DNA polymerase. For example, in some embodiments, the DNA polymerase has at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% amino acid sequence identity to a polymerase selected from the group consisting of:

[0027] (a) a *Thermotoga maritima* DNA polymerase (Tma) (SEQ ID NO:34);

[0028] (b) a *Thermotoga neopolitana* DNA polymerase (Tne) (SEQ ID NO:35);

[0029] In some embodiments, the DNA polymerase has at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% amino acid sequence identity to SEQ ID NO:1. In some embodiments, the DNA polymerase is a *Thermus* sp. Z05 DNA polymerase (Z05) DNA polymerase (i.e., SEQ ID NO:1), and the amino acid at position 640 is any amino acid other than I. For example, in some embodiments, the amino acid at position 640 is selected from G, A, V, R, F, W, P, S, T, C, Y, N, Q, D, E, K, L, M or H. In some embodiments, the DNA polymerase is a Z05 DNA polymerase, and the amino acid at position 640 is F. In some embodiments, the DNA polymerase is a Z05 DNA polymerase further comprising a substitution at position 580, and the amino acid at position 580 is any amino acid other than D or E. In some embodiments, the DNA polymerase is a Z05 DNA polymerase, and the amino acid at position 580 is any amino acid other than D. In some embodiments, the DNA polymerase is a Z05 DNA polymerase, and the amino acid at position 580 is selected from the group consisting of L, G, T, Q, A, S, N, R, and K. In some embodiments, the DNA polymerase is a Z05 DNA polymerase, and the amino acid at position 580 is G. In some embodiments, the DNA polymerase is a Z05 DNA polymerase further comprising a substitution at position 709, and the amino acid at position 709 is any amino acid other than I. In some embodiments, the DNA polymerase is a Z05 DNA polymerase, and the amino acid at position 709 is selected from the group consisting of K, R, S, G, and A. In some embodiments, the DNA polymerase is a Z05 DNA polymerase, and the amino acid at position 709 is K.

[0030] In some embodiments, the control DNA polymerase is a Z05, Z05 D580G, or Z05 D580G I709K polymerase.

[0031] The mutant or improved polymerases can include other, non-substitutional modifications. One such modification is a thermally reversible covalent modification that inactivates the enzyme, but which is reversed to activate the enzyme upon incubation at an elevated temperature, such as a temperature typically used for polynucleotide extension. Exemplary reagents for such thermally reversible modifications are described in U.S. Pat. Nos. 5,773,258 and 5,677,152 to Birch et al., which are expressly incorporated by reference herein in their entirety.

[0032] In some embodiments, the reverse transcriptase activity is determined by performing real-time RT-PCR amplification and detection of a Hepatitis C Virus (HCV) transcript generated from the first 800 bases of HCV genotype Ib 5'NTR in pSP64 poly(A) (Promega). Two or more reaction mixtures can have titrated numbers of copies of the Hepatitis C Virus (HCV) transcript (e.g., 1:5 titrations, 1:10 titrations, e.g., 10,000 copies, 1000 copies, 100 copies, 10 copies, 1 copy, 0 copies in several reaction mixtures). The reverse transcriptase ability of a polymerase of the invention can be compared to the reverse transcriptase ability of a reference polymerase (e.g., a naturally occurring, unmodified, or control polymerase), over a preselected unit of time, as described herein. Polymerases with improved reverse transcriptase ability will amplify the transcript with greater efficiency, or will require a lower number of PCR cycles to amplify the transcript (i.e., exhibit a lower Cp value, as calculated herein), in comparison to a naturally occurring or unmodified polymerase. Moreover, in some embodiments, polymerases with improved RT function also have improved replication of long RNA (e.g., at least 500 or 1000 or 2000 or 5000 or more nucleotides long) templates. In some embodiments, the improved reverse transcriptase efficiency includes a shorter reverse transcription time in comparison to a control polymerase. Thus, in some embodiments, polymerases with increased reverse transcriptase efficiency will reverse transcribe an RNA template faster than a control or reference polymerase.

[0033] In various other aspects, the present invention provides a recombinant nucleic acid encoding a mutant or improved DNA polymerase as described herein, a vector comprising the recombinant nucleic acid, and a host cell transformed with the vector. In certain embodiments, the vector is an expression vector. Host cells comprising such expression vectors are useful in methods of the invention for producing the mutant or improved polymerase by culturing the host cells under conditions suitable for expression of the recombinant nucleic acid. The polymerases of the invention may be contained in reaction mixtures and/or kits. The embodiments of the recombinant nucleic acids, host cells, vectors, expression vectors, reaction mixtures and kits are as described above and herein.

[0034] In yet another aspect, a method for conducting polynucleotide extension is provided. The method generally includes contacting a DNA polymerase having increased reverse transcriptase efficiency, mismatch tolerance, extension rate and/or tolerance of RT and polymerase inhibitors as described herein with a primer, a polynucleotide template, and nucleoside triphosphates under conditions suitable for extension of the primer, thereby producing an extended primer. The polynucleotide template can be, for example, an RNA or DNA template. The nucleotide triphosphates can include unconventional nucleotides such as, e.g., ribonucleotides and/or labeled nucleotides. Further, the primer and/or template can include one or more nucleotide analogs. In some variations, the polynucleotide extension method is a method for polynucleotide amplification that includes contacting the mutant or improved DNA polymerase with a primer pair, the polynucleotide template, and the nucleoside triphosphates under conditions suitable for amplification of the polynucleotide. The polynucleotide extension reaction can be, e.g., PCR, isothermal extension, or sequencing (e.g., 454 sequencing reaction). The polynucleotide template can be from any type of biological sample.

[0035] Optionally, the primer extension reaction comprises an actual or potential inhibitor of a reference or unmodified polymerase. The inhibitor can inhibit the nucleic acid extension rate and/or the reverse transcription efficiency of a reference or unmodified (control) polymerase. In some embodiments, the inhibitor is hemoglobin, or a degradation product thereof. For example, in some embodiments, the hemoglobin degradation product is a heme breakdown product, such as hemin, hematoporphyrin, or bilirubin. In some embodiments, the inhibitor is an iron-chelator or a purple pigment. In other embodiments, the inhibitor is heparin or melanin. In certain embodiments, the inhibitor is an intercalating dye. In some embodiments, the intercalating dye is [2-[N-bis-(3-dimethylaminopropyl)-amino]-4-[2,3-dihydro-3-methyl-(benzo-1,3-thiazol-2-yl)-methylidene]-1-phenyl-quinolinium]+. In some embodiments, the intercalating dye is [2-[N-(3-dimethylaminopropyl)-N-propylamino]-4-[2,3-dihydro-3-methyl-(benzo-1,3-thiazol-2-yl)-methylidene]-1-phenyl-quinolinium]+. In some embodiments, the intercalating dye is not [2-[N-(3-dimethylaminopropyl)-N-propylamino]-4-[2,3-dihydro-3-methyl-(benzo-1,3-thiazol-2-yl)-methylidene]-1-phenyl-quinolinium]+. In some embodiments, the conditions suitable for extension comprise Mg^{++}. In some embodiments, the conditions suitable for extension comprise Mn^{++}.

[0036] The present invention also provides a kit useful in such a polynucleotide extension method. Generally, the kit includes at least one container providing a mutant or improved DNA polymerase as described herein. In certain embodiments, the kit further includes one or more additional containers providing one or more additional reagents. For example, in specific variations, the one or more additional containers provide nucleoside triphosphates; a buffer suitable for polynucleotide extension; and/or one or more primer or probe polynucleotides, hybridizable, under polynucleotide extension conditions, to a predetermined polynucleotide template. The polynucleotide template can be from any type of biological sample.

[0037] Further provided are reaction mixtures comprising the polymerases of the invention. The reaction mixtures can also contain a template nucleic acid (DNA and/or RNA), one or more primer or probe polynucleotides, nucleoside triphosphates (including, e.g., deoxyribonucleoside triphosphates, ribonucleoside triphosphates, labeled nucleoside triphosphates, unconventional nucleoside triphosphates), buffers, salts, labels (e.g., fluorophores). In some embodiments, the reaction mixtures comprise an iron chelator or a purple dye. In certain embodiments, the reaction mixtures comprise hemoglobin, or a degradation product of hemoglobin. For example, in certain embodiments, the degradation products of hemoglobin include heme breakdown products such as hemin, hematin, hematophoryn, and bilirubin. In other embodiments, the reaction mixtures comprise heparin or a salt thereof. Optionally, the reaction mixture comprises an intercalating dye (including but not limited to those described above or elsewhere herein). In certain embodiments, the reaction mixture contains a template nucleic acid that is isolated from blood. In other embodiments, the template nucleic acid is RNA and the reaction mixture comprises heparin or a salt thereof.

[0038] In some embodiments, the reaction mixture comprises two or more polymerases. For example, in some embodiments, the reaction mixture comprises an improved DNA polymerase having increased reverse transcription efficiency (e.g., increased activity extending an RNA-template) as described herein, and another polymerase having DNA-dependent polymerase activity. In one embodiment, the reaction mixture comprises a blend of an improved DNA polymerase having increased reverse transcription efficiency as described herein, and a second thermostable DNA-dependent polymerase. The second thermostable DNA-dependent polymerase can be a reversibly modified polymerase as described above such that the enzyme is inactive at temperatures suitable for the reverse transcription step, but is activated under suitable conditions, for example, at elevated temperatures of about 90° C. to 100° C. for a period of time up to about 12 minutes. Suitable conditions for activation of a reversibly inactivated thermostable polymerase are provided, for example, in a Hot Start PCR reaction, as described in the Examples. Examples of suitable second thermostable DNA-dependent polymerases are described in U.S. Pat. Nos. 5,773,258 and 5,677,152 to Birch et al., supra.

[0039] Further embodiments of the invention are described herein.

DEFINITIONS

[0040] Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention pertains. Although essentially any methods and materials similar to those described herein can be used in the practice or testing of the present invention, only exemplary methods and materials are described. For purposes of the present invention, the following terms are defined below.

[0041] The terms "a," "an," and "the" include plural referents, unless the context clearly indicates otherwise.

[0042] An "amino acid" refers to any monomer unit that can be incorporated into a peptide, polypeptide, or protein. As used herein, the term "amino acid" includes the following twenty natural or genetically encoded alpha-amino acids: alanine (Ala or A), arginine (Arg or R), asparagine (Asn or N), aspartic acid (Asp or D), cysteine (Cys or C), glutamine (Gln or Q), glutamic acid (Glu or E), glycine (Gly or G), histidine (His or H), isoleucine (Ile or I), leucine (Leu or L), lysine (Lys or K), methionine (Met or M), phenylalanine (Phe or F), proline (Pro or P), serine (Ser or S), threonine (Thr or T), tryptophan (Trp or W), tyrosine (Tyr or Y), and valine (Val or V). In cases where "X" residues are undefined, these should be defined as "any amino acid." The structures of these twenty natural amino acids are shown in, e.g., Stryer et al., *Biochemistry*, 5th ed., Freeman and Company (2002), which is incorporated by reference. Additional amino acids, such as selenocysteine and pyrrolysine, can also be genetically coded for (Stadtman (1996) "Selenocysteine," *Annu Rev Biochem.* 65:83-100 and Ibba et al. (2002) "Genetic code: introducing pyrrolysine," *Curr Biol.* 12(13):R464-R466, which are both incorporated by reference). The term "amino acid" also includes unnatural amino acids, modified amino acids (e.g., having modified side chains and/or backbones), and amino acid analogs. See, e.g., Zhang et al. (2004) "Selective incorporation of 5-hydroxytryptophan into proteins in mammalian cells," *Proc. Natl. Acad. Sci. U.S.A.* 101(24):8882-8887, Anderson et al. (2004) "An expanded genetic code with a functional quadruplet codon" *Proc. Natl. Acad. Sci. U.S.A.* 101(20):7566-7571, Ikeda et al. (2003) "Synthesis of a novel histidine analogue and its efficient incorporation into a protein in vivo," *Protein Eng. Des. Sel.* 16(9):699-706, Chin et al. (2003) "An Expanded Eukaryotic Genetic Code," *Science* 301(5635):964-967, James et al. (2001) "Kinetic characterization of ribonuclease S mutants containing photoisomerizable phenylazophenylalanine residues," *Protein Eng. Des. Sel.* 14(12):983-991, Kohrer et al. (2001) "Import of amber and ochre suppressor tRNAs into mammalian cells: A general approach to site-specific insertion of amino acid analogues into proteins," *Proc. Natl. Acad. Sci. U.S.A.* 98(25):14310-14315, Bacher et al. (2001) "Selection and Characterization of *Escherichia coli* Variants Capable of Growth on an Otherwise Toxic Tryptophan Analogue," *J. Bacteriol.* 183(18):

5414-5425, Hamano-Takaku et al. (2000) "A Mutant *Escherichia coli* Tyrosyl-tRNA Synthetase Utilizes the Unnatural Amino Acid Azatyrosine More Efficiently than Tyrosine," *J. Biol. Chem.* 275(51):40324-40328, and Budisa et al. (2001) "Proteins with {beta}-(thienopyrrolyl)alanines as alternative chromophores and pharmaceutically active amino acids," *Protein Sci.* 10(7):1281-1292, which are each incorporated by reference.

[0043] To further illustrate, an amino acid is typically an organic acid that includes a substituted or unsubstituted amino group, a substituted or unsubstituted carboxy group, and one or more side chains or groups, or analogs of any of these groups. Exemplary side chains include, e.g., thiol, seleno, sulfonyl, alkyl, aryl, acyl, keto, azido, hydroxyl, hydrazine, cyano, halo, hydrazide, alkenyl, alkynl, ether, borate, boronate, phospho, phosphono, phosphine, heterocyclic, enone, imine, aldehyde, ester, thioacid, hydroxylamine, or any combination of these groups. Other representative amino acids include, but are not limited to, amino acids comprising photoactivatable cross-linkers, metal binding amino acids, spin-labeled amino acids, fluorescent amino acids, metal-containing amino acids, amino acids with novel functional groups, amino acids that covalently or noncovalently interact with other molecules, photocaged and/or photoisomerizable amino acids, radioactive amino acids, amino acids comprising biotin or a biotin analog, glycosylated amino acids, other carbohydrate modified amino acids, amino acids comprising polyethylene glycol or polyether, heavy atom substituted amino acids, chemically cleavable and/or photocleavable amino acids, carbon-linked sugar-containing amino acids, redox-active amino acids, amino thioacid containing amino acids, and amino acids comprising one or more toxic moieties.

[0044] The term "biological sample" encompasses a variety of sample types obtained from an organism and can be used in a diagnostic or monitoring assay. The term encompasses urine, urine sediment, blood, saliva, and other liquid samples of biological origin, solid tissue samples, such as a biopsy specimen or tissue cultures or cells derived therefrom and the progeny thereof. The term encompasses samples that have been manipulated in any way after their procurement, such as by treatment with reagents, solubilization, sedimentation, or enrichment for certain components. The term encompasses a clinical sample, and also includes cells in cell culture, cell supernatants, cell lysates, serum, plasma, biological fluids, and tissue samples.

[0045] The term "mutant," in the context of DNA polymerases of the present invention, means a polypeptide, typically recombinant, that comprises one or more amino acid substitutions relative to a corresponding, functional DNA polymerase.

[0046] The term "unmodified form," in the context of a mutant polymerase, is a term used herein for purposes of defining a mutant DNA polymerase of the present invention: the term "unmodified form" refers to a functional DNA polymerase that has the amino acid sequence of the mutant polymerase except at one or more amino acid position(s) specified as characterizing the mutant polymerase. Thus, reference to a mutant DNA polymerase in terms of (a) its unmodified form and (b) one or more specified amino acid substitutions means that, with the exception of the specified amino acid substitution(s), the mutant polymerase otherwise has an amino acid sequence identical to the unmodified form in the specified motif. The "unmodified polymerase" (and therefore also the modified form having increased reverse transcriptase efficiency, mismatch tolerance, extension rate and/or tolerance of RT and polymerase inhibitors) may contain additional mutations to provide desired functionality, e.g., improved incorporation of dideoxyribonucleotides, ribonucleotides, ribonucleotide analogs, dye-labeled nucleotides, modulating 5'-nuclease activity, modulating 3'-nuclease (or proofreading) activity, or the like. Accordingly, in carrying out the present invention as described herein, the unmodified form of a DNA polymerase is predetermined. The unmodified form of a DNA polymerase can be, for example, a wild-type and/or a naturally occurring DNA polymerase, or a DNA polymerase that has already been intentionally modified. An unmodified form of the polymerase is preferably a thermostable DNA polymerase, such as DNA polymerases from various thermophilic bacteria, as well as functional variants thereof having substantial sequence identity to a wild-type or naturally occurring thermostable polymerase. Such variants can include, for example, chimeric DNA polymerases such as, for example, the chimeric DNA polymerases described in U.S. Pat. Nos. 6,228,628 and 7,148,049, which are incorporated by reference herein in their entirety. In certain embodiments, the unmodified form of a polymerase has reverse transcriptase (RT) activity.

[0047] The term "thermostable polymerase," refers to an enzyme that is stable to heat, is heat resistant, and retains sufficient activity to effect subsequent polynucleotide extension reactions and does not become irreversibly denatured (inactivated) when subjected to the elevated temperatures for the time necessary to effect denaturation of double-stranded nucleic acids. The heating conditions necessary for nucleic acid denaturation are well known in the art and are exemplified in, e.g., U.S. Pat. Nos. 4,683,202, 4,683,195, and 4,965,188, which are incorporated herein by reference. As used herein, a thermostable polymerase is suitable for use in a temperature cycling reaction such as the polymerase chain reaction ("PCR"). Irreversible denaturation for purposes herein refers to permanent and complete loss of enzymatic activity. For a thermostable polymerase, enzymatic activity refers to the catalysis of the combination of the nucleotides in the proper manner to form polynucleotide extension products that are complementary to a template nucleic acid strand. Thermostable DNA polymerases from thermophilic bacteria include, e.g., DNA polymerases from *Thermotoga maritima, Thermus aquaticus, Thermus thermophilus, Thermus flavus, Thermus filiformis, Thermus* species sps17, *Thermus species* Z05, *Thermus caldophilus, Bacillus caldotenax, Thermotoga neopolitana*, and *Thermosipho africanus.*

[0048] The term "thermoactive" refers to an enzyme that maintains catalytic properties at temperatures commonly used for reverse transcription or anneal/extension steps in RT-PCR and/or PCR reactions (i.e., 45-80° C.). Thermostable enzymes are those which are not irreversibly inactivated or denatured when subjected to elevated temperatures necessary for nucleic acid denaturation. Thermoactive enzymes may or may not be thermostable. Thermoactive DNA polymerases can be DNA or RNA dependent from thermophilic species or from mesophilic species including, but not limited to, *Escherichia coli*, Moloney murine leukemia viruses, and Avian myoblastosis virus.

[0049] As used herein, a "chimeric" protein refers to a protein whose amino acid sequence represents a fusion product of subsequences of the amino acid sequences from at least two distinct proteins. A chimeric protein typically is not produced by direct manipulation of amino acid sequences, but, rather, is expressed from a "chimeric" gene that encodes the chimeric amino acid sequence. In certain embodiments, for example, an unmodified form of a mutant DNA polymerase of the present invention is a chimeric protein that consists of an amino-terminal (N-terminal) region derived from a *Thermus* species DNA polymerase and a carboxy-terminal (C-terminal) region derived from Tma DNA polymerase. The N-terminal region refers to a region extending from the N-terminus (amino acid position 1) to an internal amino acid. Similarly, the C-terminal region refers to a region extending from an internal amino acid to the C-terminus.

[0050] The term "aptamer" refers to a single-stranded DNA that recognizes and binds to DNA polymerase, and efficiently inhibits the polymerase activity as described in U.S. Pat. No. 5,693,502, hereby expressly incorporated by reference herein in its entirety. Use of aptamer and dUTP/UNG in RT-PCR is also discussed, for example, in Smith, E. S. et al, (Amplification of RNA: High-temperature Reverse Transcription and DNA Amplification with a Magnesium-activated Thermostable DNA Polymerase, in PCR Primer: A Laboratory Manual, 2nd Edition, Dieffenbach, C. W. and Dveksler, G. S., Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 211-219, (2003)).

[0051] In the context of mutant DNA polymerases, "correspondence" to another sequence (e.g., regions, fragments, nucleotide or amino acid positions, or the like) is based on the convention of numbering according to nucleotide or amino acid position number and then aligning the sequences in a manner that maximizes the percentage of sequence identity. An amino acid "corresponding to position [X] of [specific sequence]" refers to an amino acid in a polypeptide of interest that aligns with the equivalent amino acid of a specified sequence. Generally, as described herein, the amino acid corresponding to a position of a polymerase can be determined using an alignment algorithm such as BLAST as described below. Because not all positions within a given "corresponding region" need be identical, non-matching positions within a corresponding region may be regarded as "corresponding positions." Accordingly, as used herein, referral to an "amino acid position corresponding to amino acid position [X]" of a specified DNA polymerase refers to equivalent positions, based on alignment, in other DNA polymerases and structural homologues and families. In some embodiments of the present invention, "correspondence" of amino acid positions are determined with respect to a region of the polymerase comprising one or more motifs of SEQ ID NO:1, 2, 3, 4, 5, 6, 7, 32, 33, 34, 35, 36, 37, or 39. When a polymerase polypeptide sequence differs from SEQ ID NOS:1, 2, 3, 4, 5, 6, 7, 32, 33, 34, 35, 36, 37, or 39 (e.g., by changes in amino acids or addition or deletion of amino acids), it may be that a particular mutation associated with improved activity as discussed herein will not be in the same position number as it is in SEQ ID NOS:1, 2, 3, 4, 5, 6, 7, 32, 33, 34, 35, 36, 37, or 39. This is illustrated, for example, in Table 1.

[0052] "Recombinant," as used herein, refers to an amino acid sequence or a nucleotide sequence that has been intentionally modified by recombinant methods. By the term "recombinant nucleic acid" herein is meant a nucleic acid, originally formed in vitro, in general, by the manipulation of a nucleic acid by restriction endonucleases, in a form not normally found in nature. Thus an isolated, mutant DNA polymerase nucleic acid, in a linear form, or an expression vector formed in vitro by ligating DNA molecules that are not normally joined, are both considered recombinant for the purposes of this invention. It is understood that once a recombinant nucleic acid is made and reintroduced into a host cell, it will replicate non-recombinantly, i.e., using the in vivo cellular machinery of the host cell rather than in vitro manipulations; however, such nucleic acids, once produced recombinantly, although subsequently replicated non-recombinantly, are still considered recombinant for the purposes of the invention. A "recombinant protein" is a protein made using recombinant techniques, i.e., through the expression of a recombinant nucleic acid as depicted above.

[0053] A nucleic acid is "operably linked" when it is placed into a functional relationship with another nucleic acid sequence. For example, a promoter or enhancer is operably linked to a coding sequence if it affects the transcription of the sequence; or a ribosome binding site is operably linked to a coding sequence if it is positioned so as to facilitate translation.

[0054] The term "host cell" refers to both single-cellular prokaryote and eukaryote organisms (e.g., bacteria, yeast, and actinomycetes) and single cells from higher order plants or animals when being grown in cell culture.

[0055] The term "vector" refers to a piece of DNA, typically double-stranded, which may have inserted into it a piece of foreign DNA. The vector or may be, for example, of plasmid origin. Vectors contain "replicon" polynucleotide sequences that facilitate the autonomous replication of the vector in a host cell. Foreign DNA is defined as heterologous DNA, which is DNA not naturally found in the host cell, which, for example, replicates the vector molecule, encodes a selectable or screenable marker, or encodes a transgene. The vector is used to transport the foreign or heterologous DNA into a suitable host cell. Once in the host cell, the vector can replicate independently of or coincidental with the host chromosomal DNA, and several copies of the vector and its inserted DNA can be generated. In addition, the vector can also contain the necessary elements that permit transcription of the inserted DNA into an mRNA molecule or otherwise cause replication of the inserted DNA into multiple copies of RNA. Some expression vectors additionally contain sequence elements adjacent to the inserted DNA that increase the half-life of the expressed mRNA and/or allow translation of the mRNA into a protein molecule. Many molecules of mRNA and polypeptide encoded by the inserted DNA can thus be rapidly synthesized.

[0056] The term "nucleotide," in addition to referring to the naturally occurring ribonucleotide or deoxyribonucleotide monomers, shall herein be understood to refer to related structural variants thereof, including derivatives and analogs, that are functionally equivalent with respect to the particular context in which the nucleotide is being used (e.g., hybridization to a complementary base), unless the context clearly indicates otherwise.

[0057] The term "nucleic acid" or "polynucleotide" refers to a polymer that can be corresponded to a ribose nucleic acid (RNA) or deoxyribose nucleic acid (DNA) polymer, or an analog thereof. This includes polymers of nucleotides such as RNA and DNA, as well as synthetic forms, modified (e.g., chemically or biochemically modified) forms thereof, and mixed polymers (e.g., including both RNA and DNA subunits). Exemplary modifications include methylation, substitution of one or more of the naturally occurring nucleotides with an analog, internucleotide modifications such as uncharged linkages (e.g., methyl phosphonates, phosphotriesters, phosphoamidates, carbamates, and the like), pendent moieties (e.g., polypeptides), intercalators (e.g., acridine, psoralen, and the like), chelators, alkylators, and modified linkages (e.g., alpha anomeric nucleic acids and the like). Also included are synthetic molecules that mimic polynucleotides in their ability to bind to a designated sequence via hydrogen bonding and other chemical interactions. Typically, the nucleotide monomers are linked via phosphodiester bonds, although synthetic forms of nucleic acids can comprise other linkages (e.g., peptide nucleic acids as described in Nielsen et al. (*Science* 254:1497-1500, 1991). A nucleic acid can be or can include, e.g., a chromosome or chromosomal segment, a vector (e.g., an expression vector), an expression cassette, a naked DNA or RNA polymer, the product of a polymerase chain reaction (PCR), an oligonucleotide, a probe, and a primer. A nucleic acid can be, e.g., single-stranded, double-stranded, or triple-stranded and is not limited to any particular length. Unless otherwise indicated, a particular nucleic acid sequence optionally comprises or encodes complementary sequences, in addition to any sequence explicitly indicated.

[0058] The term "oligonucleotide" refers to a nucleic acid that includes at least two nucleic acid monomer units (e.g., nucleotides). An oligonucleotide typically includes from about six to about 175 nucleic acid monomer units, more typically from about eight to about 100 nucleic acid monomer units, and still more typically from about 10 to about 50 nucleic acid monomer units (e.g., about 15, about 20, about 25, about 30, about 35, or more nucleic acid monomer units). The exact size of an oligonucleotide will depend on many factors, including the ultimate function or use of the oligonucleotide. Oligonucleotides are optionally prepared by any suitable method, including, but not limited to, isolation of an existing or natural sequence, DNA replication or amplification, reverse transcription, cloning and restriction digestion of appropriate sequences, or direct chemical synthesis by a method such as the phosphotriester method of Narang et al. (*Meth. Enzymol.* 68:90-99, 1979); the phosphodiester method of Brown et al. (*Meth. Enzymol.* 68:109-151, 1979); the diethylphosphoramidite method of Beaucage et al. (*Tetrahedron Lett.* 22:1859-1862, 1981); the triester method of Matteucci et al. (*J. Am. Chem. Soc.* 103:3185-3191, 1981); automated synthesis methods; or the solid support method of U.S. Pat. No. 4,458,066, entitled "PROCESS FOR PREPARING POLYNUCLEOTIDES," issued Jul. 3, 1984 to Caruthers et al., or other methods known to those skilled in the art. All of these references are incorporated by reference.

[0059] The term "primer" as used herein refers to a polynucleotide capable of acting as a point of initiation of template-directed nucleic acid synthesis when placed under conditions in which polynucleotide extension is initiated (e.g., under conditions comprising the presence of requisite nucleoside triphosphates (as dictated by the template that is copied) and a polymerase in an appropriate buffer and at a suitable temperature or cycle(s) of temperatures (e.g., as in a polymerase chain reaction)). To further illustrate, primers can also be used in a variety of other oligonuceotide-mediated synthesis processes, including as initiators of de novo RNA synthesis and in vitro transcription-related processes (e.g., nucleic acid sequence-based amplification (NASBA), transcription mediated amplification (TMA), etc.). A primer is typically a single-stranded oligonucleotide (e.g., oligodeoxyribonucleotide). The appropriate length of a primer depends on the intended use of the primer but typically ranges from 6 to 40 nucleotides, more typically from 15 to 35 nucleotides. Short primer molecules generally require cooler temperatures to form sufficiently stable hybrid complexes with the template. A primer need not reflect the exact sequence of the template but must be sufficiently complementary to hybridize with a template for primer elongation to occur. In certain embodiments, the term "primer pair" means a set of primers including a 5' sense primer (sometimes called "forward") that hybridizes with the complement of the 5' end of the nucleic acid sequence to be amplified and a 3' antisense primer (sometimes called "reverse") that hybridizes with the 3' end of the sequence to be amplified (e.g., if the target sequence is expressed as RNA or is an RNA). A primer can be labeled, if desired, by incorporating a label detectable by spectroscopic, photochemical, biochemical, immunochemical, or chemical means. For example, useful labels include ^{32}P, fluorescent dyes, electron-dense reagents, enzymes (as commonly used in ELISA assays), biotin, or haptens and proteins for which antisera or monoclonal antibodies are available.

[0060] The term "conventional" or "natural" when referring to nucleic acid bases, nucleoside triphosphates, or nucleotides refers to those which occur naturally in the polynucleotide being described (i.e., for DNA these are dATP, dGTP, dCTP and dTTP). Additionally, dITP, and 7-deaza-dGTP are frequently utilized in place of dGTP and 7-deaza-dATP can be utilized in place of dATP in in vitro DNA synthesis reactions, such as sequencing. Collectively, these may be referred to as dNTPs.

[0061] The term "unconventional" or "modified" when referring to a nucleic acid base, nucleoside, or nucleotide includes modification, derivations, or analogues of conventional bases, nucleosides, or nucleotides that naturally occur in a particular polynucleotide. Certain unconventional nucleotides are modified at the 2' position of the ribose sugar in comparison to conventional dNTPs. Thus, although for RNA the naturally occurring nucleotides are ribonucleotides (i.e., ATP, GTP, CTP, UTP, collectively rNTPs), because these nucleotides have a hydroxyl group at the 2' position of the sugar, which, by comparison is absent in dNTPs, as used herein, ribonucleotides are unconventional nucleotides as substrates for DNA polymerases. As used herein, unconventional nucleotides include, but are not limited to, compounds used as terminators for nucleic acid sequencing. Exemplary terminator compounds include but are not limited to those compounds that have a 2',3' dideoxy structure and are referred to as dideoxynucleoside triphosphates. The dideoxynucleoside triphosphates ddATP, ddTTP, ddCTP and ddGTP are referred to collectively as ddNTPs. Additional examples of terminator compounds include 2'-PO_4 analogs of ribonucleotides (see, e.g., U.S. Application Publication Nos. 2005/0037991 and 2005/0037398, which are both incorporated by reference). Other unconventional nucleotides include phosphorothioate dNTPs ([α-S]dNTPs), 5'-[α-borano]-dNTPs, [α]-methyl-phosphonate dNTPs, and ribonucleoside triphosphates (rNTPs). Unconventional bases may be labeled with radioactive isotopes such as ^{32}P, ^{33}P, or ^{35}S; fluorescent labels; chemiluminescent labels; bioluminescent labels; hapten labels such as biotin; or enzyme labels such as streptavidin or avidin. Fluorescent labels may include dyes that are negatively charged, such as dyes of the fluorescein family, or dyes that are neutral in charge, such as dyes of the rhodamine family, or dyes that are positively charged, such as dyes of the cyanine family. Dyes of the fluorescein family include, e.g., FAM, HEX, TET, JOE, NAN and ZOE. Dyes of the rhodamine family include Texas Red, ROX, R110, R6G, and TAMRA. Various dyes or nucleotides labeled with FAM, HEX, TET, JOE, NAN, ZOE, ROX, R110, R6G, Texas Red and TAMRA are marketed by Perkin-Elmer (Boston, Mass.), Applied Biosystems (Foster City, Calif.), or Invitrogen/Molecular Probes (Eugene, Oreg.). Dyes of the cyanine family include Cy2, Cy3, Cy5, and Cy7 and are marketed by GE Healthcare UK Limited (Amersham Place, Little Chalfont, Buckinghamshire, England).

[0062] As used herein, "percentage of sequence identity" is determined by comparing two optimally aligned sequences over a comparison window, wherein the portion of the sequence in the comparison window can comprise additions or deletions (i.e., gaps) as compared to the reference sequence (which does not comprise additions or deletions) for optimal alignment of the two sequences. The percentage is calculated by determining the number of positions at which the identical nucleic acid base or amino acid residue occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the window of comparison and multiplying the result by 100 to yield the percentage of sequence identity.

[0063] The terms "identical" or percent "identity," in the context of two or more nucleic acids or polypeptide sequences, refer to two or more sequences or subsequences that are the same. Sequences are "substantially identical" to each other if they have a specified percentage of nucleotides or amino acid residues that are the same (e.g., at least 20%, at least 25%, at least 30%, at least 35%, at least 40%, at least 45%, at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, or at least 95% identity over a specified region)), when compared and aligned for maximum correspondence over a comparison window, or designated region as measured using one of the following sequence comparison algorithms or by manual alignment and visual inspection. Sequences are "substantially identical" to each other if they are at least 20%, at least 25%, at least 30%, at least 35%, at least 40%, at least 45%, at least 50%, or at least 55% identical. These definitions also refer to the complement of a test sequence. Optionally, the identity exists over a region that is at least about 50 nucleotides in length, or more typically over a region that is 100 to 500 or 1000 or more nucleotides in length.

[0064] The terms "similarity" or "percent similarity," in the context of two or more polypeptide sequences, refer to two or more sequences or subsequences that have a specified percentage of amino acid residues that are either the same or similar as defined by a conservative amino acid substitutions (e.g., 60% similarity, optionally 65%, 70%, 75%, 80%, 85%, 90%, or 95% similar over a specified region), when compared and aligned for maximum correspondence over a comparison window, or designated region as measured using one of the following sequence comparison algorithms or by manual alignment and visual inspection. Sequences are "substantially similar" to each other if they are at least 20%, at least 25%, at least 30%, at least 35%, at least 40%, at least 45%, at least 50%, or at least 55% similar to each other. Optionally, this similarly exists over a region that is at least about 50 amino acids in length, or more typically over a region that is at least about 100 to 500 or 1000 or more amino acids in length.

[0065] For sequence comparison, typically one sequence acts as a reference sequence, to which test sequences are compared. When using a sequence comparison algorithm, test and reference sequences are entered into a computer, subsequence coordinates are designated, if necessary, and sequence algorithm program parameters are designated. Default program parameters are commonly used, or alternative parameters can be designated. The sequence comparison algorithm then calculates the percent sequence identities or similarities for the test sequences relative to the reference sequence, based on the program parameters.

[0066] A "comparison window," as used herein, includes reference to a segment of any one of the number of contiguous positions selected from the group consisting of from 20 to 600, usually about 50 to about 200, more usually about 100 to about 150 in which a sequence may be compared to a reference sequence of the same number of contiguous positions after the two sequences are optimally aligned. Methods of alignment of sequences for comparison are well known in the art. Optimal alignment of sequences for comparison can be conducted, for example, by the local homology algorithm of Smith and Waterman (*Adv. Appl. Math*. 2:482, 1970), by the homology alignment algorithm of Needleman and Wunsch (*J. Mol. Biol*. 48:443, 1970), by the search for similarity method of Pearson and Lipman (*Proc. Natl. Acad. Sci. USA* 85:2444, 1988), by computerized implementations of these algorithms (e.g., GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, Wis.), or by manual alignment and visual inspection (see, e.g., Ausubel et al., *Current Protocols in Molecular Biology* (1995 supplement)).

[0067] Examples of an algorithm that is suitable for determining percent sequence identity and sequence similarity are the BLAST and BLAST 2.0 algorithms, which are described in Altschul et al. (*Nuc. Acids Res*. 25:3389-402, 1977), and Altschul et al. (*J. Mol. Biol*. 215:403-10, 1990), respectively. Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information (see the internet at www.ncbi.nlm.nih.gov/). This algorithm involves first identifying high scoring sequence pairs (HSPs) by identifying short words of length W in the query sequence, which either match or satisfy some positive-valued threshold score T when aligned with a word of the same length in a database sequence. T is referred to as the neighborhood word score threshold (Altschul et al., supra). These initial neighborhood word hits act as seeds for initiating searches to find longer HSPs containing them. The word hits are extended in both directions along each sequence for as far as the cumulative alignment score can be increased. Cumulative scores are calculated using, for nucleotide sequences, the parameters M (reward score for a pair of matching residues; always >0) and N (penalty score for mismatching residues; always <0). For amino acid sequences, a scoring matrix is used to calculate the cumulative score. Extension of the word hits in each direction are halted when: the cumulative alignment score falls off by the quantity X from its maximum achieved value; the cumulative score goes to zero or below, due to the accumulation of one or more negative-scoring residue alignments; or the end of either sequence is reached. The BLAST algorithm parameters W, T, and X determine the sensitivity and speed of the alignment. The BLASTN program (for nucleotide sequences) uses as defaults a wordlength (W) of 11, an expectation (E) or 10, M=5, N=−4 and a comparison of both strands. For amino acid sequences, the BLASTP program uses as defaults a wordlength of 3, and expectation (E) of 10, and the BLOSUM62 scoring matrix (see Henikoff and Henikoff, *Proc. Natl. Acad. Sci. USA*

89:10915, 1989) alignments (B) of 50, expectation (E) of 10, M=5, N=−4, and a comparison of both strands.

[0068] The BLAST algorithm also performs a statistical analysis of the similarity between two sequences (see, e.g., Karlin and Altschul, *Proc. Natl. Acad. Sci. USA* 90:5873-87, 1993). One measure of similarity provided by the BLAST algorithm is the smallest sum probability (P(N)), which provides an indication of the probability by which a match between two nucleotide or amino acid sequences would occur by chance. For example, a nucleic acid is considered similar to a reference sequence if the smallest sum probability in a comparison of the test nucleic acid to the reference nucleic acid is less than about 0.2, typically less than about 0.01, and more typically less than about 0.001.

[0069] The term "reverse transcription efficiency" refers to the fraction of RNA molecules that are reverse transcribed as cDNA in a given reverse transcription reaction. In certain embodiments, the mutant DNA polymerases of the invention have improved reverse transcription efficiencies relative to unmodified forms of these DNA polymerases. That is, these mutant DNA polymerases reverse transcribe a higher fraction of RNA templates than their unmodified forms under a particular set of reaction conditions. Without being limited by theory, the ability of a mutant DNA polymerase described herein to reverse transcribe a higher fraction of RNA templates can be due to an increased reverse transcription activity, for example, an increased nucleotide incorporation rate and/or increased processivity of the enzyme. Reverse transcription efficiency can be measured, for example, by measuring the crossing point (Cp) of a PCR reaction using a RNA template, and comparing the Cp value to a Cp value of a control reaction in which a DNA template of the same sequence (except U's are replaced with T's) is amplified, wherein the RNA and DNA amplifications use a common primer set and the same polymerase, e.g., as described in the examples. A test polymerase has improved RT efficiency when the test polymerase has a decreased Cp value compared to a control polymerase when RNA is used as a template, but has a substantially unchanged Cp value relative to the control polymerase when DNA is used as a template. In some embodiments a polymerase of the invention has an improved RT efficiency such that the Cp is at least one, two, three, four, five, six, seven, eight, nine, ten or more units less than the corresponding control polymerase on the RNA template. Improved RT efficiency of a test polymerase can be measured as described in the Examples.

[0070] The term "mismatch tolerance" refers to the ability of a polymerase to tolerate a mismatch-containing sequence when extending a nucleic acid (e.g., a primer or other oligonucleotide) in a template-dependent manner by attaching (e.g., covalently) one or more nucleotides to the nucleic acid. The term "3' mismatch tolerance" refers to the ability of a polymerase to tolerate a mismatch-containing (nearly complementary) sequence where the nucleic acid to be extended (e.g., a primer or other oligonucleotide) has a mismatch with its template at the 3' terminal nucleotide of the primer. Mismatches to the template may also be located at the 3' penultimate nucleotide of the primer, or at another position within the sequence of the primer.

[0071] The term "mismatch discrimination" refers to the ability of a polymerase to distinguish a fully complementary sequence from a mismatch-containing sequence when extending a nucleic acid (e.g., a primer or other oligonucleotide) in a template-dependent manner by attaching (e.g., covalently) one or more nucleotides to the nucleic acid. The term "3'-mismatch discrimination" refers to the ability of a polymerase to distinguish a fully complementary sequence from a mismatch-containing (nearly complementary) sequence where the nucleic acid to be extended (e.g., a primer or other oligonucleotide) has a mismatch at the nucleic acid's 3' terminus compared to the template to which the nucleic acid hybridizes. The term "mismatch" refers to the existence of one or more base mispairings (or "noncomplementary base oppositions") within a stretch of otherwise complementary duplex-forming (or potentially duplex-forming) sequences.

[0072] The term "Cp value" or "crossing point" value refers to a value that allows quantification of input target nucleic acids. The Cp value can be determined according to the second-derivative maximum method (Van Luu-The, et al., "Improved real-time RT-PCR method for high-throughput measurements using second derivative calculation and double correction," BioTechniques, Vol. 38, No. 2, February 2005, pp. 287-293). In the second derivative method, a Cp corresponds to the first peak of a second derivative curve. This peak corresponds to the beginning of a log-linear phase. The second derivative method calculates a second derivative value of the real-time fluorescence intensity curve, and only one value is obtained. The original Cp method is based on a locally defined, differentiable approximation of the intensity values, e.g., by a polynomial function. Then the third derivative is computed. The Cp value is the smallest root of the third derivative. The Cp can also be determined using the fit point method, in which the Cp is determined by the intersection of a parallel to the threshold line in the log-linear region (Van Luu-The, et al., BioTechniques, Vol. 38, No. 2, February 2005, pp. 287-293). The Cp value provided by the LightCycler instrument offered by Roche by calculation according to the second-derivative maximum method.

[0073] The term "PCR efficiency" refers to an indication of cycle to cycle amplification efficiency. PCR efficiency is calculated for each condition using the equation: % PCR efficiency=$(10^{(-slope)}-1)\times100$, wherein the slope was calculated by linear regression with the log copy number plotted on the y-axis and Cp plotted on the x-axis. PCR efficiency can be measured using a perfectly matched or mismatched primer template.

[0074] The term "nucleic acid extension rate" refers the rate at which a biocatalyst (e.g., an enzyme, such as a polymerase, ligase, or the like) extends a nucleic acid (e.g., a primer or other oligonucleotide) in a template-dependent or template-independent manner by attaching (e.g., covalently) one or more nucleotides to the nucleic acid. To illustrate, certain mutant DNA polymerases described herein have improved nucleic acid extension rates relative to unmodified forms of these DNA polymerases, such that they can extend primers at higher rates than these unmodified forms under a given set of reaction conditions.

[0075] The term "tolerance of RT and polymerase inhibitors" refers to the ability of a polymerase to maintain activity (polymerase or reverse transcription activity) in the presence of an amount of an inhibitor that would inhibit the polymerase activity or reverse transcription activity of a control polymerase. In some embodiments, the improved polymerase is capable of polymerase or reverse transcription activity in the presence of an amount of the inhibitor that would essentially eliminate the control polymerase activity.

[0076] The term "5'-nuclease probe" refers to an oligonucleotide that comprises at least one light emitting labeling moiety and that is used in a 5'-nuclease reaction to effect target nucleic acid detection. In some embodiments, for example, a 5'-nuclease probe includes only a single light emitting moiety (e.g., a fluorescent dye, etc.). In certain embodiments, 5'-nuclease probes include regions of self-complementarity such that the probes are capable of forming hairpin structures under selected conditions. To further illustrate, in some embodiments a 5'-nuclease probe comprises at least two labeling moieties and emits radiation of increased intensity after one of the two labels is cleaved or otherwise separated from the oligonucleotide. In certain embodiments, a 5'-nuclease probe is labeled with two different fluorescent dyes, e.g., a 5' terminus reporter dye and the 3' terminus quencher dye or moiety. In some embodiments, 5'-nuclease probes are labeled at one or more positions other than, or in addition to, terminal positions. When the probe is intact, energy transfer typically occurs between the two fluorophores such that fluorescent emission from the reporter dye is quenched at least in part. During an extension step of a polymerase chain reaction, for example, a 5'-nuclease probe bound to a template nucleic acid is cleaved by the 5' to 3' nuclease activity of, e.g., a Taq polymerase or another polymerase having this activity such that the fluorescent emission of the reporter dye is no longer quenched. Exemplary 5'-nuclease probes are also described in, e.g., U.S. Pat. No. 5,210,015, entitled "Homogeneous assay system using the nuclease activity of a nucleic acid polymerase," issued May 11, 1993 to Gelfand et al., U.S. Pat. No. 5,994,056, entitled "Homogeneous methods for nucleic acid amplification and detection," issued Nov. 30, 1999 to Higuchi, and U.S. Pat. No. 6,171,785, entitled "Methods and devices for homogeneous nucleic acid amplification and detector," issued Jan. 9, 2001 to Higuchi, which are each incorporated by reference herein. In other embodiments, a 5' nuclease probe may be labeled with two or more different reporter dyes and a 3' terminus quencher dye or moiety.

[0077] The term "FRET" or "fluorescent resonance energy transfer" or "Foerster resonance energy transfer" refers to a transfer of energy between at least two chromophores, a donor chromophore and an acceptor chromophore (referred to as a quencher). The donor typically transfers the energy to the acceptor when the donor is excited by light radiation with a suitable wavelength. The acceptor typically re-emits the transferred energy in the form of light radiation with a different wavelength. When the acceptor is a "dark" quencher, it dissipates the transferred energy in a form other than light. Whether a particular fluorophore acts as a donor or an acceptor depends on the properties of the other member of the FRET pair. Commonly used donor-acceptor pairs include the FAM-TAMRA pair. Commonly used quenchers are DABCYL and TAMRA. Commonly used dark quenchers include BlackHole Quenchers™ (BHQ), (Biosearch Technologies, Inc., Novato, Calif.), Iowa Black™ (Integrated DNA Tech., Inc., Coralville, Iowa), and BlackBerry™ Quencher 650 (BBQ-650) (Berry & Assoc., Dexter, Mich.).

BRIEF DESCRIPTION OF THE DRAWINGS

[0078] FIG. **1** depicts an amino acid sequence alignment of a region from the polymerase domain of exemplary DNA polymerases from various species of bacteria: *Thermus* species Z05 (Z05) (SEQ ID NO:12), *Thermus aquaticus* (Taq) (SEQ ID NO:13), *Thermus filiformus* (Tfi) (SEQ ID NO:14), *Thermus flavus* (Tfl) (SEQ ID NO:15), *Thermus* species sps17 (Sps17) (SEQ ID NO:16), *Thermus thermophilus* (Tth) (SEQ ID NO:17), *Thermus caldophilus* (Tca) (SEQ ID NO:18), *Thermotoga maritima* (Tma) (SEQ ID NO:19), *Thermotoga neopolitana* (Tne) (SEQ ID NO:20), *Thermosipho africanus* (Taf) (SEQ ID NO:21), *Deinococcus radiodurans* (Dra) (SEQ ID NO:23), *Bacillus stearothermophilus* (Bst) (SEQ ID NO:24), and *Bacillus caldotenax* (Bca) (SEQ ID NO:25). In addition, the polypeptide regions shown comprise the amino acid motif X_1-X_2-X_3-F-X_4-X_5-X_6-X_7-D-X_8-H-T-X_9-T-A-X_{10}-X_{11} (SEQ ID NO:26), the variable positions of which are further defined herein. This motif is highlighted in bold type for each polymerase sequence. Amino acid positions amenable to mutation in accordance with the present invention are indicated with an asterisk (*). Gaps in the alignments are indicated with a dot (.).

[0079] FIG. **2** provides sequence identities among the following DNA Polymerase I enzymes: *Thermus* sp. Z05 DNA polymerase (Z05); *Thermus aquaticus* DNA polymerase (Taq); *Thermus filiformis* DNA polymerase (Tfi); *Thermus flavus* DNA polymerase (Tfl); *Thermus* sp. sps17 DNA polymerase (Sps17); *Thermus thermophilus* DNA polymerase (Tth); *Thermus caldophilus* DNA polymerase (Tca); *Deinococcus radiodurans* DNA polymerase (Dra); *Thermotoga maritima* DNA polymerase (Tma); *Thermotoga neopolitana* DNA polymerase (Tne); *Thermosipho africanus* DNA polymerase (Taf); *Bacillus stearothermophilus* DNA polymerase (Bst); and *Bacillus caldotenax* DNA polymerase (Bca). (A) sequence identities over the entire polymerase I enzyme (corresponding to amino acids 1-834 of Z05); and (B) sequence identities over the polymerase sub domain corresponding to amino acids 420-834 of Z05.

[0080] FIG. **3** provides sequence identities among various *Thermus* sp DNA Polymerase I enzymes: *Thermus* sp. Z05 DNA polymerase (Z05); *Thermus aquaticus* DNA polymerase (Taq); *Thermus filiformis* DNA polymerase (Tfi); *Thermus flavus* DNA polymerase (Tfl); *Thermus* sp. sps17 DNA polymerase (Sps17); *Thermus thermophilus* DNA polymerase (Tth); and *Thermus caldophilus* DNA polymerase (Tca). (A) sequence identities over the entire polymerase I enzyme (corresponding to amino acids 1-834 of Z05); and (B) sequence identities over the polymerase sub domain corresponding to amino acids 420-834 of Z05.

DETAILED DESCRIPTION

[0081] The present invention provides improved DNA polymerases in which one or more amino acids in the polymerase domain have been mutated relative to a functional DNA polymerase. The DNA polymerases of the invention are active enzymes having increased reverse transcriptase efficiency (e.g., in the presence of Mn^{2+} and Mg^{2+} divalent cations) relative to the unmodified form of the polymerase and/or increased mismatch tolerance, extension rate and tolerance of RT and polymerase inhibitors. In certain embodiments, the mutant DNA polymerases may be used at lower concentrations for superior or equivalent performance as the parent enzymes. In some embodiments, the mutant DNA polymerases have increased reverse transcriptase efficiency while retaining substantially the same DNA-dependent polymerase activity relative to an unmodified or control polymerase.

[0082] DNA polymerases that more efficiently perform reverse transcription are helpful, for example, in a variety of applications involving assays that employ RT-PCR to detect and/or quantify RNA targets. The DNA polymerases are therefore useful in a variety of applications involving polynucleotide extension as well as reverse transcription or amplification of polynucleotide templates, including, for example, applications in recombinant DNA studies and medical diagnosis of disease. The mutant DNA polymerases are also particularly useful, because of their tolerance for mis-matches, for detecting targets that possibly have variable sequences (e.g., viral targets, or cancer and other disease genetic markers).

[0083] In some embodiments, DNA polymerases of the invention can be characterized by having the following motif:

[0084] X_1-X_2-X_3-Phe-X_4-X_5-X_6-X_7-Asp-X_8-His-Thr-X_9-Thr-Ala-X_{10}-X_{11} (also referred to herein in the one-letter code as X_1-X_2-X_3-F-X_4-X_5-X_6-X_7-D-X_8-H-T-X_9-T-A-X_{10}-X_{11} (SEQ ID NO:8);

[0085] wherein:

[0086] X_1 is Ile (I), Leu (L), Val (V), Gln (Q) or Met (M);

[0087] X_2 is Arg (R), Lys (K), Gln (Q), or Glu (E);

[0088] X_3 is Val (V) or Ala (A);

[0089] X_4 is Gln (Q), Arg (R), Glu (E), Lys (K) or Val (V);

[0090] X_5 is Glu (E) or Arg (R);

[0091] X_6 is Gly (G) or Asp (D);

[0092] X_7 is Lys (K), Arg (R), Ile (I), Leu (L), Ala (A);

[0093] X_8 is any amino acid other than Ile (I) or Val (V);

[0094] X_9 is Gln (Q), Glu (E), Leu (L), Ile (I), Arg (R) or Lys (K);

[0095] X_{10} is Ser (S), Ala (A) or Met (M);

[0096] X_{11} is Trp (W), Arg (R), Lys (K), Gln (Q) or Asp (D).

[0097] In some embodiments, X_8 is selected from G, A, W, P, S, T, F, Y, C, N, Q, D, E, K, R, L, M, or H.

[0098] In some embodiments, DNA polymerases of the invention can be characterized by having the following motif:

[0099] Ile-Arg-Val-Phe-X_4-Glu-Gly-X_7-Asp-X_8-His-Thr-X_9-Thr-Ala-X_{10}-Trp (also referred to herein in the one-letter code as I-R-V-F-X_4-E-G-X_7-D-X_8-H-T-X_9-T-A-X_{10}-W (SEQ ID NO:9);

[0100] wherein:

[0101] X_4 is Gln (Q) or Arg (R);

[0102] X_7 is Lys (K) or Arg (R);

[0103] X_8 is any amino acid other than Ile (I);

[0104] X_9 is Gln (Q) or Glu (E);

[0105] X_{10} is Ser (S) or Ala (A).

[0106] In some embodiments, DNA polymerases of the invention can be characterized by having the following motif:

[0107] Ile-Arg-Val-Phe-Gln-Glu-Gly-Lys-Asp-X_8-His-Thr-Gln-Thr-Ala-Ser-Trp (also referred to herein in the one-letter code as I-R-V-F-Q-E-G-K-D-X_8-H-T-Q-T-A-S-W (SEQ ID NO:10);

[0108] wherein:

[0109] X_8 is any amino acid other than Ile (I).

[0110] In some embodiments, DNA polymerases of the invention can be characterized by having the following motif:

[0111] Ile-Arg-Val-Phe-Gln-Glu-Gly-Lys-Asp-X_8-His-Thr-Gln-Thr-Ala-Ser-Trp (also referred to herein in the one-letter code as I-R-V-F-Q-E-G-K-D-X_8-H-T-Q-T-A-S-W (SEQ ID NO:11);

[0112] wherein:

[0113] X_8 is Phe (F).

[0114] In some embodiments, DNA polymerases of the invention can be characterized by having the above motifs (e.g., SEQ ID NOs:8, 9, 10, and 11), optionally in combination with additional motifs described below. For example, in some embodiments, the DNA polymerase further comprises the motif of SEQ ID NO:29 and/or SEQ ID NO:38.

[0115] This motif is present within the "fingers" domain (N alpha helix) of many Family A type DNA-dependent DNA polymerases, particularly thermostable DNA polymerases from thermophilic bacteria (Li et al., *EMBO J.* 17:7514-7525, 1998). For example, FIG. **1** shows an amino acid sequence alignment of a region from the "fingers" domain of DNA polymerases from several species of bacteria: *Bacillus caldotenax, Bacillus stearothermophilus, Deinococcus radiodurans, Thermosipho africanus, Thermotoga maritima, Thermotoga neopolitana, Thermus aquaticus, Thermus caldophilus, Thermus filiformus, Thermus flavus, Thermus* sp. sps17, *Thermus* sp. Z05, and *Thermus thermophilus*. As shown, the native sequence corresponding to the motif above is present in each of these polymerases, indicating a conserved function for this region of the polymerase. FIG. **2** provides sequence identities among these DNA polymerases.

[0116] Accordingly, in some embodiments, the invention provides for a polymerase comprising SEQ ID NO:8, 9, 10, or 11, having the improved activity and/or characteristics described herein, and wherein the DNA polymerase is otherwise a wild-type or a naturally occurring DNA polymerase, such as, for example, a polymerase from any of the species of thermophilic bacteria listed above, or is substantially identical to such a wild-type or a naturally occurring DNA polymerase. For example, in some embodiments, the polymerase of the invention comprises SEQ ID NO:8, 9, 10, or 11 and is at least 80%, 85%, 90%, or 95% identical to SEQ ID NO:1, 2, 3, 4, 5, 6, 7, 32, 33, 34, 35, 36, 37, or 39. In one variation, the unmodified form of the polymerase is from a species of the genus *Thermus*. In other embodiments of the invention, the unmodified polymerase is from a thermophilic species other than *Thermus*, e.g., *Thermotoga*. The full nucleic acid and amino acid sequence for numerous thermostable DNA polymerases are available. The sequences each of *Thermus aquaticus* (Taq) (SEQ ID NO:2), *Thermus thermophilus* (Tth) (SEQ ID NO:6), *Thermus* species Z05 (SEQ ID NO:1), *Thermus* species sps17 (SEQ ID NO:5), *Thermotoga maritima* (Tma) (SEQ ID NO:34), and *Thermosipho africanus* (Taf) (SEQ ID NO:33) polymerase have been published in PCT International Patent Publication No. WO 92/06200, which is incorporated herein by reference. The sequence for the DNA polymerase from *Thermus flavus* (SEQ ID NO:4) has been published in Akhmetzjanov and Vakhitov (*Nucleic Acids Research* 20:5839, 1992), which is incorporated herein by reference. The sequence of the thermostable DNA polymerase from *Thermus caldophilus* (SEQ ID NO:7) is found in EMBL/GenBank Accession No. U62584. The sequence of the thermostable DNA polymerase from *Thermus filiformis* can be recovered from ATCC Deposit No. 42380 using, e.g., the methods provided in U.S. Pat. No. 4,889,818, as well as the sequence information provided in Table 1. The sequence of the *Thermotoga neapolitana* DNA polymerase (SEQ ID NO:35) is from GeneSeq Patent Data Base Accession No. R98144 and PCT WO 97/09451, each incorporated herein by reference. The sequence of the thermostable DNA polymerase from *Bacillus caldotenax* (SEQ ID NO:37 is described in, e.g., Uemori et al. (*J Biochem* (*Tokyo*) 113(3): 401-410, 1993; see also, Swiss-Prot database Accession No. Q04957 and GenBank Accession Nos. D12982 and BAA02361), which are each incorporated by reference. Examples of unmodified forms of DNA polymerases that can be modified as described herein are also described in, e.g., U.S. Pat. No. 6,228,628, entitled "Mutant chimeric DNA polymerase" issued May 8, 2001 to Gelfand et al.; U.S. Pat.

No. 6,346,379, entitled "Thermostable DNA polymerases incorporating nucleoside triphosphates labeled with fluorescein family dyes" issued Feb. 12, 2002 to Gelfand et al.; U.S. Pat. No. 7,030,220, entitled "Thermostable enzyme promoting the fidelity of thermostable DNA polymerases-for improvement of nucleic acid synthesis and amplification in vitro" issued Apr. 18, 2006 to Ankenbauer et al.; U.S. Pat. No. 6,881,559 entitled "Mutant B-type DNA polymerases exhibiting improved performance in PCR" issued Apr. 19, 2005 to Sobek et al.; U.S. Pat. No. 6,794,177 entitled "Modified DNA-polymerase from *carboxydothermus hydrogenoformans* and its use for coupled reverse transcription and polymerase chain reaction" issued Sep. 21, 2004 to Markau et al.; U.S. Pat. No. 6,468,775, entitled "Thermostable DNA polymerase from *carboxydothermus hydrogenoformans*" issued Oct. 22, 2002 to Ankenbauer et al.; and U.S. Pat. No. 7,148,049 entitled "Thermostable or thermoactive DNA polymerase molecules with attenuated 3'-5' exonuclease activity" issued Dec. 12, 2006 to Schoenbrunner et al.; U.S. Pat. No. 7,179,590 entitled "High temperature reverse transcription using mutant DNA polymerases" issued Feb. 20, 2007 to Smith et al.; U.S. Pat. No. 7,410,782 entitled "Thermostable enzyme promoting the fidelity of thermostable DNA polymerases-for improvement of nucleic acid synthesis and amplification in vitro" issued Aug. 12, 2008 to Ankenbauer et al.; U.S. Pat. No. 7,378,262 entitled "Reversibly modified thermostable enzymes for DNA synthesis and amplification in vitro" issued May 27, 2008 to Sobek et al., which are each incorporated by reference. Representative full length polymerase sequences are also provided in the sequence listing.

[0117] Also amenable to the mutations described herein are functional DNA polymerases that have been previously modified (e.g., by amino acid substitution, addition, or deletion). In some embodiments, such functional modified polymerases retain the amino acid motif of SEQ ID NO:8 (or a motif of SEQ ID NO:9, 10 or 11), and optionally the amino acid motif of SEQ ID NO:38. Thus, suitable unmodified DNA polymerases also include functional variants of wild-type or naturally occurring polymerases. Such variants typically will have substantial sequence identity or similarity to the wild-type or naturally occurring polymerase, typically at least 80% sequence identity and more typically at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% sequence identity.

[0118] In some embodiments, the polymerase of the invention, as well as having a polymerase domain comprising SEQ ID NOS:8, 9, 10, or 11 also comprises a nuclease domain (e.g., corresponding to positions 1 to 291 of Z05)

[0119] In some embodiments, a polymerase of the invention is a chimeric polymerase, i.e., comprising polypeptide regions from two or more enzymes. Examples of such chimeric DNA polymerases are described in, e.g., U.S. Pat. No. 6,228,628, which is incorporated by reference herein in its entirety. Particularly suitable are chimeric CS-family DNA polymerases, which include the CS5 (SEQ ID NO:27) and CS6 (SEQ ID NO:28) polymerases and variants thereof having substantial amino acid sequence identity or similarity to SEQ ID NO:27 or SEQ ID NO:28 (typically at least 80% amino acid sequence identity and more typically at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% amino acid sequence identity) and can thus be modified to contain SEQ ID NO:8. The CS5 and CS6 DNA polymerases are chimeric enzymes derived from *Thermus* sp. Z05 and *Thermotoga maritima* (Tma) DNA polymerases. They comprise the N-terminal 5'-nuclease domain of the *Thermus* enzyme and the C-terminal 3'-5' exonuclease and the polymerase domains of the Tma enzyme. These enzymes have efficient reverse transcriptase activity, can extend nucleotide analog-containing primers, and can incorporate alpha-phosphorothioate dNTPs, dUTP, dITP, and also fluorescein- and cyanine-dye family labeled dNTPs. The CS5 and CS6 polymerases are also efficient Mg^{2+}-activated PCR enzymes. The CS5 and CS6 chimeric polymerases are further described in, e.g., U.S. Pat. No. 7,148,049, which is incorporated by reference herein in its entirety.

[0120] In some embodiments, the amino acid substitutions are single amino acid substitutions. The DNA polymerases provided herein can comprise one or more amino acid substitutions in the active site relative to the unmodified polymerase. In some embodiments, the amino acid substitution(s) comprise at least position X_8 of the motif set forth in SEQ ID NO:8 (or a motif of SEQ ID NO:9, 10 or 11). Amino acid substitution at this position confers increased reverse transcriptase efficiency, mismatch tolerance, extension rate and/or tolerance of RT and polymerase inhibitors, yielding a mutant DNA polymerase with an increased reverse transcriptase efficiency, mismatch tolerance, extension rate and/or tolerance of RT and polymerase inhibitors relative to the unmodified polymerase. Typically, the amino acid at position X_8 is substituted with an amino acid that does not correspond to the native sequence within the motif set forth in SEQ ID NO:8 (or a motif of SEQ ID NO:9, 10 or 11). Thus, typically, the amino acid at position X_8, if substituted, is not Ile (I) or Val (V), as I or V occurs at this position in naturally-occurring polymerases. See, e.g., FIG. **1**. In certain embodiments, amino acid substitutions include G, A, W, P, S, T, F, Y, C, N, Q, D, E, K, R, L, M, or H at position X_8. In certain embodiments, amino acid substitutions include Phenylalanine (F) at position X_8. Other suitable amino acid substitution(s) at one or more of the identified sites can be determined using, e.g., known methods of site-directed mutagenesis and determination of polynucleotide extension performance in assays described further herein or otherwise known to persons of skill in the art.

[0121] In some embodiments, the polymerase of the invention comprises SEQ ID NO:8, 9, 10, or 11 and further comprises one or more additional amino acid changes (e.g., by amino acid substitution, addition, or deletion) compared to a native polymerase. In some embodiments, such polymerases retain the amino acid motif of SEQ ID NO:8 (or a motif of SEQ ID NO:9, 10 or 11), and further comprise the amino acid motif of SEQ ID NO:38 (corresponding to the D580X mutation of Z05 (SEQ ID NO:1)) as follows:

[0122] Thr-Gly-Arg-Leu-Ser-Ser-X_{b7}-X_{b8}-Pro-Asn-Leu-Gln-Asn (also referred to herein in the one-letter code as T-G-R-L-S-S-X_{b7}-X_{b8}-P-N-L-Q-N) (SEQ ID NO:38); wherein

[0123] X_{b7} is Ser (S) or Thr (T); and

[0124] X_{b8} is any amino acid other than Asp (D) or Glu (E)

The mutation characterized by SEQ ID NO:38 is discussed in more detail in, e.g., US Patent Publication No. 2009/0148891. Such functional variant polymerases typically will have substantial sequence identity or similarity to the wild-type or naturally occurring polymerase (e.g., SEQ ID NO:1, 2, 3, 4, 5, 6, 7, 32, 33, 34, 35, 36, 37, or 39), typically at least 80% amino acid sequence identity and more typically at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% amino acid sequence identity.

[0125] In some embodiments, the polymerase of the invention comprises SEQ ID NO:8, 9, 10, or 11 and further comprises the amino acid motif of SEQ ID NO:29 (corresponding to the I709X mutation of Z05 (SEQ ID NO:1) as follows:

[0126] X_1-X_2-X_3-X_4-X_5-X_6-X_7-X_8-X_9-X_{10}-X_{11}-X_{12}-X_{13}-Gly-Tyr-Val-X_{14}-Thr-Leu (also referred to herein in the one-letter code as X_1-X_2-X_3-X_{a4}-X_5-X_6-X_7-X_8-X_9-X_{10}-X_{11}-X_{12}-X_{13}-Gly-Val-X_{14}-T-L) (SEQ ID NO:29); wherein

[0127] X_1 is Ala (A), Asp (D), Ser (S), Glu (E), Arg (R) or Gln (Q);

[0128] X_2 is Trp (W) or Tyr (Y);

[0129] X_3 is any amino acid other than Ile (I), Leu (L) or Met (M);

[0130] X_4 is Glu (E), Ala (A), Gln (Q), Lys (K), Asn (N) or Asp (D);

[0131] X_5 is Lys (K), Gly (G), Arg (R), Gln (Q), His (H) or Asn (N);

[0132] X_6 is Thr (T), Val (V), Met (M) or Ile (I);

[0133] X_7 is Leu (L), Val (V) or Lys (K);

[0134] X_8 is Glu (E), Ser (S), Ala (A), Asp (D) or Gln (Q);

[0135] X_9 is Glu (E) or Phe (F);

[0136] X_{10} is Gly (G) or Ala (A);

[0137] X_{11} is Arg (R) or Lys (K);

[0138] X_{12} is Lys (K), Arg (R), Glu (E), Thr (T) or Gln (Q);

[0139] X_{13} is Arg (R), Lys (K) or His (H); and

[0140] X_{14} is Glu (E), Arg (R) or Thr (T).

[0141] In some embodiments, such functional variant polymerases typically will have substantial sequence identity or similarity to the wild-type or naturally occurring polymerase (e.g., SEQ ID NO: 1, 2, 3, 4, 5, 6, 7, 32, 33, 34, 35, 36, 37, or 39), typically at least 80% amino acid sequence identity and more typically at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% amino acid sequence identity.

[0142] In some embodiments, the DNA polymerase of the invention comprises an amino acid substitution at position X_8 (e.g., as in a motif selected from SEQ ID NO:8, 9, 10 or 11) and comprises an amino acid substitution corresponding to SEQ ID NO:38 and SEQ ID NO:29.

[0143] In some embodiments, the amino acid at position X_{b8} is substituted with an amino acid as set forth in SEQ ID NO:8, 9, 10 or 11, and the amino acid at position X_{b8} is substituted with an amino acid as set forth in SEQ ID NO:38. Thus, in some embodiments, the amino acid at position X_8 is any amino acid other than Ile (I) and the amino acid at position X_{b8} is any amino acid other than Asp (D) or Glu (E). In some embodiments, amino acid substitutions include Leucine (L), Glycine (G), Threonine (T), Glutamine (Q), Alanine (A), Serine (S), Asparagine (N), Arginine (R), and Lysine (K) at position X_{b8} of SEQ ID NO:38. In certain embodiments, amino acid substitutions independently include Methionine (M) at position X_8 of SEQ ID NO:8, 9, 10 or 11, and Glycine (G) at position X_{b8} of SEQ ID NO:38.

[0144] In some embodiments, the amino acid at position X_8 is substituted with an amino acid as set forth in SEQ ID NO:8, 9, 10 or 11, and the amino acid at position X_3 (of SEQ ID NO:29) is substituted with an amino acid as set forth in SEQ ID NO:29. Thus, in some embodiments, the amino acid at position X_8 is any amino acid other than Ile (I) and the amino acid at position X_3 is any amino acid other than Ile (I), Leu (L) or Met (M). In some embodiments, amino acid substitutions include Lysine (K), Arginine (R), Serine (S), Glycine (G) or Alanine (A) at position X_3 of SEQ ID NO:29. In certain embodiments, amino acid substitutions independently include Methionine (M) at position X_8 of SEQ ID NO:8, 9, 10 or 11, and Lysine (K) at position X_3 of SEQ ID NO:29.

[0145] Other suitable amino acid substitution(s) at one or more of the identified sites can be determined using, e.g., known methods of site-directed mutagenesis and determination of polynucleotide extension performance in assays described further herein or otherwise known to persons of skill in the art, e.g., amino acid substitutions described in U.S. Pat. Application Publication Nos. 2009/0148891 and 2009/0280539, which are incorporated by reference herein in its entirety.

[0146] Because the precise length of DNA polymerases vary, the precise amino acid positions corresponding to each of X_8 (SEQ ID NO:8), X_{b8} (SEQ ID NO:38) and X_3 (SEQ ID NO:29) can vary depending on the particular mutant polymerase used. Amino acid and nucleic acid sequence alignment programs are readily available (see, e.g., those referred to supra) and, given the particular motifs identified herein, serve to assist in the identification of the exact amino acids (and corresponding codons) for modification in accordance with the present invention. The positions corresponding to each of X_8, X_{b8} and X_3 are shown in Table 1 for representative chimeric thermostable DNA polymerases and thermostable DNA polymerases from exemplary thermophilic species.

TABLE 1

Amino Acid Positions Corresponding to Motif Positions X_8 (e.g. , of SEQ ID NOs: 8, 9, 10, and 11), X_{b8} (of SEQ ID NO: 38) and X_3 (of SEQ ID NO: 29) in Exemplary Polymerases.

Organism or Chimeric	Amino Acid Position		
Sequence Consensus (SEQ ID NO:)	X_8	X_{b8} (of SEQ ID NO: 38)	X_3 (of SEQ ID NO: 29)
T. thermophilus (6)	640	580	709
T. caldophilus (7)	640	580	709
T. sp. Z05 (1)	640	580	709
T. aquaticus (2)	638	578	707
T. flavus (4)	637	577	706
T. filiformis (3)	636	576	705
T. sp. sps17 (5)	636	576	705
T. maritima (34)	701	640	770
T. neapolitana (35)	701	640	770
T. africanus (33)	700	639	769
B. caldotenax (37)	682	621	751
B. stearothermophilus (36)	681	620	750
CS5 (27)	701	640	770
CS6 (28)	701	640	770

[0147] In some embodiments, the DNA polymerase of the present invention is derived from *Thermus* sp. Z05 DNA polymerase (SEQ ID NO:1) or a variant thereof (e.g., carrying the D580G mutation or the like). As referred to above, in *Thermus* sp. Z05 DNA polymerase, position X_8 corresponds to Isoleucine (I) at position 640; position X_{b8} corresponds to Aspartate (D) at position 580, and position X_3 corresponds to Isoleucine (I) at position 709. Thus, in certain variations of the invention, the mutant polymerase comprises at least one amino acid substitution, relative to a *Thermus* sp. Z05 DNA polymerase (or a DNA polymerase that is substantially identical (e.g., at least about 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% identical) to SEQ ID NO:1), at I640, D580 and/or I709. Thus, typically, the amino acid at position 640 of SEQ ID NO:1 is not I. In some embodiments, the amino acid at position 640 of SEQ ID NO:1 is selected from G, A, V, R,

F, W, P, S, T, C, Y, N, Q, D, E, K, L, M, or H. In certain embodiments, the amino acid residue at position 640 of SEQ ID NO:1 is F. In certain embodiments, amino acid residues at position D580 of SEQ ID NO:1 can be selected from Leucine (L), Glycine (G), Threonine (T), Glutamine (Q), Alanine (A), Serine (S), Asparagine (N), Arginine (R), and Lysine (K). Thus, in some embodiments, the amino acid residue at position 580 of SEQ ID NO:1 is Glycine (G). Further, in certain embodiments, the amino acid at position 709 of SEQ ID NO:1 is not I. In some embodiments, the amino acid at position 709 of SEQ ID NO:1 is selected from G, A, V, R, F, W, P, S, T, C, Y, N, Q, D, E, K, L, M, or H. In some embodiments, the amino acid at position 709 of SEQ ID NO: 1 is K, R, S, G or A. In some embodiments, the amino acid at position 709 of SEQ ID NO: 1 is K.

[0148] Exemplary *Thermus* sp. Z05 DNA polymerase mutants include those comprising the amino acid substitution(s) I640F, and/or I709K (or I709R, I709S, I709G, I709A), and/or D580G. In some embodiments, the mutant *Thermus* sp. Z05 DNA polymerase comprises, e.g., amino acid residue substitutions I640F and D580G. In some embodiments, the mutant *Thermus* sp. Z05 DNA polymerase comprises, e.g., amino acid residue substitutions I640 F and I709K. In some embodiments, the mutant *Thermus* sp. Z05 DNA polymerase comprises, e.g., amino acid residue substitutions I640F, I709K, and D580G. In certain embodiments, the mutant *Thermus* sp. Z05 DNA polymerase comprises, e.g., amino acid residue substitutions independently selected from I640F, I709K, and/or D580G.

[0149] In some embodiments, the amino acid corresponding to position 324 of SEQ ID NO:1 is Lys (K). In some embodiments, the amino acid corresponding to position 324 of SEQ ID NO:1 is not Met (M). In some embodiments, the amino acid corresponding to position 461 of SEQ ID NO:1 is Leu (L). In some embodiments, the amino acid corresponding to position 461 of SEQ ID NO:1 is not Met (M). In some embodiments, the amino acid corresponding to position 517 of SEQ ID NO:1 is Ser (S). In some embodiments, the amino acid corresponding to position 517 of SEQ ID NO:1 is not Arg (R). In some embodiments, the amino acid corresponding to position 741 of SEQ ID NO:1 is Ser (S). In some embodiments, the amino acid corresponding to position 741 of SEQ ID NO:1 is not Gly (G). In some embodiments, the amino acid corresponding to position 775 of SEQ ID NO:1 is Arg (R). In some embodiments, the amino acid corresponding to position 775 of SEQ ID NO:1 is not Gly (G). In some embodiments, the amino acid corresponding to position 791 of SEQ ID NO:1 is Leu (L). In some embodiments, the amino acid corresponding to position 789 of SEQ ID NO:1 is not Phe (F).

[0150] The inventors have shown that substitutions at the amino acid corresponding to position 709 of SEQ ID NO:1 described above can result in DNA polymerases having improved (i.e., increased) reverse transcription efficiency, increased RT-PCR activity (e.g., more efficient amplification of an RNA template without compromising PCR efficiency on a DNA template), increased RT-PCR efficiency in the presence of Mg^{2+}, increased reverse transcriptase activity in the presence of inhibitors (e.g., breakdown products of hemoglobin such as hemin, and/or heparin), increased extension rate and improved 3'-mismatch tolerance compared to a control polymerase. See U.S. Patent Application No. 61/474,160, filed Apr. 11, 2011, the contents of which are incorporated by reference herein in its entirety. Thus, it is expected that the improved polymerases that comprise substitutions at the amino acid corresponding to position 709 of SEQ ID NO:1 described herein will also have the improved properties described above.

[0151] In addition to the mutations and substitutions described herein, the DNA polymerases of the present invention can also include other, non-substitutional modification(s). Such modifications can include, for example, covalent modifications known in the art to confer an additional advantage in applications comprising polynucleotide extension. For example, one such modification is a thermally reversible covalent modification that inactivates the enzyme, but which is reversed to activate the enzyme upon incubation at an elevated temperature, such as a temperature typically used for polynucleotide extension. Exemplary reagents for such thermally reversible modifications are described in U.S. Pat. Nos. 5,773,258 and 5,677,152 to Birch et al., which are expressly incorporated by reference herein in their entirety.

[0152] The DNA polymerases of the present invention can be constructed by mutating the DNA sequences that encode the corresponding unmodified polymerase (e.g., a wild-type polymerase or a corresponding variant from which the polymerase of the invention is derived), such as by using techniques commonly referred to as site-directed mutagenesis. Nucleic acid molecules encoding the unmodified form of the polymerase can be mutated by a variety of polymerase chain reaction (PCR) techniques well-known to one of ordinary skill in the art. (See, e.g., PCR Strategies (M. A. Innis, D. H. Gelfand, and J. J. Sninsky eds., 1995, Academic Press, San Diego, Calif.) at Chapter 14*; PCR Protocols: A Guide to Methods and Applications* (M. A. Innis, D. H. Gelfand, J. J. Sninsky, and T. J. White eds., Academic Press, NY, 1990).

[0153] By way of non-limiting example, the two primer system, utilized in the Transformer Site-Directed Mutagenesis kit from Clontech, may be employed for introducing site-directed mutants into a polynucleotide encoding an unmodified form of the polymerase. Following denaturation of the target plasmid in this system, two primers are simultaneously annealed to the plasmid; one of these primers contains the desired site-directed mutation, the other contains a mutation at another point in the plasmid resulting in elimination of a restriction site. Second strand synthesis is then carried out, tightly linking these two mutations, and the resulting plasmids are transformed into a mutS strain of *E. coli*. Plasmid DNA is isolated from the transformed bacteria, restricted with the relevant restriction enzyme (thereby linearizing the unmutated plasmids), and then retransformed into *E. coli*. This system allows for generation of mutations directly in an expression plasmid, without the necessity of subcloning or generation of single-stranded phagemids. The tight linkage of the two mutations and the subsequent linearization of unmutated plasmids result in high mutation efficiency and allow minimal screening. Following synthesis of the initial restriction site primer, this method requires the use of only one new primer type per mutation site. Rather than prepare each positional mutant separately, a set of "designed degenerate" oligonucleotide primers can be synthesized in order to introduce all of the desired mutations at a given site simultaneously. Transformants can be screened by sequencing the plasmid DNA through the mutagenized region to identify and sort mutant clones. Each mutant DNA can then be restricted and analyzed by electrophoresis, such as for example, on a Mutation Detection Enhancement gel (Mallinckrodt Baker, Inc., Phillipsburg, N.J.) to confirm that no other alterations in the sequence have occurred (by band shift comparison to the unmutagenized control). Alternatively, the entire DNA region can be sequenced to confirm that no additional mutational events have occurred outside of the targeted region.

[0154] DNA polymerases with more than one amino acid substituted can be generated in various ways. In the case of amino acids located close together in the polypeptide chain, they may be mutated simultaneously using one oligonucleotide that codes for all of the desired amino acid substitutions. If however, the amino acids are located some distance from each other (separated by more than ten amino acids, for example) it is more difficult to generate a single oligonucleotide that encodes all of the desired changes. Instead, one of two alternative methods may be employed. In the first method, a separate oligonucleotide is generated for each amino acid to be substituted. The oligonucleotides are then annealed to the single-stranded template DNA simultaneously, and the second strand of DNA that is synthesized from the template will encode all of the desired amino acid substitutions. An alternative method involves two or more rounds of mutagenesis to produce the desired mutant. The first round is as described for the single mutants: DNA encoding the unmodified polymerase is used for the template, an oligonucleotide encoding the first desired amino acid substitution(s) is annealed to this template, and the heteroduplex DNA molecule is then generated. The second round of mutagenesis utilizes the mutated DNA produced in the first round of mutagenesis as the template. Thus, this template already contains one or more mutations. The oligonucleotide encoding the additional desired amino acid substitution(s) is then annealed to this template, and the resulting strand of DNA now encodes mutations from both the first and second rounds of mutagenesis. This resultant DNA can be used as a template in a third round of mutagenesis, and so on. Alternatively, the multi-site mutagenesis method of Seyfang & Jin (*Anal. Biochem*. 324:285-291. 2004) may be utilized.

[0155] Accordingly, also provided are recombinant nucleic acids encoding any of the DNA polymerases of the present invention. Using a nucleic acid of the present invention, encoding a DNA polymerase, a variety of vectors can be made. Any vector containing replicon and control sequences that are derived from a species compatible with the host cell can be used in the practice of the invention. Generally, expression vectors include transcriptional and translational regulatory nucleic acid regions operably linked to the nucleic acid encoding the DNA polymerase. The term "control sequences" refers to DNA sequences necessary for the expression of an operably linked coding sequence in a particular host organism. The control sequences that are suitable for prokaryotes, for example, include a promoter, optionally an operator sequence, and a ribosome binding site. In addition, the vector may contain a Positive Retroregulatory Element (PRE) to enhance the half-life of the transcribed mRNA (see Gelfand et al. U.S. Pat. No. 4,666,848). The transcriptional and translational regulatory nucleic acid regions will generally be appropriate to the host cell used to express the polymerase. Numerous types of appropriate expression vectors, and suitable regulatory sequences are known in the art for a variety of host cells. In general, the transcriptional and translational regulatory sequences may include, e.g., promoter sequences, ribosomal binding sites, transcriptional start and stop sequences, translational start and stop sequences, and enhancer or activator sequences. In typical embodiments, the regulatory sequences include a promoter and transcriptional start and stop sequences. Vectors also typically include a polylinker region containing several restriction sites for insertion of foreign DNA. In certain embodiments, "fusion flags" are used to facilitate purification and, if desired, subsequent removal of tag/flag sequence, e.g., "His-Tag". However, these are generally unnecessary when purifying a thermoactive and/or thermostable protein from a mesophilic host (e.g., *E. coli*) where a "heat-step" may be employed. The construction of suitable vectors containing DNA encoding replication sequences, regulatory sequences, phenotypic selection genes, and the polymerase of interest are prepared using standard recombinant DNA procedures. Isolated plasmids, viral vectors, and DNA fragments are cleaved, tailored, and ligated together in a specific order to generate the desired vectors, as is well-known in the art (see, e.g., Sambrook et al., *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory Press, New York, N.Y., 2nd ed. 1989)).

[0156] In certain embodiments, the expression vector contains a selectable marker gene to allow the selection of transformed host cells. Selection genes are well known in the art and will vary with the host cell used. Suitable selection genes can include, for example, genes coding for ampicillin and/or tetracycline resistance, which enables cells transformed with these vectors to grow in the presence of these antibiotics.

[0157] In one aspect of the present invention, a nucleic acid encoding a DNA polymerase is introduced into a cell, either alone or in combination with a vector. By "introduced into" or grammatical equivalents herein is meant that the nucleic acids enter the cells in a manner suitable for subsequent integration, amplification, and/or expression of the nucleic acid. The method of introduction is largely dictated by the targeted cell type. Exemplary methods include $CaPO_4$ precipitation, liposome fusion, LIPOFECTIN®, electroporation, viral infection, and the like.

[0158] In some embodiments, prokaryotes are typically used as host cells for the initial cloning steps of the present invention. They are particularly useful for rapid production of large amounts of DNA, for production of single-stranded DNA templates used for site-directed mutagenesis, for screening many mutants simultaneously, and for DNA sequencing of the mutants generated. Suitable prokaryotic host cells include *E. coli* K12 strain 94 (ATCC No. 31,446), *E. coli* strain W3110 (ATCC No. 27,325), *E. coli* K12 strain DG116 (ATCC No. 53,606), *E. coli* X1776 (ATCC No. 31,537), and *E. coli* B; however many other strains of *E. coli*, such as HB101, JM101, NM522, NM538, NM539, and many other species and genera of prokaryotes including bacilli such as *Bacillus subtilis*, other enterobacteriaceae such as *Salmonella typhimurium* or *Serratia marcesans*, and various *Pseudomonas* species can all be used as hosts. Prokaryotic host cells or other host cells with rigid cell walls are typically transformed using the calcium chloride method as described in section 1.82 of Sambrook et al., supra. Alternatively, electroporation can be used for transformation of these cells. Prokaryote transformation techniques are set forth in, for example Dower, in *Genetic Engineering, Principles and Methods* 12:275-296 (Plenum Publishing Corp., 1990); Hanahan et al., *Meth. Enzymol.*, 204:63, 1991. Plasmids typically used for transformation of *E. coli* include pBR322, pUCI8, pUCI9, pUCI18, pUC119, and Bluescript M13, all of which are described in sections 1.12-1.20 of Sambrook et al., supra. However, many other suitable vectors are available as well.

[0159] The DNA polymerases of the present invention are typically produced by culturing a host cell transformed with an expression vector containing a nucleic acid encoding the DNA polymerase, under the appropriate conditions to induce or cause expression of the DNA polymerase. Methods of culturing transformed host cells under conditions suitable for protein expression are well-known in the art (see, e.g., Sambrook et al., supra). Suitable host cells for production of the polymerases from lambda pL promotor-containing plasmid vectors include *E. coli* strain DG116 (ATCC No. 53606) (see U.S. Pat. No. 5,079,352 and Lawyer, F. C. et al., *PCR Methods and Applications* 2:275-87, 1993, which are both incorporated herein by reference). Following expression, the polymerase can be harvested and isolated. Methods for purifying the thermostable DNA polymerase are described in, for example, Lawyer et al., supra. Once purified, the ability of the DNA polymerases to have improved RT efficiency, increased mis-match tolerance, extension rate and/or tolerance of RT and polymerase inhibitors can be tested (e.g., as described in the examples).

[0160] The improved DNA polymerases of the present invention may be used for any purpose in which such enzyme activity is necessary or desired. Accordingly, in another aspect of the invention, methods of polynucleotide extension (e.g., PCR) using the polymerases are provided. Conditions suitable for polynucleotide extension are known in the art. (See, e.g., Sambrook et al., supra. See also Ausubel et al., *Short Protocols in Molecular Biology* (4th ed., John Wiley & Sons 1999). Generally, a primer is annealed, i.e., hybridized, to a target nucleic acid to form a primer-template complex. The primer-template complex is contacted with the DNA polymerase and nucleoside triphosphates in a suitable environment to permit the addition of one or more nucleotides to the 3' end of the primer, thereby producing an extended primer complementary to the target nucleic acid. The primer can include, e.g., one or more nucleotide analog(s). In addition, the nucleoside triphosphates can be conventional nucleotides, unconventional nucleotides (e.g., ribonucleotides or labeled nucleotides), or a mixture thereof. In some variations, the polynucleotide extension reaction comprises amplification of a target nucleic acid. Conditions suitable for nucleic acid amplification using a DNA polymerase and a primer pair are also known in the art (e.g., PCR amplification methods). (See, e.g., Sambrook et al., supra; Ausubel et al., supra; *PCR Applications: Protocols for Functional Genomics* (Innis et al. eds., Academic Press 1999). In other, non-mutually exclusive embodiments, the polynucleotide extension reaction comprises reverse transcription of an RNA template (e.g., RT-PCR). In some embodiments, the improved polymerases find use in 454 sequencing (Margulies, M et al. 2005, Nature, 437, 376-380).

[0161] Optionally, the primer extension reaction comprises an actual or potential inhibitor of a reference or unmodified polymerase. The inhibitor can inhibit, for example, the nucleic acid extension rate and/or the reverse transcription efficiency of a reference or unmodified (control) polymerase. In some embodiments, the inhibitor is hemoglobin, or a degradation product thereof. For example, in some embodiments, the hemoglobin degradation product is a heme breakdown product, such as hemin, hematoporphyrin, or bilirubin. In some embodiments, the inhibitor is an iron-chelator or a purple pigment. In other embodiments, the inhibitor is heparin. In certain embodiments, the inhibitor is an intercalating dye. In certain embodiments, the inhibitor is melanin, which has been described as a polymerase inhibitor. See, e.g, Ekhardt, et al., *Biochem Biophys Res Commun.* 271(3):726-30 (2000).

[0162] The DNA polymerases of the present invention can be used to extend templates in the presence of polynucleotide templates isolated from samples comprising polymerase inhibitors, e.g., such as blood. For example, the DNA polymerases of the present invention can be used to extend templates in the presence of hemoglobin, a major component of blood, or in the presence of a hemoglobin degradation product. Hemoglobin can be degraded to various heme breakdown products, such as hemin, hematin, hematoporphyrin, and bilirubin. Thus, in certain embodiments, the DNA polymerases of the present invention can be used to extend templates in the presence of hemoglobin degradation products, including but not limited to, hemin, hematin, hematoporphyrin, and bilirubin. In certain embodiments, the hemoglobin degradation product is hemin. In some embodiments, the DNA polymerases of the present invention can be used to extend templates in the presence of about 0.5 to 20.0 μM, about 0.5 to 10.0 μM, about 0.5 to 5.0 μM, about 1.0 to 10.0 μM, about 1.0 to 5.0 μM, about 2.0 to 5.0 μM, or about 2.0 to 3.0 μM hemin. In other embodiments, the DNA polymerases of the present invention can be used to extend templates in the presence of at least about 0.5, 1.0, 1.5, 2.0, 2.5, 3.0, 4.0, 5.0, 10.0, 20.0, or greater than 20 μM hemin. The breakdown products of hemoglobin include iron-chelators and purple pigments. Thus, in some embodiments, the DNA polymerases of the present invention can be used to extend templates in the presence of iron-chelators and/or purple pigments. In other embodiments, the DNA polymerases of the present invention can be used to extend templates in the presence of amounts of hemoglobin degradation products that would inhibit extension of the same template by a reference or control DNA polymerase.

[0163] The DNA polymerases of the present invention can be used to extend templates in the presence of heparin. Heparin is commonly present as an anticoagulant in samples isolated from blood. In some embodiments, the DNA polymerases of the present invention can be used to extend templates in the presence of about 1.0 to 400 ng/μl, 1.0 to 300 ng/μl, 1.0 to 200 ng/μl, 5.0 to 400 ng/μl, 5.0 to 300 ng/μl, 5.0 to 200 ng/μl, 10.0 to 400 ng/μl, 10.0 to 300 ng/μl, or 10.0 to 200 ng/μl heparin. In some embodiments, the DNA polymerases of the present invention can be used to extend templates in the presence of at least about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 20, 30, 40, 50, 100, 150, 200, 250, 300, 350, 400 ng/μl, or greater than 400 ng/μl of heparin. In other embodiments, the DNA polymerases of the present invention can be used to extend templates in the presence of amounts of heparin that would inhibit extension of the same template by a reference or control DNA polymerase.

[0164] In some embodiments, an improved polymerase of the invention is used in a reverse transcription reaction. In some embodiments, the reverse transcription reaction is carried out in a mixture containing the RNA template, one or more primer(s), and a thermostable DNA polymerase of the invention. The reaction mixture typically contains all four standard deoxyribonucleoside triphosphates (dNTPs) and a buffer containing a divalent cation and a monovalent cation. Exemplary cations include, e.g., Mg^{2+}, although other cations, such as Mn^{2+} or Co^{2+} can activate DNA polymerases. In other embodiments, the reverse transcription reaction is carried out with a thermo-active DNA polymerase of the invention. In particular embodiments, the improved polymerase of the invention allows for more efficient amplification of RNA templates without compromising the efficient amplification of a DNA template in the presence of Mn^{2+} or Mg^{2+}, as described in the examples.

[0165] In some embodiments, an improved polymerase of the invention increases reverse transcription efficiency by reducing the reaction time required for extending an RNA template. For example, an improved polymerase described herein can significantly reduce the reaction time required to transcribe RNA to cDNA as compared to a control polymerase, thereby increasing the reverse transcriptase efficiency. Without being limited by theory, the improved polymerase can increase RT efficiency by, for example, increasing the activity of the enzyme on an RNA template, such as increasing the rate of nucleotide incorporation and/or increasing the processivity of the polymerase, thereby effectively shortening the extension time of an RNA template or population of RNA templates. Reaction times for the initial RT step are typically on the order of 30 minutes or longer at 65 degrees C. when using an unmodified or control polymerase. Thus, in some embodiments, the improved polymerase can transcribe an RNA template into cDNA in less than about 10 minutes, less than about 8 minutes, less than about 5 minutes, less than about 4 minutes, less than about 3 minutes, or less than about 2 minutes at 65 degrees C. In some embodiments, the improved polymerase can transcribe an RNA template derived from Hepatitis C Virus (HCV) transcript JP2-5, containing the first 800 bases of HCV genotype Ib 5'NTR, into cDNA in less time or faster than a control polymerase. For example, the improved polymerase can transcribe 240 bases of the HCV JP2-5 RNA template into full-length cDNA in about 15 seconds less, 30 seconds less, one minute less, two minutes less, 3 minutes less, 4 minutes less, 5 minutes less, or about 10 minutes less than a control polymerase under identical reaction conditions. In some embodiments, the improved polymerase can transcribe 240 bases of the HCV JP2-5 RNA template into full-length cDNA faster than a control polymerase, for example, about 5 seconds, 10 seconds, 15 seconds, 30 seconds, 45 seconds, or 60 seconds or more faster than a control polymerase under identical reaction conditions. In some embodiments, the reaction conditions are those described in the Examples. In some embodiments, an improved polymerase described herein is contacted with an RNA template at 65 degrees C. for 2 minutes in the reaction mixture described above. The extension step can be followed by PCR amplification of the extended template, as described in the examples.

[0166] The most efficient RT activity in thermostable DNA polymerases has been achieved using Mn^{2+} as the divalent metal ion activator. However, it is well known that when Mn^{2+} is present in reactions the fidelity of DNA polymerases is lower. Unless one is trying to generate mutations, it is generally favored to maintain a higher fidelity. Fortunately, most conventional sequencing, PCR and RT-PCR applications do not require high fidelity conditions because the detection systems generally are looking at a population of products. With the advent of next generation sequencing, digital PCR, etc., the fidelity of the product is more important and methods that allow for higher fidelity DNA synthesis are critical. Achieving efficient RT activity using Mg^{2+} as the divalent metal ion activator is an excellent way to substantially increase the fidelity of the DNA polymerase and allow for more reliable copying of the nucleic acid target. Accordingly, in some embodiments, the improved polymerase of the invention allows for efficient extension and/or amplification of RNA templates using Mg^{2+} as the divalent metal ion activator, as described in the examples.

[0167] Because the polymerases described herein can also have increased mismatch tolerance, the polymerases find use in methods where variation of the target template is likely and yet the template is nevertheless desired to be amplified regardless of the variation at the target template. An example of such templates can include, for example, viral, bacterial, or other pathogen sequences. In many embodiments, it is desirable to determine simply whether an individual (human or non-human animal) has a viral or other infection, regardless of the precise viral variant that has infected the individual. As an example, one can use a primer pair to amplify HCV using a polymerase of the invention and detect the presence of the HCV even if the particular virus infecting the individual has a mutation resulting in a mismatch at the primer hybridization site.

[0168] Target nucleic acids can come from a biological or synthetic source. The target can be, for example, DNA or RNA. Generally, where amplicons are generated, the amplicons will be composed of DNA, though ribonucleotides or synthetic nucleotides can also be incorporated into the amplicon. Where one wishes to detect an RNA, the amplification process will typically involve the use of reverse transcription, including for example, reverse transcription PCR (RT-PCR).

[0169] Specific target sequences can include, e.g., viral nucleic acids (e.g., human immunodeficiency virus (HIV), hepatitis virus B (HBV), (cytomegalovirus (CMV), parvo B19 virus, Epstein-Barr virus, hepatitis virus C (HCV), human papilloma virus (HPV), Japanese encephalitis virus (JEV), West Nile virus (WNV), St. Louis encephalitis virus (SLEV), Murray Valley encephalitis virus, and Kunjin virus), bacterial nucleic acids (e.g., *S. aureus, Neisseria meningitidis, Plasmodium falciparum, Chlamydia muridarum, Chlamydia trachomatis*), mycobacteria, fungal nucleic acids, or nucleic acids from animals or plants. In some embodiments, the target nucleic acids are animal (e.g., human) nucleic acids or are derived from an animal (e.g., human) sample (i.e., viral or other pathogenic organism nucleic acids may be present in a sample from an animal biopsy, blood sample, urine sample, fecal sample, saliva, etc.). In some embodiments, the target nucleic acids are, for example, human genetic regions that may include variants associated with disease (e.g., cancer, diabetes, etc.). Because in some embodiments the polymerases of the invention have mismatch tolerance, such enzymes are particularly useful, for example, where a diversity of related sequences could be in a target sequence. As an example, the invention can be used to detect viral pathogens, where the viral pathogens have sufficient variation in their genomes to make it difficult or impossible to design a single or small set of primers that will amplify most or all possible viral genomes or in cancer or other disease genetic markers where variation in sequence is known or likely to occur.

[0170] Other methods for detecting extension products or amplification products using the improved polymerases described herein include the use of fluorescent double-stranded nucleotide binding dyes or fluorescent double-stranded nucleotide intercalating dyes. Examples of fluorescent double-stranded DNA binding dyes include SYBR-green (Molecular Probes). The double stranded DNA binding dyes can be used in conjunction with melting curve analysis to measure primer extension products and/or amplification products. The melting curve analysis can be performed on a real-time PCR instrument, such as the ABI 5700/7000 (96 well format) or ABI 7900 (384 well format) instrument with onboard software (SDS 2.1). Alternatively, the melting curve analysis can be performed as an end point analysis. Exemplary methods of melting point analysis are described in U.S. Patent Publication No. 2006/0172324, the contents of which are expressly incorporated by reference herein in its entirety.

[0171] In another aspect of the present invention, kits are provided for use in primer extension methods described herein. In some embodiments, the kit is compartmentalized for ease of use and contains at least one container providing an improved DNA polymerase in accordance with the present invention. One or more additional containers providing additional reagent(s) can also be included. In some embodiments, the kit can also include a blood collection tube, container, or unit that comprises heparin or a salt thereof, or releases heparin into solution. The blood collection unit can be a heparinized tube. Such additional containers can include any reagents or other elements recognized by the skilled artisan for use in primer extension procedures in accordance with the methods described above, including reagents for use in, e.g., nucleic acid amplification procedures (e.g., PCR, RT-PCR), DNA sequencing procedures, or DNA labeling procedures. For example, in certain embodiments, the kit further includes a container providing a 5' sense primer hybridizable, under primer extension conditions, to a predetermined polynucleotide template, or a primer pair comprising the 5' sense primer and a corresponding 3' antisense primer. In other, non-mutually exclusive variations, the kit includes one or more containers providing nucleoside triphosphates (conventional and/or unconventional). In specific embodiments, the kit includes alpha-phosphorothioate dNTPs, dUTP, dITP, and/or labeled dNTPs such as, e.g., fluorescein- or cyanin-dye family dNTPs. In still other, non-mutually exclusive embodiments, the kit includes one or more containers providing a buffer suitable for a primer extension reaction.

[0172] In another aspect of the present invention, reaction mixtures are provided comprising the polymerases with increased reverse transcriptase efficiency, mismatch tolerance, extension rate and/or tolerance of RT and polymerase inhibitors as described herein. The reaction mixtures can further comprise reagents for use in, e.g., nucleic acid amplification procedures (e.g., PCR, RT-PCR), DNA sequencing procedures, or DNA labeling procedures. For example, in certain embodiments, the reaction mixtures comprise a buffer suitable for a primer extension reaction. The reaction mixtures can also contain a template nucleic acid (DNA and/or RNA), one or more primer or probe polynucleotides, nucleoside triphosphates (including, e.g., deoxyribonucleotides, ribonucleotides, labeled nucleotides, unconventional nucleotides), salts (e.g., Mn^{2+}, Mg^{2+}), labels (e.g., fluorophores). In some embodiments, the reaction mixtures contain a 5'-sense primer hybridizable, under primer extension conditions, to a predetermined polynucleotide template, or a primer pair comprising the 5'-sense primer and a corresponding 3' antisense primer. In some embodiments, the reaction mixtures contain alpha-phosphorothioate dNTPs, dUTP, dITP, and/or labeled dNTPs such as, e.g., fluorescein- or cyanin-dye family dNTPs. In some embodiments, the reaction mixtures comprise an iron chelator or a purple dye. In certain embodiments, the reaction mixtures comprise hemoglobin, or a degradation product of hemoglobin. For example, in certain embodiments, the degradation products of hemoglobin include heme breakdown products such as hemin, hematin, hematophoryn, and bilirubin. In other embodiments, the reaction mixtures comprise heparin or a salt thereof. In certain embodiments, the reaction mixture contains a template nucleic acid that is isolated from blood. In other embodiments, the template nucleic acid is RNA and the reaction mixture comprises heparin or a salt thereof.

[0173] In some embodiments, the reaction mixture comprises two or more polymerases. For example, in some embodiments, the reaction mixture comprises a first DNA polymerase having increased reverse transcriptase efficiency compared to a control polymerase, and a second DNA polymerase having DNA-dependent polymerase activity. The second DNA polymerase can be a wild-type or unmodified polymerase, or can be an improved polymerase having increased DNA-dependent polymerase activity. Such reaction mixtures are useful for amplification of RNA templates (e.g., RT-PCR) by providing both a polymerase having increased reverse transcriptase activity and a polymer are having DNA-dependent polymerase activity.

EXAMPLES

[0174] The following examples are offered to illustrate, but not to limit the claimed invention.

Example 1

Library Generation

[0175] In brief, the steps in this screening process included library generation, expression and partial purification of the mutant enzymes, screening of the enzymes for the desired properties, DNA sequencing, clonal purification, and further characterization of selected candidate mutants. Each of these steps is described further below.

[0176] Clonal Library Generation:

[0177] A nucleic acid encoding the polymerase domain of Z05 D580G_I709K DNA polymerase was subjected to error-prone (mutagenic) PCR between Blp I and Bgl II restriction sites of a plasmid including this nucleic acid sequence. The primers used for this are given below:

```
Forward Primer:
                                        (SEQ ID NO: 30)
5'-CTACCTCCTGGACCCCTCCAA-3';
and,

Reverse Primer:
                                        (SEQ ID NO: 31)
5'-ATAACCAACTGGTAGTGGCGTGTAA-3'
```

PCR was performed using a Mg^{2+} concentration of 1.8 mM, in order to generate a library with a desired mutation rate. Buffer conditions were 50 mM Bicine pH 8.2, 115 mM KOAc, 8% w/v glycerol, and 0.2 mM each dNTPs. A GeneAmp® AccuRT Hot Start PCR enzyme was used at 0.15 U/μL. Starting with 5×10^5 copies of linearized Z05 D580G_I709K plasmid DNA per reaction volume of 50 μL, reactions were denatured using a temperature of 94° C. for 60 seconds, then 30 cycles of amplification were performed, using a denaturation temperature of 94° C. for 15 seconds, an annealing temperature of 60° C. for 15 seconds, an extension temperature of 72° C. for 120 seconds, and followed by a final extension at a temperature of 72° C. for 5 minutes.

[0178] The resulting amplicon was purified with a QIAquick PCR Purification Kit (Qiagen, Inc., Valencia, Calif., USA) and cut with Blp I and Bgl II, and then repurified with a QIAquick PCR Purification Kit. A Z05 D580G_I709K vector plasmid was prepared by cutting with the same two restriction enzymes and treating with alkaline phosphatase, recombinant (RAS, cat#03359123001) and purified with a QIAquick PCR Purification Kit. The cut vector and the mutated insert were mixed at a 1:3 ratio and treated with T4 DNA ligase for 5 minutes at room temperature (NEB Quick Ligation™ Kit). The ligations were purified with a QIAquick PCR Purification Kit and transformed into an *E.coli* host strain by electroporation.

[0179] Aliquots of the expressed cultures were plated on ampicillin-selective medium in order to determine the number of unique transformants in each transformation. Transformations were pooled and stored at −70° C. to −80° C. in the presence of glycerol as a cryo-protectant.

[0180] The library was then spread on large format ampicillin-selective agar plates. Individual colonies were transferred to 384-well plates containing 2× Luria broth with ampicillin and 10% w/v glycerol using an automated colony picker (QPix2, Genetix Ltd). These plates were incubated overnight at 30° C. to allow the cultures to grow and then stored at −70° C. to −80° C. The glycerol added to the 2× Luria broth was low enough to permit culture growth and yet high enough to provide cryo-protection. Several thousand colonies were prepared in this way for later use.

[0181] Extract Library Preparation Part 1—Fermentation:

[0182] From the clonal libraries described above, a corresponding library of partially purified extracts suitable for screening purposes was prepared. The first step of this process was to make small-scale expression cultures of each clone. These cultures were grown in 96-well format; therefore there were 4 expression culture plates for each 384-well library plate. 0.5 µL was transferred from each well of the clonal library plate to a well of a 96 well seed plate, containing 150 µL of Medium A (see Table 3 below). This seed plate was shaken overnight at 1150 rpm at 30° C., in an iEMS plate incubator/shaker (ThermoElectron). These seed cultures were then used to inoculate the same medium, this time inoculating 20 µL into 250 µL Medium A in large format 96 well plates (Nunc #267334). These plates were incubated overnight at 37° C. with shaking. The expression plasmid contained transcriptional control elements, which allow for expression at 37° C. but not at 30° C. After overnight incubation, the cultures expressed the clone protein at typically 1-10% of total cell protein. The cells from these cultures were harvested by centrifugation. These cells were either frozen (−20° C.) or processed immediately, as described below.

TABLE 2

Medium A (Filter-sterilized prior to use)

Component	Concentration
$MgSO_4 \cdot 7H_2O$	0.2 g/L
Citric acid•H_2O	2 g/L
K_2HPO_4	10 g/L
$NaNH_4PO_4 \cdot 4H_2O$	3.5 g/L
$MgSO_4$	2 mM
Casamino acids	2.5 g/L
Glucose	2 g/L
Thiamine•HCl	10 mg/L
Ampicillin	100 mg/L

[0183] Extract Library Preparation Part 2—Extraction:

[0184] Cell pellets from the fermentation step were resuspended in 25 µL Lysis buffer (Table 3 below) and transferred to 384-well thermocycler plates and sealed. Note that the buffer contained lysozyme to assist in cell lysis, and DNase to remove DNA from the extract. To lyse the cells the plates were incubated at 37° C. for 15 minutes, frozen overnight at −20° C., and incubated again at 37° C. for 15 minutes. Ammonium sulfate was added (1.5 µL of a 2M solution) and the plates incubated at 75° C. for 15 minutes in order to precipitate and inactivate contaminating proteins, including the exogenously added nucleases. The plates were centrifuged at 3000×g for 15 minutes at 4° C. and the supernatants transferred to a fresh 384-well thermocycler plate. These extract plates were frozen at −20° C. for later use in screens. Each well contained about 0.5-3 µM of the mutant library polymerase enzyme.

TABLE 3

Lysis Buffer

Component	Concentration or Percentage
Tris pH 7.5	50 mM
EDTA	1 mM
$MgCl_2$	6 mM
Tween 20	0.5% v/v
Lysozyme (from powder)	1 mg/mL
DNase I	0.05 Units/µL

Example 2

Identification of Mutant DNA Polymerases with Improved Reverse Transcription Efficiency

[0185] Screening Crude Protein Extract Libraries for Improved Reverse Transcription Efficiency:

[0186] The extract library was screened by comparing Cp (Crossing Point) values from growth curves generated by fluorescent 5' nuclease (TaqMan) activity of crude enzyme extracts in a RT-PCR system from amplification of a 240 base pair amplicon from Hepatitis C Virus (HCV) transcript JP2-5, containing the first 800 bases of HCV genotype Ib 5'NTR in pSP64 poly(A) (Promega).

[0187] Reactions were carried out on the Roche LC 480 kinetic thermocycler in 384 well format with each well containing 3 µL of an individual enzyme extract diluted 10-fold with buffer containing 20 mM Tris-HCl, pH 8, 100 mM KCl, 0.1 mM EDTA, and 0.1% Tween-20 added to 12 µL of RT-PCR master mix described in Table 4. The thermocycling conditions were: 2 minute at 50° C. ("UNG" step); 2 minute at 65° C. ("RT" step); 5 cycles of 94° C. for 15 seconds followed by 62° C. for 30 seconds; and 45 cycles of 91° C. for 15 seconds followed by 62° C. for 30 seconds.

TABLE 4

Component	Concentration
Tricine pH 8.3	50 mM
KOAc	60 mM
Glycerol	5% (v/v)
DMSO	2% (v/v)
Primer 1	200 nM
Primer 2	200 nM
TaqMan Probe	100 nM
Aptamer	200 nM
dATP	200 μM
dCTP	200 μM
dGTP	200 μM
dUTP	400 μM
UNG	.2 Units/μL
RNA Target	6666 copies/μL
$Mg(OAc)_2$	2 mM

[0188] Approximately 5000 clones were screened using the above protocol. Forty clones were chosen from the original pool for rescreening based on earliest Crossing Point (Cp) values and fluorescent plateau values above an arbitrary cut off as calculated by the Abs Quant/2^{nd} derivative max method. Culture wells corresponding to the top extracts were sampled to fresh growth medium and re-grown to produce new culture plates containing the best mutants, as well as a number of parental Z05 D580G_I709K cultures to be used for comparison controls. These culture plates were then used to make fresh crude extracts which were rescreened with the same RNA target and conditions as previously described for the original screen. Table 5 shows average Cp values obtained from the fluorescent signal increase due to 5' hydrolysis of a FAM labeled probe. Results show that clone 0691-L24 amplifies the RNA target with higher efficiency than the Z05_D580G_I709K parental.

TABLE 5

Clone	Average Cp
0691-L24	19.1
Z05 D580G_I709K	28.0

[0189] The DNA sequence of the mutated region of the polymerase gene was sequenced to determine the mutation(s) that were present in any single clone. Clone 0691-L24 was chosen for further testing, so mutant polymerase protein was expressed in flask culture, purified to homogeneity, and quantified.

[0190] Use of Fully Purified 0691-L24 Mutant in Mg^{2+}-Based RT-PCR:

[0191] Purified and quantified mutant 0691-L24 was compared to parental Z05_D580G_I709K in TaqMan Mg^{2+}-based RT-PCR. Reverse transcription and PCR efficiencies were measured by comparing Cp values from amplifications of JP2-5 RNA transcript and pJP2-5 DNA linear plasmid digested with the restriction endonuclease EcoRI. Oligonucleotides and Master Mix conditions (Table 4) were the same as used in the original screen. Each reaction had either 100,000 copies of JP2-5 transcript RNA, 100,000 copies of pJP2-5 linear plasmid DNA, or 1000 copies of pJP2-5 linear plasmid DNA. All targets were amplified with Primer 1 and Primer 2, as described above, in duplicate reactions to generate a 240 base pair amplicon. All reactions were performed on the Roche Light Cycler 480 thermal cycler with a reaction volume of 15 μL. Crossing Point (Cps) were calculated by the Abs Quant/2^{nd} derivative max method and averaged. Master Mix conditions were the same as those described previously in Table 4 except reactions were carried out using a range of DNA Polymerase concentrations from 5 nM-40 nM. The thermocycling conditions were: 2 minute at 50° C. ("UNG" step); 2 minute at 65° C. ("RT" step); 5 cycles of 94° C. for 15 seconds followed by 62° C. for 30 seconds; and 45 cycles of 91° C. for 15 seconds followed by 62° C. for 30 seconds. Table 6 shows Cp values obtained from fluorescent signal increase due to cleavage of the TaqMan probe at 20 nM enzyme condition.

TABLE 6

Enzyme	RNA 10^5 copies Cp	DNA 10^5 copies Cp	DNA 10^3 copies Cp
Z05 D580G_I709K	28.8	17.5	24.4
0691-L24	18.9	17.3	24.0

[0192] The results indicate that mutant DNA polymerase 0691-L24 allows for more efficient amplification of RNA target without compromise of PCR efficiency on DNA target, as compared to parental Z05 D580G_I709K.

[0193] Determination of Phenotype-Conferring Mutation (s):

[0194] Sequencing results revealed that the polymerase expressed by clone 0691-L24 carries mutations N629D and I640F in addition to the parental D580G and I709K mutations. A Z05 D580G_I709K_I640F mutant was constructed by subcloning, purified, quantified, and compared to 0691-L24 (Z05 D580G_I709K_N629D_I640F) and parental Z05 D580G_I709K in Mg^{2+} activated TaqMan RT-PCR with varying KOAc concentration from 40 mM-110 mM and 25 nM purified enzyme. Master Mix conditions were the same as those described previously in Table 4. Each reaction had either 100,000 copies of JP2-5 RNA transcript, 100,000 copies of pJP2-5 linear plasmid DNA, or 1000 copies of pJP2-5 linear plasmid DNA. All targets were amplified with the same primer set in duplicate reactions to generate a 240 base pair amplicon. The PCR and RT-PCR efficiencies were determined by comparing Cp values between DNA and RNA. All reactions were performed on the Roche Light Cycler 480 thermal cycler with a reaction volume of 15 μL. Crossing Point (Cps) were calculated by the Abs Quant/2^{nd} derivative max method and Cps were averaged. The thermocycling conditions were: 2 minutes at 50° C. ("UNG" step); 65° C. for 2 minutes ("RT"step); 5 cycles of 94° C. for 15 seconds followed by 62° C. for 30 seconds; and 45 cycles of 91° C. for 15 seconds followed by 62° C. for 30 seconds. Table 8 shows the Cp values obtained from fluorescent signal increase due to cleavage of the TaqMan probe at the 60 mM KOAc condition.

TABLE 7

Enzyme	RNA 10^5 copies Cp	DNA 10^5 copies Cp	DNA 10^3 copies Cp
Z05 D580G_I709K	28.1	17.2	24.1
0691-L24	18.8	17.2	24.3
Z05 D580G_I709K_I640F	19.6	17.0	24.2

[0195] 0691-L24 (Z05 D580G_I709K_N629D_I640F) and Z05 D580G_I709K_I640F have similar Cp values on both RNA and DNA targets, demonstrating that the I640F mutation confers the observed improvement in RT-PCR performance.

[0196] It is understood that the examples and embodiments described herein are for illustrative purposes only and that various modifications or changes in light thereof will be suggested to persons skilled in the art and are to be included within the spirit and purview of this application and scope of the appended claims. All publications, sequence accession numbers, patents, and patent applications cited herein are hereby incorporated by reference in their entirety for all purposes.

```
                              SEQUENCE LISTING


<160> NUMBER OF SEQ ID NOS: 39

<210> SEQ ID NO 1
<211> LENGTH: 834
<212> TYPE: PRT
<213> ORGANISM: Thermus sp.
<220> FEATURE:
<223> OTHER INFORMATION: Thermus sp. Z05 DNA polymerase (Z05)

<400> SEQUENCE: 1

Met Lys Ala 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 Phe 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 Tyr Lys Ala Val Phe
    50                  55                  60

Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Glu
65                  70                  75                  80

Ala Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln
                85                  90                  95

Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Phe Thr Arg Leu
            100                 105                 110

Glu Val Pro Gly Phe Glu Ala Asp Asp Val Leu Ala Thr Leu Ala Lys
        115                 120                 125

Lys Ala Glu Arg Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Arg
    130                 135                 140

Asp Leu Tyr Gln Leu Val Ser Asp Arg Val Ala Val Leu His Pro Glu
145                 150                 155                 160

Gly His Leu Ile Thr Pro Glu Trp Leu Trp Glu Lys Tyr Gly Leu Lys
                165                 170                 175

Pro Glu Gln Trp Val Asp Phe Arg Ala Leu Val Gly Asp Pro Ser Asp
            180                 185                 190

Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Leu Lys Leu
        195                 200                 205

Leu Lys Glu Trp Gly Ser Leu Glu Asn Ile Leu Lys Asn Leu Asp Arg
    210                 215                 220

Val Lys Pro Glu Ser Val Arg Glu Arg Ile Lys Ala His Leu Glu Asp
225                 230                 235                 240

Leu Lys Leu Ser Leu Glu Leu Ser Arg Val Arg Ser Asp Leu Pro Leu
                245                 250                 255

Glu Val Asp Phe Ala Arg Arg Arg Glu Pro Asp Arg Glu Gly Leu Arg
            260                 265                 270

Ala Phe Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly
        275                 280                 285

Leu Leu Glu Ala Pro Ala Pro Leu Glu Glu Ala Pro Trp Pro Pro Pro
    290                 295                 300
```

-continued

```
Glu Gly Ala Phe Val Gly Phe Val Leu Ser Arg Pro Glu Pro Met Trp
305                 310                 315                 320

Ala Glu Leu Lys Ala Leu Ala Ala Cys Lys Glu Gly Arg Val His Arg
                325                 330                 335

Ala Lys Asp Pro Leu Ala Gly Leu Lys Asp Leu Lys Glu Val Arg Gly
            340                 345                 350

Leu Leu Ala Lys Asp Leu Ala Val Leu Ala Leu Arg Glu Gly Leu Asp
        355                 360                 365

Leu Ala Pro Ser Asp Asp Pro Met Leu Leu Ala Tyr Leu Leu Asp Pro
    370                 375                 380

Ser Asn Thr Thr Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly Glu Trp
385                 390                 395                 400

Thr Glu Asp Ala Ala His Arg Ala Leu Leu Ala Glu Arg Leu Gln Gln
                405                 410                 415

Asn Leu Leu Glu Arg Leu Lys Gly Glu Glu Lys Leu Leu Trp Leu Tyr
            420                 425                 430

Gln Glu Val Glu Lys Pro Leu Ser Arg Val Leu Ala His Met Glu Ala
        435                 440                 445

Thr Gly Val Arg Leu Asp Val Ala Tyr Leu Lys Ala Leu Ser Leu Glu
    450                 455                 460

Leu Ala Glu Glu Ile Arg Arg Leu Glu Glu Glu Val Phe Arg Leu Ala
465                 470                 475                 480

Gly His Pro Phe Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg Val Leu
                485                 490                 495

Phe Asp Glu Leu Arg Leu Pro Ala Leu Gly Lys Thr Gln Lys Thr Gly
            500                 505                 510

Lys Arg Ser Thr Ser Ala Ala Val Leu Glu Ala Leu Arg Glu Ala His
        515                 520                 525

Pro Ile Val Glu Lys Ile Leu Gln His Arg Glu Leu Thr Lys Leu Lys
    530                 535                 540

Asn Thr Tyr Val Asp Pro Leu Pro Gly Leu Val His Pro Arg Thr Gly
545                 550                 555                 560

Arg Leu His Thr Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly Arg Leu
                565                 570                 575

Ser Ser Ser Asp Pro Asn Leu Gln Asn Ile Pro Ile Arg Thr Pro Leu
            580                 585                 590

Gly Gln Arg Ile Arg Arg Ala Phe Val Ala Glu Ala Gly Trp Ala Leu
        595                 600                 605

Val Ala Leu Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala His Leu
    610                 615                 620

Ser Gly Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly Lys Asp Ile
625                 630                 635                 640

His Thr Gln Thr Ala Ser Trp Met Phe Gly Val Ser Pro Glu Ala Val
                645                 650                 655

Asp Pro Leu Met Arg Arg Ala Ala Lys Thr Val Asn Phe Gly Val Leu
            660                 665                 670

Tyr Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile Pro Tyr
        675                 680                 685

Glu Glu Ala Val Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys
    690                 695                 700
```

-continued

```
Val Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Lys Arg Gly
705                 710                 715                 720

Tyr Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Asn
                725                 730                 735

Ala Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn
            740                 745                 750

Met Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val
        755                 760                 765

Lys Leu Phe Pro His Leu Arg Glu Met Gly Ala Arg Met Leu Leu Gln
    770                 775                 780

Val His Asp Glu Leu Leu Leu Glu Ala Pro Gln Ala Arg Ala Glu Glu
785                 790                 795                 800

Val Ala Ala Leu Ala Lys Glu Ala Met Glu Lys Ala Tyr Pro Leu Ala
                805                 810                 815

Val Pro Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala
            820                 825                 830

Lys Gly


<210> SEQ ID NO 2
<211> LENGTH: 832
<212> TYPE: PRT
<213> ORGANISM: Thermus aquaticus
<220> FEATURE:
<223> OTHER INFORMATION: Thermus aquaticus DNA polymerase (Taq)

<400> SEQUENCE: 2

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
```

-continued

```
    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 Glu Leu Arg Val Leu Ala His Leu Ser Gly
    610                 615                 620
```

-continued

```
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


<210> SEQ ID NO 3
<211> LENGTH: 830
<212> TYPE: PRT
<213> ORGANISM: Thermus filiformis
<220> FEATURE:
<223> OTHER INFORMATION: Thermus filiformis DNA polymerase (Tfi)

<400> SEQUENCE: 3

Met Leu Pro Leu Leu Glu Pro Lys Gly Arg Val Leu Leu Val Asp Gly
1               5                   10                  15

His His Leu Ala Tyr Arg Thr Phe Phe Ala Leu Lys Gly Leu Thr Thr
            20                  25                  30

Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe Ala Lys Ser Leu
        35                  40                  45

Leu Lys Ala Leu Lys Glu Asp Gly Glu Val Ala Ile Val Val Phe Asp
    50                  55                  60

Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Glu Ala Tyr Lys Ala
65                  70                  75                  80

Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln Leu Ala Leu Ile
                85                  90                  95

Lys Glu Leu Val Asp Leu Leu Gly Leu Val Arg Leu Glu Val Pro Gly
            100                 105                 110

Phe Glu Ala Asp Asp Val Leu Ala Thr Leu Ala Arg Lys Ala Glu Arg
        115                 120                 125

Glu Gly Tyr Glu Val Arg Ile Leu Ser Ala Asp Arg Asp Leu Tyr Gln
    130                 135                 140
```

-continued

```
Leu Leu Ser Asp Arg Ile His Leu Leu His Pro Glu Gly Glu Val Leu
145                 150                 155                 160

Thr Pro Gly Trp Leu Gln Glu Arg Tyr Gly Leu Ser Pro Glu Arg Trp
                165                 170                 175

Val Glu Tyr Arg Ala Leu Val Gly Asp Pro Ser Asp Asn Leu Pro Gly
            180                 185                 190

Val Pro Gly Ile Gly Glu Lys Thr Ala Leu Lys Leu Leu Lys Glu Trp
        195                 200                 205

Gly Ser Leu Glu Ala Ile Leu Lys Asn Leu Asp Gln Val Lys Pro Glu
    210                 215                 220

Arg Val Trp Glu Ala Ile Arg Asn Asn Leu Asp Lys Leu Gln Met Ser
225                 230                 235                 240

Leu Glu Leu Ser Arg Leu Arg Thr Asp Leu Pro Leu Glu Val Asp Phe
                245                 250                 255

Ala Lys Arg Arg Glu Pro Thr Gly Lys Gly Leu Lys Ala Phe Leu Glu
            260                 265                 270

Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly Leu Leu Glu Ala
        275                 280                 285

Pro Lys Glu Ala Glu Glu Ala Pro Trp Pro Pro Pro Gly Gly Ala Phe
    290                 295                 300

Leu Gly Phe Leu Leu Ser Arg Pro Glu Pro Met Trp Ala Glu Leu Leu
305                 310                 315                 320

Ala Leu Ala Gly Ala Lys Glu Gly Arg Val His Arg Ala Glu Asp Pro
                325                 330                 335

Val Gly Ala Leu Lys Asp Leu Lys Glu Ile Arg Gly Leu Leu Ala Lys
            340                 345                 350

Asp Leu Ser Val Leu Ala Leu Arg Glu Gly Arg Glu Ile Pro Pro Gly
        355                 360                 365

Asp Asp Pro Met Leu Leu Ala Tyr Leu Leu Asp Pro Gly Asn Thr Asn
    370                 375                 380

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

Ala Ala Arg Ala Leu Leu Ser Glu Arg Leu Trp Gln Ala Leu Tyr Pro
                405                 410                 415

Arg Val Ala Glu Glu Glu Arg Leu Leu Trp Leu Tyr Arg Glu Val Glu
            420                 425                 430

Arg Pro Leu Ala Gln Val Leu Ala His Met Glu Ala Thr Gly Val Arg
        435                 440                 445

Leu Asp Val Pro Tyr Leu Glu Ala Leu Ser Gln Glu Val Ala Phe Glu
    450                 455                 460

Leu Glu Arg Leu Glu Ala Glu Val His Arg Leu Ala Gly His Pro Phe
465                 470                 475                 480

Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg Val Leu Phe Asp Glu Leu
                485                 490                 495

Gly Leu Pro Pro Ile Gly Lys Thr Glu Lys Thr Gly Lys Arg Ser Thr
            500                 505                 510

Ser Ala Ala Val Leu Glu Leu Leu Arg Glu Ala His Pro Ile Val Gly
        515                 520                 525

Arg Ile Leu Glu Tyr Arg Glu Leu Met Lys Leu Lys Ser Thr Tyr Ile
    530                 535                 540
```

-continued

```
Asp Pro Leu Pro Arg Leu Val His Pro Lys Thr Gly Arg Leu His Thr
545                 550                 555                 560

Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly Arg Leu Ser Ser Ser Asp
                565                 570                 575

Pro Asn Leu Gln Asn Ile Pro Val Arg Thr Pro Leu Gly Gln Arg Ile
            580                 585                 590

Arg Lys Ala Phe Ile Ala Glu Glu Gly His Leu Leu Val Ala Leu Asp
        595                 600                 605

Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala His Leu Ser Gly Asp Glu
    610                 615                 620

Asn Leu Ile Arg Val Phe Arg Glu Gly Lys Asp Ile His Thr Glu Thr
625                 630                 635                 640

Ala Ala Trp Met Phe Gly Val Pro Pro Glu Gly Val Asp Gly Ala Met
                645                 650                 655

Arg Arg Ala Ala Lys Thr Val Asn Phe Gly Val Leu Tyr Gly Met Ser
            660                 665                 670

Ala His Arg Leu Ser Gln Glu Leu Ser Ile Pro Tyr Glu Glu Ala Ala
        675                 680                 685

Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys Val Arg Ala Trp
    690                 695                 700

Ile Ala Lys Thr Leu Glu Glu Gly Arg Lys Lys Gly Tyr Val Glu Thr
705                 710                 715                 720

Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Asn Ala Arg Val Lys
                725                 730                 735

Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met Pro Val Gln
            740                 745                 750

Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys Leu Phe Pro
        755                 760                 765

Arg Leu Arg Pro Leu Gly Val Arg Ile Leu Leu Gln Val His Asp Glu
    770                 775                 780

Leu Val Leu Glu Ala Pro Lys Ala Arg Ala Glu Glu Ala Ala Gln Leu
785                 790                 795                 800

Ala Lys Glu Thr Met Glu Gly Val Tyr Pro Leu Ser Val Pro Leu Glu
                805                 810                 815

Val Glu Val Gly Met Gly Glu Asp Trp Leu Ser Ala Lys Glu
            820                 825                 830


<210> SEQ ID NO 4
<211> LENGTH: 831
<212> TYPE: PRT
<213> ORGANISM: Thermus flavus
<220> FEATURE:
<223> OTHER INFORMATION: Thermus flavus DNA polymerase (Tfl)

<400> SEQUENCE: 4

Met Ala Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu Leu Val
1               5                   10                  15

Asp Gly His His Leu Ala Tyr Arg Thr Phe Phe Ala Leu Lys Gly Leu
            20                  25                  30

Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe Ala Lys
        35                  40                  45

Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Asp Val Val Val Val Val
    50                  55                  60
```

-continued

```
Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Glu Ala Tyr
65                  70                  75                  80

Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln Leu Ala
                85                  90                  95

Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Leu Val Arg Leu Glu Val
            100                 105                 110

Pro Gly Phe Glu Ala Asp Asp Val Leu Ala Thr Leu Ala Lys Arg Ala
        115                 120                 125

Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Arg Asp Leu
    130                 135                 140

Tyr Gln Leu Leu Ser Glu Arg Ile Ala Ile Leu His Pro Glu Gly Tyr
145                 150                 155                 160

Leu Ile Thr Pro Ala Trp Leu Tyr Glu Lys Tyr Gly Leu Arg Pro Glu
                165                 170                 175

Gln Trp Val Asp Tyr Arg Ala Leu Ala Gly Asp Pro Ser Asp Asn Ile
            180                 185                 190

Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Gln Arg Leu Ile Arg
        195                 200                 205

Glu Trp Gly Ser Leu Glu Asn Leu Phe Gln His Leu Asp Gln Val Lys
    210                 215                 220

Pro Ser Leu Arg Glu Lys Leu Gln Ala Gly Met Glu Ala Leu Ala Leu
225                 230                 235                 240

Ser Arg Lys Leu Ser Gln Val His Thr Asp Leu Pro Leu Glu Val Asp
                245                 250                 255

Phe Gly Arg Arg Arg Thr Pro Asn Leu Glu Gly Leu Arg Ala Phe Leu
            260                 265                 270

Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly Leu Leu Glu
        275                 280                 285

Gly Pro Lys Ala Ala Glu Glu Ala Pro Trp Pro Pro Pro Glu Gly Ala
    290                 295                 300

Phe Leu Gly Phe Ser Phe Ser Arg Pro Glu Pro Met Trp Ala Glu Leu
305                 310                 315                 320

Leu Ala Leu Ala Gly Ala Trp Glu Gly Arg Leu His Arg Ala Gln Asp
                325                 330                 335

Pro Leu Arg Gly Leu Arg Asp Leu Lys Gly Val Arg Gly Ile Leu Ala
            340                 345                 350

Lys Asp Leu Ala Val Leu Ala Leu Arg Glu Gly Leu Asp Leu Phe Pro
        355                 360                 365

Glu Asp Asp Pro Met Leu Leu Ala Tyr Leu Leu Asp Pro Ser Asn Thr
    370                 375                 380

Thr Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly Glu Trp Thr Glu Asp
385                 390                 395                 400

Ala Gly Glu Arg Ala Leu Leu Ala Glu Arg Leu Phe Gln Thr Leu Lys
                405                 410                 415

Glu Arg Leu Lys Gly Glu Glu Arg Leu Leu Trp Leu Tyr Glu Glu Val
            420                 425                 430

Glu Lys Pro Leu Ser Arg Val Leu Ala Arg Met Glu Ala Thr Gly Val
        435                 440                 445

Arg Leu Asp Val Ala Tyr Leu Gln Ala Leu Ser Leu Glu Val Glu Ala
    450                 455                 460

Glu Val Arg Gln Leu Glu Glu Glu Val Phe Arg Leu Ala Gly His Pro
```

-continued

```
465                 470                 475                 480

Phe Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg Val Leu Phe Asp Glu
                485                 490                 495

Leu Gly Leu Pro Ala Ile Gly Lys Thr Glu Lys Thr Gly Lys Arg Ser
            500                 505                 510

Thr Ser Ala Ala Val Leu Glu Ala Leu Arg Glu Ala His Pro Ile Val
        515                 520                 525

Asp Arg Ile Leu Gln Tyr Arg Glu Leu Thr Lys Leu Lys Asn Thr Tyr
    530                 535                 540

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

Thr Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly Arg Leu Ser Ser Ser
                565                 570                 575

Asp Pro Asn Leu Gln Asn Ile Pro Val Arg Thr Pro Leu Gly Gln Arg
            580                 585                 590

Ile Arg Arg Ala Phe Val Ala Glu Glu Gly Trp Val Leu Val Val Leu
        595                 600                 605

Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala His Leu Ser Gly Asp
    610                 615                 620

Glu Asn Leu Ile Arg Val Phe Gln Glu Gly Arg Asp Ile His Thr Gln
625                 630                 635                 640

Thr Ala Ser Trp Met Phe Gly Val Ser Pro Glu Gly Val Asp Pro Leu
                645                 650                 655

Met Arg Arg Ala Ala Lys Thr Ile Asn Phe Gly Val Leu Tyr Gly Met
            660                 665                 670

Ser Ala His Arg Leu Ser Gly Glu Leu Ser Ile Pro Tyr Glu Glu Ala
        675                 680                 685

Val Ala Phe Ile Glu Arg Tyr Phe Gln Ser Tyr Pro Lys Val Arg Ala
    690                 695                 700

Trp Ile Glu Gly Thr Leu Glu Glu Gly Arg Arg Arg Gly Tyr Val Glu
705                 710                 715                 720

Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Asn Ala Arg Val
                725                 730                 735

Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met Pro Val
            740                 745                 750

Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Arg Leu Phe
        755                 760                 765

Pro Arg Leu Gln Glu Leu Gly Ala Arg Met Leu Leu Gln Val His Asp
    770                 775                 780

Glu Leu Val Leu Glu Ala Pro Lys Asp Arg Ala Glu Arg Val Ala Ala
785                 790                 795                 800

Leu Ala Lys Glu Val Met Glu Gly Val Trp Pro Leu Gln Val Pro Leu
                805                 810                 815

Glu Val Glu Val Gly Leu Gly Glu Asp Trp Leu Ser Ala Lys Glu
            820                 825                 830


<210> SEQ ID NO 5
<211> LENGTH: 830
<212> TYPE: PRT
<213> ORGANISM: Thermus sp.
<220> FEATURE:
<223> OTHER INFORMATION: Thermus sp. sps17 DNA polymerase (Sps17)
```

-continued

```
<400> SEQUENCE: 5

Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu Leu Val Asp Gly
1               5                   10                  15

His His Leu Ala Tyr Arg Thr Phe Phe Ala Leu Lys Gly Leu Thr Thr
            20                  25                  30

Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe Ala Lys Ser Leu
        35                  40                  45

Leu Lys Ala Leu Lys Glu Asp Gly Glu Val Ala Ile Val Val Phe Asp
    50                  55                  60

Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Glu Ala Tyr Lys Ala
65                  70                  75                  80

Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln Leu Ala Leu Ile
                85                  90                  95

Lys Glu Leu Val Asp Leu Leu Gly Leu Val Arg Leu Glu Val Pro Gly
            100                 105                 110

Phe Glu Ala Asp Asp Val Leu Ala Thr Leu Ala Lys Lys Ala Glu Arg
        115                 120                 125

Glu Gly Tyr Glu Val Arg Ile Leu Ser Ala Asp Arg Asp Leu Tyr Gln
    130                 135                 140

Leu Leu Ser Asp Arg Ile His Leu Leu His Pro Glu Gly Glu Val Leu
145                 150                 155                 160

Thr Pro Gly Trp Leu Gln Glu Arg Tyr Gly Leu Ser Pro Glu Arg Trp
                165                 170                 175

Val Glu Tyr Arg Ala Leu Val Gly Asp Pro Ser Asp Asn Leu Pro Gly
            180                 185                 190

Val Pro Gly Ile Gly Glu Lys Thr Ala Leu Lys Leu Leu Lys Glu Trp
        195                 200                 205

Gly Ser Leu Glu Ala Ile Leu Lys Asn Leu Asp Gln Val Lys Pro Glu
    210                 215                 220

Arg Val Arg Glu Ala Ile Arg Asn Asn Leu Asp Lys Leu Gln Met Ser
225                 230                 235                 240

Leu Glu Leu Ser Arg Leu Arg Thr Asp Leu Pro Leu Glu Val Asp Phe
                245                 250                 255

Ala Lys Arg Arg Glu Pro Asp Trp Glu Gly Leu Lys Ala Phe Leu Glu
            260                 265                 270

Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly Leu Leu Glu Ala
        275                 280                 285

Pro Lys Glu Ala Glu Glu Ala Pro Trp Pro Pro Pro Gly Gly Ala Phe
    290                 295                 300

Leu Gly Phe Leu Leu Ser Arg Pro Glu Pro Met Trp Ala Glu Leu Leu
305                 310                 315                 320

Ala Leu Ala Gly Ala Lys Glu Gly Arg Val His Arg Ala Glu Asp Pro
                325                 330                 335

Val Gly Ala Leu Lys Asp Leu Lys Glu Ile Arg Gly Leu Leu Ala Lys
            340                 345                 350

Asp Leu Ser Val Leu Ala Leu Arg Glu Gly Arg Glu Ile Pro Pro Gly
        355                 360                 365

Asp Asp Pro Met Leu Leu Ala Tyr Leu Leu Asp Pro Gly Asn Thr Asn
    370                 375                 380

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

-continued

```
Ala Ala Arg Ala Leu Leu Ser Glu Arg Leu Trp Gln Ala Leu Tyr Pro
                405                 410                 415

Arg Val Ala Glu Glu Glu Arg Leu Leu Trp Leu Tyr Arg Glu Val Glu
            420                 425                 430

Arg Pro Leu Ala Gln Val Leu Ala His Met Glu Ala Thr Gly Val Arg
        435                 440                 445

Leu Asp Val Pro Tyr Leu Glu Ala Leu Ser Gln Glu Val Ala Phe Glu
    450                 455                 460

Leu Glu Arg Leu Glu Ala Glu Val His Arg Leu Ala Gly His Pro Phe
465                 470                 475                 480

Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg Val Leu Phe Asp Glu Leu
                485                 490                 495

Gly Leu Pro Pro Ile Gly Lys Thr Glu Lys Thr Gly Lys Arg Ser Thr
            500                 505                 510

Ser Ala Ala Val Leu Glu Leu Leu Arg Glu Ala His Pro Ile Val Gly
        515                 520                 525

Arg Ile Leu Glu Tyr Arg Glu Leu Met Lys Leu Lys Ser Thr Tyr Ile
    530                 535                 540

Asp Pro Leu Pro Arg Leu Val His Pro Lys Thr Gly Arg Leu His Thr
545                 550                 555                 560

Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly Arg Leu Ser Ser Ser Asp
                565                 570                 575

Pro Asn Leu Gln Asn Ile Pro Val Arg Thr Pro Leu Gly Gln Arg Ile
            580                 585                 590

Arg Lys Ala Phe Ile Ala Glu Glu Gly His Leu Leu Val Ala Leu Asp
        595                 600                 605

Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala His Leu Ser Gly Asp Glu
    610                 615                 620

Asn Leu Ile Arg Val Phe Arg Glu Gly Lys Asp Ile His Thr Glu Thr
625                 630                 635                 640

Ala Ala Trp Met Phe Gly Val Pro Pro Glu Gly Val Asp Gly Ala Met
                645                 650                 655

Arg Arg Ala Ala Lys Thr Val Asn Phe Gly Val Leu Tyr Gly Met Ser
            660                 665                 670

Ala His Arg Leu Ser Gln Glu Leu Ser Ile Pro Tyr Glu Glu Ala Ala
        675                 680                 685

Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys Val Arg Ala Trp
    690                 695                 700

Ile Ala Lys Thr Leu Glu Glu Gly Arg Lys Lys Gly Tyr Val Glu Thr
705                 710                 715                 720

Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Asn Ala Arg Val Lys
                725                 730                 735

Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met Pro Val Gln
            740                 745                 750

Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys Leu Phe Pro
        755                 760                 765

Arg Leu Arg Pro Leu Gly Val Arg Ile Leu Leu Gln Val His Asp Glu
    770                 775                 780

Leu Val Leu Glu Ala Pro Lys Ala Arg Ala Glu Glu Ala Ala Gln Leu
785                 790                 795                 800
```

-continued

```
Ala Lys Glu Thr Met Glu Gly Val Tyr Pro Leu Ser Val Pro Leu Glu
                805                 810                 815

Val Glu Val Gly Met Gly Glu Asp Trp Leu Ser Ala Lys Ala
            820                 825                 830


<210> SEQ ID NO 6
<211> LENGTH: 834
<212> TYPE: PRT
<213> ORGANISM: Thermus thermophilus
<220> FEATURE:
<223> OTHER INFORMATION: Thermus thermophilus DNA polymerase (Tth)

<400> SEQUENCE: 6

Met Glu Ala 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 Phe 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 Tyr Lys Ala Val Phe
    50                  55                  60

Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Glu
65                  70                  75                  80

Ala Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln
                85                  90                  95

Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Phe Thr Arg Leu
            100                 105                 110

Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Thr Leu Ala Lys
        115                 120                 125

Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Arg
    130                 135                 140

Asp Leu Tyr Gln Leu Val Ser Asp Arg Val Ala Val Leu His Pro Glu
145                 150                 155                 160

Gly His Leu Ile Thr Pro Glu Trp Leu Trp Glu Lys Tyr Gly Leu Arg
                165                 170                 175

Pro Glu Gln Trp Val Asp Phe Arg Ala Leu Val Gly Asp Pro Ser Asp
            180                 185                 190

Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Leu Lys Leu
        195                 200                 205

Leu Lys Glu Trp Gly Ser Leu Glu Asn Leu Leu Lys Asn Leu Asp Arg
    210                 215                 220

Val Lys Pro Glu Asn Val Arg Glu Lys Ile Lys Ala His Leu Glu Asp
225                 230                 235                 240

Leu Arg Leu Ser Leu Glu Leu Ser Arg Val Arg Thr Asp Leu Pro Leu
                245                 250                 255

Glu Val Asp Leu Ala Gln Gly Arg Glu Pro Asp Arg Glu Gly Leu Arg
            260                 265                 270

Ala Phe Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly
        275                 280                 285

Leu Leu Glu Ala Pro Ala Pro Leu Glu Glu Ala Pro Trp Pro Pro Pro
    290                 295                 300

Glu Gly Ala Phe Val Gly Phe Val Leu Ser Arg Pro Glu Pro Met Trp
305                 310                 315                 320
```

-continued

```
Ala Glu Leu Lys Ala Leu Ala Ala Cys Arg Asp Gly Arg Val His Arg
                325                 330                 335

Ala Ala Asp Pro Leu Ala Gly Leu Lys Asp Leu Lys Glu Val Arg Gly
            340                 345                 350

Leu Leu Ala Lys Asp Leu Ala Val Leu Ala Ser Arg Glu Gly Leu Asp
        355                 360                 365

Leu Val Pro Gly Asp Asp Pro Met Leu Leu Ala Tyr Leu Leu Asp Pro
    370                 375                 380

Ser Asn Thr Thr Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly Glu Trp
385                 390                 395                 400

Thr Glu Asp Ala Ala His Arg Ala Leu Leu Ser Glu Arg Leu His Arg
                405                 410                 415

Asn Leu Leu Lys Arg Leu Glu Gly Glu Glu Lys Leu Leu Trp Leu Tyr
            420                 425                 430

His Glu Val Glu Lys Pro Leu Ser Arg Val Leu Ala His Met Glu Ala
        435                 440                 445

Thr Gly Val Arg Arg Asp Val Ala Tyr Leu Gln Ala Leu Ser Leu Glu
    450                 455                 460

Leu Ala Glu Glu Ile Arg Arg Leu Glu Glu Glu Val Phe Arg Leu Ala
465                 470                 475                 480

Gly His Pro Phe Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg Val Leu
                485                 490                 495

Phe Asp Glu Leu Arg Leu Pro Ala Leu Gly Lys Thr Gln Lys Thr Gly
            500                 505                 510

Lys Arg Ser Thr Ser Ala Ala Val Leu Glu Ala Leu Arg Glu Ala His
        515                 520                 525

Pro Ile Val Glu Lys Ile Leu Gln His Arg Glu Leu Thr Lys Leu Lys
    530                 535                 540

Asn Thr Tyr Val Asp Pro Leu Pro Ser Leu Val His Pro Arg Thr Gly
545                 550                 555                 560

Arg Leu His Thr Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly Arg Leu
                565                 570                 575

Ser Ser Ser Asp Pro Asn Leu Gln Asn Ile Pro Val Arg Thr Pro Leu
            580                 585                 590

Gly Gln Arg Ile Arg Arg Ala Phe Val Ala Glu Ala Gly Trp Ala Leu
        595                 600                 605

Val Ala Leu Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala His Leu
    610                 615                 620

Ser Gly Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly Lys Asp Ile
625                 630                 635                 640

His Thr Gln Thr Ala Ser Trp Met Phe Gly Val Pro Pro Glu Ala Val
                645                 650                 655

Asp Pro Leu Met Arg Arg Ala Ala Lys Thr Val Asn Phe Gly Val Leu
            660                 665                 670

Tyr Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile Pro Tyr
        675                 680                 685

Glu Glu Ala Val Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys
    690                 695                 700

Val Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Lys Arg Gly
705                 710                 715                 720

Tyr Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Asn
```

-continued

725 730 735

Ala Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn
740 745 750

Met Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val
755 760 765

Lys Leu Phe Pro Arg Leu Arg Glu Met Gly Ala Arg Met Leu Leu Gln
770 775 780

Val His Asp Glu Leu Leu Leu Glu Ala Pro Gln Ala Arg Ala Glu Glu
785 790 795 800

Val Ala Ala Leu Ala Lys Glu Ala Met Glu Lys Ala Tyr Pro Leu Ala
805 810 815

Val Pro Leu Glu Val Glu Val Gly Met Gly Glu Asp Trp Leu Ser Ala
820 825 830

Lys Gly

<210> SEQ ID NO 7
<211> LENGTH: 834
<212> TYPE: PRT
<213> ORGANISM: Thermus caldophilus
<220> FEATURE:
<223> OTHER INFORMATION: Thermus caldophilus DNA polymerase (Tca)

<400> SEQUENCE: 7

Met Glu Ala 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 Phe 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 Tyr Lys Ala Val Phe
50 55 60

Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Glu
65 70 75 80

Ala Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln
85 90 95

Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Phe Thr Arg Leu
100 105 110

Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Thr Leu Ala Lys
115 120 125

Asn Pro Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Arg
130 135 140

Asp Leu Asp Gln Leu Val Ser Asp Arg Val Ala Val Leu His Pro Glu
145 150 155 160

Gly His Leu Ile Thr Pro Glu Trp Leu Trp Gln Lys Tyr Gly Leu Lys
165 170 175

Pro Glu Gln Trp Val Asp Phe Arg Ala Leu Val Gly Asp Pro Ser Asp
180 185 190

Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Leu Lys Leu
195 200 205

Leu Lys Glu Trp Gly Ser Leu Glu Asn Leu Leu Lys Asn Leu Asp Arg
210 215 220

Val Lys Pro Glu Asn Val Arg Glu Lys Ile Lys Ala His Leu Glu Asp
225 230 235 240

-continued

```
Leu Arg Leu Ser Leu Glu Leu Ser Arg Val Arg Thr Asp Leu Pro Leu
                245                 250                 255

Glu Val Asp Leu Ala Gln Gly Arg Glu Pro Asp Arg Glu Gly Leu Arg
            260                 265                 270

Ala Phe Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly
        275                 280                 285

Leu Leu Glu Ala Pro Ala Pro Leu Glu Glu Ala Pro Trp Pro Pro Pro
    290                 295                 300

Glu Gly Ala Phe Val Gly Phe Val Leu Ser Arg Pro Glu Pro Met Trp
305                 310                 315                 320

Ala Glu Leu Lys Ala Leu Ala Ala Cys Arg Asp Gly Arg Val His Arg
                325                 330                 335

Ala Ala Asp Pro Leu Ala Gly Leu Lys Asp Leu Lys Glu Val Arg Gly
            340                 345                 350

Leu Leu Ala Lys Asp Leu Ala Val Leu Ala Ser Arg Glu Gly Leu Asp
        355                 360                 365

Leu Val Pro Gly Asp Asp Pro Met Leu Leu Ala Tyr Leu Leu Asp Pro
    370                 375                 380

Ser Asn Thr Thr Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly Glu Trp
385                 390                 395                 400

Thr Glu Asp Ala Ala His Arg Ala Leu Leu Ser Glu Arg Leu His Arg
                405                 410                 415

Asn Leu Leu Lys Arg Leu Gln Gly Glu Glu Lys Leu Leu Trp Leu Tyr
            420                 425                 430

His Glu Val Glu Lys Pro Leu Ser Arg Val Leu Ala His Met Glu Ala
        435                 440                 445

Thr Gly Val Arg Leu Asp Val Ala Tyr Leu Gln Ala Leu Ser Leu Glu
    450                 455                 460

Leu Ala Glu Glu Ile Arg Arg Leu Glu Glu Glu Val Phe Arg Leu Ala
465                 470                 475                 480

Gly His Pro Phe Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg Val Leu
                485                 490                 495

Phe Asp Glu Leu Arg Leu Pro Ala Leu Gly Lys Thr Gln Lys Thr Gly
            500                 505                 510

Lys Arg Ser Thr Ser Ala Ala Val Leu Glu Ala Leu Arg Glu Ala His
        515                 520                 525

Pro Ile Val Glu Lys Ile Leu Gln His Arg Glu Leu Thr Lys Leu Lys
    530                 535                 540

Asn Thr Tyr Val Asp Pro Leu Pro Ser Leu Val His Pro Asn Thr Gly
545                 550                 555                 560

Arg Leu His Thr Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly Arg Leu
                565                 570                 575

Ser Ser Ser Asp Pro Asn Leu Gln Asn Ile Pro Val Arg Thr Pro Leu
            580                 585                 590

Gly Gln Arg Ile Arg Arg Ala Phe Val Ala Glu Ala Gly Trp Ala Leu
        595                 600                 605

Val Ala Leu Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala His Leu
    610                 615                 620

Ser Gly Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly Lys Asp Ile
625                 630                 635                 640
```

-continued

```
His Thr Gln Thr Ala Ser Trp Met Phe Gly Val Pro Pro Glu Ala Val
                645                 650                 655

Asp Pro Leu Met Arg Arg Ala Ala Lys Thr Val Asn Phe Gly Val Leu
            660                 665                 670

Tyr Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile Pro Tyr
        675                 680                 685

Glu Glu Ala Val Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys
    690                 695                 700

Val Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Lys Arg Gly
705                 710                 715                 720

Tyr Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Asn
                725                 730                 735

Ala Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn
            740                 745                 750

Met Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val
        755                 760                 765

Lys Leu Phe Pro Arg Leu Arg Glu Met Gly Ala Arg Met Leu Leu Gln
    770                 775                 780

Val His Asp Glu Leu Leu Leu Glu Ala Pro Gln Ala Gly Ala Glu Glu
785                 790                 795                 800

Val Ala Ala Leu Ala Lys Glu Ala Met Glu Lys Ala Tyr Pro Leu Ala
                805                 810                 815

Val Pro Leu Glu Val Glu Val Gly Met Gly Glu Asp Trp Leu Ser Ala
            820                 825                 830

Lys Gly


<210> SEQ ID NO 8
<211> LENGTH: 17
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic DNA polymerase motif
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (1)...(1)
<223> OTHER INFORMATION: Xaa = Ile, Leu, Val, Gln or Met
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (2)...(2)
<223> OTHER INFORMATION: Xaa = Arg, Lys, Gln or Glu
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (3)...(3)
<223> OTHER INFORMATION: Xaa = Val or Ala
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (5)...(5)
<223> OTHER INFORMATION: Xaa = Gln, Arg, Glu, Lys or Val
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (6)...(6)
<223> OTHER INFORMATION: Xaa = Glu or Arg
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (7)...(7)
<223> OTHER INFORMATION: Xaa = Gly or Asp
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (8)...(8)
<223> OTHER INFORMATION: Xaa = Lys, Arg, Ile, Leu or Ala
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (10)...(10)
```

-continued

```
<223> OTHER INFORMATION: Xaa = any amino acid other than Ile or Val
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (13)...(13)
<223> OTHER INFORMATION: Xaa = Gln, Glu, Leu, Ile, Arg or Lys
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (16)...(16)
<223> OTHER INFORMATION: Xaa = Ser, Ala or Met
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (17)...(17)
<223> OTHER INFORMATION: Xaa = Trp, Arg, Lys, Gln or Asp

<400> SEQUENCE: 8

Xaa Xaa Xaa Phe Xaa Xaa Xaa Xaa Asp Xaa His Thr Xaa Thr Ala Xaa
1               5                   10                  15

Xaa


<210> SEQ ID NO 9
<211> LENGTH: 17
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic DNA polymerase motif
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (5)...(5)
<223> OTHER INFORMATION: Xaa = Gln or Arg
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (8)...(8)
<223> OTHER INFORMATION: Xaa = Lys or Arg
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (10)...(10)
<223> OTHER INFORMATION: Xaa = any amino acid other than Ile
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (16)...(16)
<223> OTHER INFORMATION: Xaa = Ser or Ala

<400> SEQUENCE: 9

Ile Arg Val Phe Xaa Glu Gly Xaa Asp Xaa His Thr Glx Thr Ala Xaa
1               5                   10                  15

Trp


<210> SEQ ID NO 10
<211> LENGTH: 17
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic DNA polymerase motif
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (10)...(10)
<223> OTHER INFORMATION: Xaa = any amino acid other than Ile

<400> SEQUENCE: 10

Ile Arg Val Phe Gln Glu Gly Lys Asp Xaa His Thr Gln Thr Ala Ser
1               5                   10                  15

Trp


<210> SEQ ID NO 11
<211> LENGTH: 17
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
```

-continued

```
<220> FEATURE:
<223> OTHER INFORMATION: synthetic DNA polymerase motif

<400> SEQUENCE: 11

Ile Arg Val Phe Gln Glu Gly Lys Asp Phe His Thr Gln Thr Ala Ser
1               5                   10                  15

Trp


<210> SEQ ID NO 12
<211> LENGTH: 31
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic polymerase domain region of Thermus
      sp. Z05 DNA polymerase (Z05)

<400> SEQUENCE: 12

Ala His Leu Ser Gly Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly
1               5                   10                  15

Lys Asp Ile His Thr Gln Thr Ala Ser Trp Met Phe Gly Val Ser
            20                  25                  30


<210> SEQ ID NO 13
<211> LENGTH: 31
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic polymerase domain region of Thermus
      aquaticus DNA polymerase (Taq)

<400> SEQUENCE: 13

Ala His Leu Ser Gly Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly
1               5                   10                  15

Arg Asp Ile His Thr Glu Thr Ala Ser Trp Met Phe Gly Val Pro
            20                  25                  30


<210> SEQ ID NO 14
<211> LENGTH: 31
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic polymerase domain region of Thermus
      filiformus DNA polymerase (Tfi)

<400> SEQUENCE: 14

Ala His Leu Ser Gly Asp Glu Asn Leu Ile Arg Val Phe Arg Glu Gly
1               5                   10                  15

Lys Asp Ile His Thr Glu Thr Ala Ala Trp Met Phe Gly Val Pro
            20                  25                  30


<210> SEQ ID NO 15
<211> LENGTH: 31
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic polymerase domain region of Thermus
      flavus DNA polymerase (Tfl)

<400> SEQUENCE: 15

Ala His Leu Ser Gly Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly
1               5                   10                  15

Arg Asp Ile His Thr Gln Thr Ala Ser Trp Met Phe Gly Val Ser
```

-continued

```
            20                  25                  30


<210> SEQ ID NO 16
<211> LENGTH: 31
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic polymerase domain region of Thermus
      sp. sps17 DNA polymerase (Sps17)

<400> SEQUENCE: 16

Ala His Leu Ser Gly Asp Glu Asn Leu Ile Arg Val Phe Arg Glu Gly
1               5                   10                  15

Lys Asp Ile His Thr Glu Thr Ala Ala Trp Met Phe Gly Val Pro
            20                  25                  30


<210> SEQ ID NO 17
<211> LENGTH: 31
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic polymerase domain region of Thermus
      thermophilus DNA polymerase (Tth)

<400> SEQUENCE: 17

Ala His Leu Ser Gly Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly
1               5                   10                  15

Lys Asp Ile His Thr Gln Thr Ala Ser Trp Met Phe Gly Val Pro
            20                  25                  30


<210> SEQ ID NO 18
<211> LENGTH: 31
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic polymerase domain region of Thermus
      caldophilus DNA polymerase (Tca)

<400> SEQUENCE: 18

Ala His Leu Ser Gly Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly
1               5                   10                  15

Lys Asp Ile His Thr Gln Thr Ala Ser Trp Met Phe Gly Val Pro
            20                  25                  30


<210> SEQ ID NO 19
<211> LENGTH: 31
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic polymerase domain region of
      Thermotoga maritima DNA polymerase (Tma)

<400> SEQUENCE: 19

Ala His Leu Ser Gly Asp Glu Asn Leu Leu Arg Ala Phe Glu Glu Gly
1               5                   10                  15

Ile Asp Val His Thr Leu Thr Ala Ser Arg Ile Phe Asn Val Lys
            20                  25                  30


<210> SEQ ID NO 20
<211> LENGTH: 31
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
```

-continued

```
<223> OTHER INFORMATION: synthetic polymerase domain region of
      Thermotoga neopolitana DNA polymerase (Tne)

<400> SEQUENCE: 20

Ala His Leu Ser Gly Asp Glu Asn Leu Val Lys Ala Phe Glu Glu Gly
1               5                   10                  15

Ile Asp Val His Thr Leu Thr Ala Ser Arg Ile Tyr Asn Val Lys
            20                  25                  30


<210> SEQ ID NO 21
<211> LENGTH: 31
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic polymerase domain region of
      Thermosipho africanus DNA polymerase (Taf)

<400> SEQUENCE: 21

Ala His Val Ser Lys Asp Glu Asn Leu Leu Lys Ala Phe Lys Glu Asp
1               5                   10                  15

Leu Asp Ile His Thr Ile Thr Ala Ala Lys Ile Phe Gly Val Ser
            20                  25                  30


<210> SEQ ID NO 22
<211> LENGTH: 8
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic conserved DNA polymerase active site
      motif A

<400> SEQUENCE: 22

Asp Tyr Ser Gln Ile Glu Leu Arg
1               5


<210> SEQ ID NO 23
<211> LENGTH: 31
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic polymerase domain region of
      Deinococcus radiodurans DNA polymerase (Dra)

<400> SEQUENCE: 23

Ala His Ile Ala Asp Asp Pro Leu Met Gln Gln Ala Phe Val Glu Gly
1               5                   10                  15

Ala Asp Ile His Arg Arg Thr Ala Ala Gln Val Leu Gly Leu Asp
            20                  25                  30


<210> SEQ ID NO 24
<211> LENGTH: 31
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic polymerase domain region of Bacillus
      stearothermophilus DNA polymerase (Bst)

<400> SEQUENCE: 24

Ala His Ile Ala Glu Asp Asp Asn Leu Ile Glu Ala Phe Arg Arg Gly
1               5                   10                  15

Leu Asp Ile His Thr Lys Thr Ala Met Asp Ile Phe His Val Ser
            20                  25                  30
```

-continued

```
<210> SEQ ID NO 25
<211> LENGTH: 31
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic polymerase domain region of Bacillus
      caldotenax DNA polymerase (Bca)

<400> SEQUENCE: 25

Ala His Ile Ala Glu Asp Asp Asn Leu Met Glu Ala Phe Arg Arg Asp
1               5                   10                  15

Leu Asp Ile His Thr Lys Thr Ala Met Asp Ile Phe Gln Val Ser
            20                  25                  30


<210> SEQ ID NO 26
<211> LENGTH: 17
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic polymerase domain region native
      consensus motif
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (1)...(1)
<223> OTHER INFORMATION: Xaa = Ile, Leu, Val, Gln or Met
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (2)...(2)
<223> OTHER INFORMATION: Xaa = Arg, Lys, Gln or Glu
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (3)...(3)
<223> OTHER INFORMATION: Xaa = Val or Ala
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (5)...(5)
<223> OTHER INFORMATION: Xaa = Gln, Arg, Glu, Lys or Val
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (6)...(6)
<223> OTHER INFORMATION: Xaa = Glu or Arg
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (7)...(7)
<223> OTHER INFORMATION: Xaa = Gly or Asp
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (8)...(8)
<223> OTHER INFORMATION: Xaa = Lys, Arg, Ile, Leu or Ala
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (10)...(10)
<223> OTHER INFORMATION: Xaa = Ile or Val
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (13)...(13)
<223> OTHER INFORMATION: Xaa = Gln, Glu, Leu, Ile, Arg or Lys
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (16)...(16)
<223> OTHER INFORMATION: Xaa = Ser, Ala or Met
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (17)...(17)
<223> OTHER INFORMATION: Xaa = Trp, Arg, Lys, Gln or Asp

<400> SEQUENCE: 26

Xaa Xaa Xaa Phe Xaa Xaa Xaa Xaa Asp Xaa His Thr Xaa Thr Ala Xaa
1               5                   10                  15

Xaa
```

-continued

```
<210> SEQ ID NO 27
<211> LENGTH: 893
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic chimeric CS5 DNA polymerase derived
      from N-terminal 5'-nuclease domain of Thermus sp. Z05 and
      C-terminal 3'-5' exonuclease and polymerase domains of Thermotoga
      maritima DNA polymerases

<400> SEQUENCE: 27

Met Lys Ala 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 Phe 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 Tyr Lys Ala Val Phe
    50                  55                  60

Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Glu
65                  70                  75                  80

Ala Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln
                85                  90                  95

Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Phe Thr Arg Leu
            100                 105                 110

Glu Val Pro Gly Phe Glu Ala Asp Asp Val Leu Ala Thr Leu Ala Lys
        115                 120                 125

Lys Ala Glu Arg Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Arg
    130                 135                 140

Asp Leu Tyr Gln Leu Val Ser Asp Arg Val Ala Val Leu His Pro Glu
145                 150                 155                 160

Gly His Leu Ile Thr Pro Glu Trp Leu Trp Glu Lys Tyr Gly Leu Lys
                165                 170                 175

Pro Glu Gln Trp Val Asp Phe Arg Ala Leu Val Gly Asp Pro Ser Asp
            180                 185                 190

Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Leu Lys Leu
        195                 200                 205

Leu Lys Glu Trp Gly Ser Leu Glu Asn Ile Leu Lys Asn Leu Asp Arg
    210                 215                 220

Val Lys Pro Glu Ser Val Arg Glu Arg Ile Lys Ala His Leu Glu Asp
225                 230                 235                 240

Leu Lys Leu Ser Leu Glu Leu Ser Arg Val Arg Ser Asp Leu Pro Leu
                245                 250                 255

Glu Val Asp Phe Ala Arg Arg Arg Glu Pro Asp Arg Glu Gly Leu Arg
            260                 265                 270

Ala Phe Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly
        275                 280                 285

Leu Leu Glu Glu Ser Glu Pro Val Gly Tyr Arg Ile Val Lys Asp Leu
    290                 295                 300

Val Glu Phe Glu Lys Leu Ile Glu Lys Leu Arg Glu Ser Pro Ser Phe
305                 310                 315                 320

Ala Ile Asp Leu Glu Thr Ser Ser Leu Asp Pro Phe Asp Cys Asp Ile
                325                 330                 335
```

-continued

```
Val Gly Ile Ser Val Ser Phe Lys Pro Lys Glu Ala Tyr Tyr Ile Pro
            340                 345                 350

Leu His His Arg Asn Ala Gln Asn Leu Asp Glu Lys Glu Val Leu Lys
        355                 360                 365

Lys Leu Lys Glu Ile Leu Glu Asp Pro Gly Ala Lys Ile Val Gly Gln
    370                 375                 380

Asn Leu Lys Phe Asp Tyr Lys Val Leu Met Val Lys Gly Val Glu Pro
385                 390                 395                 400

Val Pro Pro Tyr Phe Asp Thr Met Ile Ala Ala Tyr Leu Leu Glu Pro
                405                 410                 415

Asn Glu Lys Lys Phe Asn Leu Asp Asp Leu Ala Leu Lys Phe Leu Gly
            420                 425                 430

Tyr Lys Met Thr Ser Tyr Gln Glu Leu Met Ser Phe Ser Phe Pro Leu
        435                 440                 445

Phe Gly Phe Ser Phe Ala Asp Val Pro Val Glu Lys Ala Ala Asn Tyr
    450                 455                 460

Ser Cys Glu Asp Ala Asp Ile Thr Tyr Arg Leu Tyr Lys Thr Leu Ser
465                 470                 475                 480

Leu Lys Leu His Glu Ala Asp Leu Glu Asn Val Phe Tyr Lys Ile Glu
                485                 490                 495

Met Pro Leu Val Asn Val Leu Ala Arg Met Glu Leu Asn Gly Val Tyr
            500                 505                 510

Val Asp Thr Glu Phe Leu Lys Lys Leu Ser Glu Glu Tyr Gly Lys Lys
        515                 520                 525

Leu Glu Glu Leu Ala Glu Glu Ile Tyr Arg Ile Ala Gly Glu Pro Phe
    530                 535                 540

Asn Ile Asn Ser Pro Lys Gln Val Ser Arg Ile Leu Phe Glu Lys Leu
545                 550                 555                 560

Gly Ile Lys Pro Arg Gly Lys Thr Thr Lys Thr Gly Asp Tyr Ser Thr
                565                 570                 575

Arg Ile Glu Val Leu Glu Glu Leu Ala Gly Glu His Glu Ile Ile Pro
            580                 585                 590

Leu Ile Leu Glu Tyr Arg Lys Ile Gln Lys Leu Lys Ser Thr Tyr Ile
        595                 600                 605

Asp Ala Leu Pro Lys Met Val Asn Pro Lys Thr Gly Arg Ile His Ala
    610                 615                 620

Ser Phe Asn Gln Thr Gly Thr Ala Thr Gly Arg Leu Ser Ser Ser Asp
625                 630                 635                 640

Pro Asn Leu Gln Asn Leu Pro Thr Lys Ser Glu Glu Gly Lys Glu Ile
                645                 650                 655

Arg Lys Ala Ile Val Pro Gln Asp Pro Asn Trp Trp Ile Val Ser Ala
            660                 665                 670

Asp Tyr Ser Gln Ile Glu Leu Arg Ile Leu Ala His Leu Ser Gly Asp
        675                 680                 685

Glu Asn Leu Leu Arg Ala Phe Glu Glu Gly Ile Asp Val His Thr Leu
    690                 695                 700

Thr Ala Ser Arg Ile Phe Asn Val Lys Pro Glu Glu Val Thr Glu Glu
705                 710                 715                 720

Met Arg Arg Ala Gly Lys Met Val Asn Phe Ser Ile Ile Tyr Gly Val
                725                 730                 735
```

-continued

```
Thr Pro Tyr Gly Leu Ser Val Arg Leu Gly Val Pro Val Lys Glu Ala
            740                 745                 750

Glu Lys Met Ile Val Asn Tyr Phe Val Leu Tyr Pro Lys Val Arg Asp
        755                 760                 765

Tyr Ile Gln Arg Val Val Ser Glu Ala Lys Glu Lys Gly Tyr Val Arg
    770                 775                 780

Thr Leu Phe Gly Arg Lys Arg Asp Ile Pro Gln Leu Met Ala Arg Asp
785                 790                 795                 800

Arg Asn Thr Gln Ala Glu Gly Glu Arg Ile Ala Ile Asn Thr Pro Ile
                805                 810                 815

Gln Gly Thr Ala Ala Asp Ile Ile Lys Leu Ala Met Ile Glu Ile Asp
            820                 825                 830

Arg Glu Leu Lys Glu Arg Lys Met Arg Ser Lys Met Ile Ile Gln Val
        835                 840                 845

His Asp Glu Leu Val Phe Glu Val Pro Asn Glu Glu Lys Asp Ala Leu
    850                 855                 860

Val Glu Leu Val Lys Asp Arg Met Thr Asn Val Val Lys Leu Ser Val
865                 870                 875                 880

Pro Leu Glu Val Asp Val Thr Ile Gly Lys Thr Trp Ser
                885                 890
```

<210> SEQ ID NO 28
<211> LENGTH: 893
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic chimeric CS6 DNA polymerase derived
      from N-terminal 5'-nuclease domain of Thermus sp. Z05 and
      C-terminal 3'-5' exonuclease and polymerase domains of Thermotoga
      maritima DNA polymerases

<400> SEQUENCE: 28

```
Met Lys Ala 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 Phe 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 Tyr Lys Ala Val Phe
    50                  55                  60

Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Glu
65                  70                  75                  80

Ala Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln
                85                  90                  95

Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Phe Thr Arg Leu
            100                 105                 110

Glu Val Pro Gly Phe Glu Ala Asp Asp Val Leu Ala Thr Leu Ala Lys
        115                 120                 125

Lys Ala Glu Arg Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Arg
    130                 135                 140

Asp Leu Tyr Gln Leu Val Ser Asp Arg Val Ala Val Leu His Pro Glu
145                 150                 155                 160

Gly His Leu Ile Thr Pro Glu Trp Leu Trp Glu Lys Tyr Gly Leu Lys
                165                 170                 175
```

-continued

```
Pro Glu Gln Trp Val Asp Phe Arg Ala Leu Val Gly Asp Pro Ser Asp
            180                 185                 190

Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Leu Lys Leu
        195                 200                 205

Leu Lys Glu Trp Gly Ser Leu Glu Asn Ile Leu Lys Asn Leu Asp Arg
    210                 215                 220

Val Lys Pro Glu Ser Val Arg Glu Arg Ile Lys Ala His Leu Glu Asp
225                 230                 235                 240

Leu Lys Leu Ser Leu Glu Leu Ser Arg Val Arg Ser Asp Leu Pro Leu
                245                 250                 255

Glu Val Asp Phe Ala Arg Arg Arg Glu Pro Asp Arg Glu Gly Leu Arg
            260                 265                 270

Ala Phe Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly
        275                 280                 285

Leu Leu Glu Glu Ser Glu Pro Val Gly Tyr Arg Ile Val Lys Asp Leu
    290                 295                 300

Val Glu Phe Glu Lys Leu Ile Glu Lys Leu Arg Glu Ser Pro Ser Phe
305                 310                 315                 320

Ala Ile Ala Leu Ala Thr Ser Ser Leu Asp Pro Phe Asp Cys Asp Ile
                325                 330                 335

Val Gly Ile Ser Val Ser Phe Lys Pro Lys Glu Ala Tyr Tyr Ile Pro
            340                 345                 350

Leu His His Arg Asn Ala Gln Asn Leu Asp Glu Lys Glu Val Leu Lys
        355                 360                 365

Lys Leu Lys Glu Ile Leu Glu Asp Pro Gly Ala Lys Ile Val Gly Gln
    370                 375                 380

Asn Leu Lys Phe Asp Tyr Lys Val Leu Met Val Lys Gly Val Glu Pro
385                 390                 395                 400

Val Pro Pro Tyr Phe Asp Thr Met Ile Ala Ala Tyr Leu Leu Glu Pro
                405                 410                 415

Asn Glu Lys Lys Phe Asn Leu Asp Asp Leu Ala Leu Lys Phe Leu Gly
            420                 425                 430

Tyr Lys Met Thr Ser Tyr Gln Glu Leu Met Ser Phe Ser Phe Pro Leu
        435                 440                 445

Phe Gly Phe Ser Phe Ala Asp Val Pro Val Glu Lys Ala Ala Asn Tyr
    450                 455                 460

Ser Cys Glu Asp Ala Asp Ile Thr Tyr Arg Leu Tyr Lys Thr Leu Ser
465                 470                 475                 480

Leu Lys Leu His Glu Ala Asp Leu Glu Asn Val Phe Tyr Lys Ile Glu
                485                 490                 495

Met Pro Leu Val Asn Val Leu Ala Arg Met Glu Leu Asn Gly Val Tyr
            500                 505                 510

Val Asp Thr Glu Phe Leu Lys Lys Leu Ser Glu Glu Tyr Gly Lys Lys
        515                 520                 525

Leu Glu Glu Leu Ala Glu Glu Ile Tyr Arg Ile Ala Gly Glu Pro Phe
    530                 535                 540

Asn Ile Asn Ser Pro Lys Gln Val Ser Arg Ile Leu Phe Glu Lys Leu
545                 550                 555                 560

Gly Ile Lys Pro Arg Gly Lys Thr Thr Lys Thr Gly Asp Tyr Ser Thr
                565                 570                 575

Arg Ile Glu Val Leu Glu Glu Leu Ala Gly Glu His Glu Ile Ile Pro
```

-continued

```
            580                 585                 590

Leu Ile Leu Glu Tyr Arg Lys Ile Gln Lys Leu Lys Ser Thr Tyr Ile
        595                 600                 605

Asp Ala Leu Pro Lys Met Val Asn Pro Lys Thr Gly Arg Ile His Ala
    610                 615                 620

Ser Phe Asn Gln Thr Gly Thr Ala Thr Gly Arg Leu Ser Ser Ser Asp
625                 630                 635                 640

Pro Asn Leu Gln Asn Leu Pro Thr Lys Ser Glu Glu Gly Lys Glu Ile
                645                 650                 655

Arg Lys Ala Ile Val Pro Gln Asp Pro Asn Trp Trp Ile Val Ser Ala
            660                 665                 670

Asp Tyr Ser Gln Ile Glu Leu Arg Ile Leu Ala His Leu Ser Gly Asp
        675                 680                 685

Glu Asn Leu Leu Arg Ala Phe Glu Glu Gly Ile Asp Val His Thr Leu
    690                 695                 700

Thr Ala Ser Arg Ile Phe Asn Val Lys Pro Glu Glu Val Thr Glu Glu
705                 710                 715                 720

Met Arg Arg Ala Gly Lys Met Val Asn Phe Ser Ile Ile Tyr Gly Val
                725                 730                 735

Thr Pro Tyr Gly Leu Ser Val Arg Leu Gly Val Pro Val Lys Glu Ala
            740                 745                 750

Glu Lys Met Ile Val Asn Tyr Phe Val Leu Tyr Pro Lys Val Arg Asp
        755                 760                 765

Tyr Ile Gln Arg Val Val Ser Glu Ala Lys Glu Lys Gly Tyr Val Arg
    770                 775                 780

Thr Leu Phe Gly Arg Lys Arg Asp Ile Pro Gln Leu Met Ala Arg Asp
785                 790                 795                 800

Arg Asn Thr Gln Ala Glu Gly Glu Arg Ile Ala Ile Asn Thr Pro Ile
                805                 810                 815

Gln Gly Thr Ala Ala Asp Ile Ile Lys Leu Ala Met Ile Glu Ile Asp
            820                 825                 830

Arg Glu Leu Lys Glu Arg Lys Met Arg Ser Lys Met Ile Ile Gln Val
        835                 840                 845

His Asp Glu Leu Val Phe Glu Val Pro Asn Glu Glu Lys Asp Ala Leu
    850                 855                 860

Val Glu Leu Val Lys Asp Arg Met Thr Asn Val Val Lys Leu Ser Val
865                 870                 875                 880

Pro Leu Glu Val Asp Val Thr Ile Gly Lys Thr Trp Ser
                885                 890


<210> SEQ ID NO 29
<211> LENGTH: 19
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic DNA polymerase motif
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (1)...(1)
<223> OTHER INFORMATION: Xaa = Ala, Asp, Ser, Glu, Arg or Gln
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (2)...(2)
<223> OTHER INFORMATION: Xaa = Trp or Tyr
<220> FEATURE:
<221> NAME/KEY: VARIANT
```

-continued

```
<222> LOCATION: (3)...(3)
<223> OTHER INFORMATION: Xaa = any amino acid other than Ile, Leu or Met
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (4)...(4)
<223> OTHER INFORMATION: Xaa = Glu, Ala, Gln, Lys, Asn or Asp
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (5)...(5)
<223> OTHER INFORMATION: Xaa = Lys, Gly, Arg, Gln, His or Asn
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (6)...(6)
<223> OTHER INFORMATION: Xaa = Thr, Val, Met or Ile
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (7)...(7)
<223> OTHER INFORMATION: Xaa = Leu, Val or Lys
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (8)...(8)
<223> OTHER INFORMATION: Xaa = Glu, Ser, Ala, Asp or Gln
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (9)...(9)
<223> OTHER INFORMATION: Xaa = Glu or Phe
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (10)...(10)
<223> OTHER INFORMATION: Xaa = Gly or Ala
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (11)...(11)
<223> OTHER INFORMATION: Xaa = Arg or Lys
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (12)...(12)
<223> OTHER INFORMATION: Xaa = Lys, Arg, Glu, Thr or Gln
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (13)...(13)
<223> OTHER INFORMATION: Xaa = Arg, Lys or His
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (17)...(17)
<223> OTHER INFORMATION: Xaa = Glu, Arg or Thr

<400> SEQUENCE: 29

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Gly Tyr Val
1               5                   10                  15

Xaa Thr Leu


<210> SEQ ID NO 30
<211> LENGTH: 21
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic error-prone (mutagenic) PCR
      amplification forward primer

<400> SEQUENCE: 30

ctacctcctg gacccctcca a                                                21


<210> SEQ ID NO 31
<211> LENGTH: 25
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic error-prone (mutagenic) PCR
      amplification reverse primer
```

-continued

```
<400> SEQUENCE: 31

ataaccaact ggtagtggcg tgtaa                                           25


<210> SEQ ID NO 32
<211> LENGTH: 921
<212> TYPE: PRT
<213> ORGANISM: Deinococcus radiodurans
<220> FEATURE:
<223> OTHER INFORMATION: Deinococcus radiodurans DNA polymerase (Dra)

<400> SEQUENCE: 32

Met Ala Asp Ala Ser Pro Asp Pro Ser Lys Pro Asp Ala Leu Val Leu
1               5                   10                  15

Ile Asp Gly His Ala Leu Ala Phe Arg Ser Tyr Phe Ala Leu Pro Pro
            20                  25                  30

Leu Asn Asn Ser Lys Gly Glu Met Thr Asp Ala Ile Val Gly Phe Met
        35                  40                  45

Lys Leu Leu Leu Arg Leu Ala Arg Gln Lys Ser Asn Gln Val Ile Val
    50                  55                  60

Val Phe Asp Pro Pro Val Lys Thr Leu Arg His Glu Gln Tyr Glu Gly
65                  70                  75                  80

Tyr Lys Ser Gly Arg Ala Gln Thr Pro Glu Asp Leu Arg Gly Gln Ile
                85                  90                  95

Asn Arg Ile Arg Ala Leu Val Asp Ala Leu Gly Phe Pro Arg Leu Glu
            100                 105                 110

Glu Pro Gly Tyr Glu Ala Asp Asp Val Ile Ala Ser Leu Thr Arg Met
        115                 120                 125

Ala Glu Gly Lys Gly Tyr Glu Val Arg Ile Val Thr Ser Asp Arg Asp
    130                 135                 140

Ala Tyr Gln Leu Leu Asp Glu His Val Lys Val Ile Ala Asn Asp Phe
145                 150                 155                 160

Ser Leu Ile Gly Pro Ala Gln Val Glu Glu Lys Tyr Gly Val Thr Val
                165                 170                 175

Arg Gln Trp Val Asp Tyr Arg Ala Leu Thr Gly Asp Ala Ser Asp Asn
            180                 185                 190

Ile Pro Gly Ala Lys Gly Ile Gly Pro Lys Thr Ala Ala Lys Leu Leu
        195                 200                 205

Gln Glu Tyr Gly Thr Leu Glu Lys Val Tyr Glu Ala Ala His Ala Gly
    210                 215                 220

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

Asn Val Lys Phe Ser His Asp Leu Ser Cys Met Val Thr Asp Leu Pro
                245                 250                 255

Leu Asp Ile Glu Phe Gly Val Arg Arg Leu Pro Asp Asn Pro Leu Val
            260                 265                 270

Thr Glu Asp Leu Leu Thr Glu Leu Glu Leu His Ser Leu Arg Pro Met
        275                 280                 285

Ile Leu Gly Leu Asn Gly Pro Glu Gln Asp Gly His Ala Pro Asp Asp
    290                 295                 300

Leu Leu Glu Arg Glu His Ala Gln Thr Pro Glu Glu Asp Glu Ala Ala
305                 310                 315                 320

Ala Leu Pro Ala Phe Ser Ala Pro Glu Leu Ala Glu Trp Gln Thr Pro
                325                 330                 335
```

-continued

```
Ala Glu Gly Ala Val Trp Gly Tyr Val Leu Ser Arg Glu Asp Asp Leu
            340                 345                 350

Thr Ala Ala Leu Leu Ala Ala Ala Thr Phe Glu Asp Gly Val Ala Arg
        355                 360                 365

Pro Ala Arg Val Ser Glu Pro Asp Glu Trp Ala Gln Ala Glu Ala Pro
    370                 375                 380

Glu Asn Leu Phe Gly Glu Leu Leu Pro Ser Asp Lys Pro Leu Thr Lys
385                 390                 395                 400

Lys Glu Gln Lys Ala Leu Glu Lys Ala Gln Lys Asp Ala Glu Lys Ala
                405                 410                 415

Arg Ala Lys Leu Arg Glu Gln Phe Pro Ala Thr Val Asp Glu Ala Glu
            420                 425                 430

Phe Val Gly Gln Arg Thr Val Thr Ala Ala Ala Ala Lys Ala Leu Ala
        435                 440                 445

Ala His Leu Ser Val Arg Gly Thr Val Val Glu Pro Gly Asp Asp Pro
    450                 455                 460

Leu Leu Tyr Ala Tyr Leu Leu Asp Pro Ala Asn Thr Asn Met Pro Val
465                 470                 475                 480

Val Ala Lys Arg Tyr Leu Asp Arg Glu Trp Pro Ala Asp Ala Pro Thr
                485                 490                 495

Arg Ala Ala Ile Thr Gly His Leu Val Arg Glu Leu Pro Pro Leu Leu
            500                 505                 510

Asp Asp Ala Arg Arg Lys Met Tyr Asp Glu Met Glu Lys Pro Leu Ser
        515                 520                 525

Gly Val Leu Gly Arg Met Glu Val Arg Gly Val Gln Val Asp Ser Asp
    530                 535                 540

Phe Leu Gln Thr Leu Ser Ile Gln Ala Gly Val Arg Leu Ala Asp Leu
545                 550                 555                 560

Glu Ser Gln Ile His Glu Tyr Ala Gly Glu Glu Phe His Ile Arg Ser
                565                 570                 575

Pro Lys Gln Leu Glu Thr Val Leu Tyr Asp Lys Leu Glu Leu Ala Ser
            580                 585                 590

Ser Lys Lys Thr Lys Leu Thr Gly Gln Arg Ser Thr Ala Val Ser Ala
        595                 600                 605

Leu Glu Pro Leu Arg Asp Ala His Pro Ile Ile Pro Leu Val Leu Glu
    610                 615                 620

Phe Arg Glu Leu Asp Lys Leu Arg Gly Thr Tyr Leu Asp Pro Ile Pro
625                 630                 635                 640

Asn Leu Val Asn Pro His Thr Gly Arg Leu His Thr Thr Phe Ala Gln
                645                 650                 655

Thr Ala Val Ala Thr Gly Arg Leu Ser Ser Leu Asn Pro Asn Leu Gln
            660                 665                 670

Asn Ile Pro Ile Arg Ser Glu Leu Gly Arg Glu Ile Arg Lys Gly Phe
        675                 680                 685

Ile Ala Glu Asp Gly Phe Thr Leu Ile Ala Ala Asp Tyr Ser Gln Ile
    690                 695                 700

Glu Leu Arg Leu Leu Ala His Ile Ala Asp Asp Pro Leu Met Gln Gln
705                 710                 715                 720

Ala Phe Val Glu Gly Ala Asp Ile His Arg Arg Thr Ala Ala Gln Val
                725                 730                 735
```

-continued

```
Leu Gly Leu Asp Glu Ala Thr Val Asp Ala Asn Gln Arg Arg Ala Ala
            740                 745                 750

Lys Thr Val Asn Phe Gly Val Leu Tyr Gly Met Ser Ala His Arg Leu
        755                 760                 765

Ser Asn Asp Leu Gly Ile Pro Tyr Ala Glu Ala Ala Thr Phe Ile Glu
    770                 775                 780

Ile Tyr Phe Ala Thr Tyr Pro Gly Ile Arg Arg Tyr Ile Asn His Thr
785                 790                 795                 800

Leu Asp Phe Gly Arg Thr His Gly Tyr Val Glu Thr Leu Tyr Gly Arg
                805                 810                 815

Arg Arg Tyr Val Pro Gly Leu Ser Ser Arg Asn Arg Val Gln Arg Glu
            820                 825                 830

Ala Glu Glu Arg Leu Ala Tyr Asn Met Pro Ile Gln Gly Thr Ala Ala
        835                 840                 845

Asp Ile Met Lys Leu Ala Met Val Gln Leu Asp Pro Gln Leu Asp Ala
    850                 855                 860

Ile Gly Ala Arg Met Leu Leu Gln Val His Asp Glu Leu Leu Ile Glu
865                 870                 875                 880

Ala Pro Leu Asp Lys Ala Glu Gln Val Ala Ala Leu Thr Lys Lys Val
                885                 890                 895

Met Glu Asn Val Val Gln Leu Lys Val Pro Leu Ala Val Glu Val Gly
            900                 905                 910

Thr Gly Pro Asn Trp Phe Asp Thr Lys
        915                 920


<210> SEQ ID NO 33
<211> LENGTH: 892
<212> TYPE: PRT
<213> ORGANISM: Thermosipho africanus
<220> FEATURE:
<223> OTHER INFORMATION: Thermosipho africanus DNA polymerase (Taf)

<400> SEQUENCE: 33

Met Gly Lys Met Phe Leu Phe Asp Gly Thr Gly Leu Val Tyr Arg Ala
1               5                   10                  15

Phe Tyr Ala Ile Asp Gln Ser Leu Gln Thr Ser Ser Gly Leu His Thr
            20                  25                  30

Asn Ala Val Tyr Gly Leu Thr Lys Met Leu Ile Lys Phe Leu Lys Glu
        35                  40                  45

His Ile Ser Ile Gly Lys Asp Ala Cys Val Phe Val Leu Asp Ser Lys
    50                  55                  60

Gly Gly Ser Lys Lys Arg Lys Asp Ile Leu Glu Thr Tyr Lys Ala Asn
65                  70                  75                  80

Arg Pro Ser Thr Pro Asp Leu Leu Leu Glu Gln Ile Pro Tyr Val Glu
                85                  90                  95

Glu Leu Val Asp Ala Leu Gly Ile Lys Val Leu Lys Ile Glu Gly Phe
            100                 105                 110

Glu Ala Asp Asp Ile Ile Ala Thr Leu Ser Lys Lys Phe Glu Ser Asp
        115                 120                 125

Phe Glu Lys Val Asn Ile Ile Thr Gly Asp Lys Asp Leu Leu Gln Leu
    130                 135                 140

Val Ser Asp Lys Val Phe Val Trp Arg Val Glu Arg Gly Ile Thr Asp
145                 150                 155                 160
```

-continued

```
Leu Val Leu Tyr Asp Arg Asn Lys Val Ile Glu Lys Tyr Gly Ile Tyr
                165                 170                 175

Pro Glu Gln Phe Lys Asp Tyr Leu Ser Leu Val Gly Asp Gln Ile Asp
            180                 185                 190

Asn Ile Pro Gly Val Lys Gly Ile Gly Lys Lys Thr Ala Val Ser Leu
        195                 200                 205

Leu Lys Lys Tyr Asn Ser Leu Glu Asn Val Leu Lys Asn Ile Asn Leu
    210                 215                 220

Leu Thr Glu Lys Leu Arg Arg Leu Leu Glu Asp Ser Lys Glu Asp Leu
225                 230                 235                 240

Gln Lys Ser Ile Glu Leu Val Glu Leu Ile Tyr Asp Val Pro Met Asp
                245                 250                 255

Val Glu Lys Asp Glu Ile Ile Tyr Arg Gly Tyr Asn Pro Asp Lys Leu
            260                 265                 270

Leu Lys Val Leu Lys Lys Tyr Glu Phe Ser Ser Ile Ile Lys Glu Leu
        275                 280                 285

Asn Leu Gln Glu Lys Leu Glu Lys Glu Tyr Ile Leu Val Asp Asn Glu
    290                 295                 300

Asp Lys Leu Lys Lys Leu Ala Glu Glu Ile Glu Lys Tyr Lys Thr Phe
305                 310                 315                 320

Ser Ile Asp Thr Glu Thr Thr Ser Leu Asp Pro Phe Glu Ala Lys Leu
                325                 330                 335

Val Gly Ile Ser Ile Ser Thr Met Glu Gly Lys Ala Tyr Tyr Ile Pro
            340                 345                 350

Val Ser His Phe Gly Ala Lys Asn Ile Ser Lys Ser Leu Ile Asp Lys
        355                 360                 365

Phe Leu Lys Gln Ile Leu Gln Glu Lys Asp Tyr Asn Ile Val Gly Gln
    370                 375                 380

Asn Leu Lys Phe Asp Tyr Glu Ile Phe Lys Ser Met Gly Phe Ser Pro
385                 390                 395                 400

Asn Val Pro His Phe Asp Thr Met Ile Ala Ala Tyr Leu Leu Asn Pro
                405                 410                 415

Asp Glu Lys Arg Phe Asn Leu Glu Glu Leu Ser Leu Lys Tyr Leu Gly
            420                 425                 430

Tyr Lys Met Ile Ser Phe Asp Glu Leu Val Asn Glu Asn Val Pro Leu
        435                 440                 445

Phe Gly Asn Asp Phe Ser Tyr Val Pro Leu Glu Arg Ala Val Glu Tyr
    450                 455                 460

Ser Cys Glu Asp Ala Asp Val Thr Tyr Arg Ile Phe Arg Lys Leu Gly
465                 470                 475                 480

Arg Lys Ile Tyr Glu Asn Glu Met Glu Lys Leu Phe Tyr Glu Ile Glu
                485                 490                 495

Met Pro Leu Ile Asp Val Leu Ser Glu Met Glu Leu Asn Gly Val Tyr
            500                 505                 510

Phe Asp Glu Glu Tyr Leu Lys Glu Leu Ser Lys Lys Tyr Gln Glu Lys
        515                 520                 525

Met Asp Gly Ile Lys Glu Lys Val Phe Glu Ile Ala Gly Glu Thr Phe
    530                 535                 540

Asn Leu Asn Ser Ser Thr Gln Val Ala Tyr Ile Leu Phe Glu Lys Leu
545                 550                 555                 560

Asn Ile Ala Pro Tyr Lys Lys Thr Ala Thr Gly Lys Phe Ser Thr Asn
```

-continued

```
                565                 570                 575

Ala Glu Val Leu Glu Glu Leu Ser Lys Glu His Glu Ile Ala Lys Leu
            580                 585                 590

Leu Leu Glu Tyr Arg Lys Tyr Gln Lys Leu Lys Ser Thr Tyr Ile Asp
        595                 600                 605

Ser Ile Pro Leu Ser Ile Asn Arg Lys Thr Asn Arg Val His Thr Thr
    610                 615                 620

Phe His Gln Thr Gly Thr Ser Thr Gly Arg Leu Ser Ser Ser Asn Pro
625                 630                 635                 640

Asn Leu Gln Asn Leu Pro Thr Arg Ser Glu Glu Gly Lys Glu Ile Arg
                645                 650                 655

Lys Ala Val Arg Pro Gln Arg Gln Asp Trp Trp Ile Leu Gly Ala Asp
            660                 665                 670

Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala His Val Ser Lys Asp Glu
        675                 680                 685

Asn Leu Leu Lys Ala Phe Lys Glu Asp Leu Asp Ile His Thr Ile Thr
    690                 695                 700

Ala Ala Lys Ile Phe Gly Val Ser Glu Met Phe Val Ser Glu Gln Met
705                 710                 715                 720

Arg Arg Val Gly Lys Met Val Asn Phe Ala Ile Ile Tyr Gly Val Ser
                725                 730                 735

Pro Tyr Gly Leu Ser Lys Arg Ile Gly Leu Ser Val Ser Glu Thr Lys
            740                 745                 750

Lys Ile Ile Asp Asn Tyr Phe Arg Tyr Tyr Lys Gly Val Phe Glu Tyr
        755                 760                 765

Leu Lys Arg Met Lys Asp Glu Ala Arg Lys Lys Gly Tyr Val Thr Thr
    770                 775                 780

Leu Phe Gly Arg Arg Arg Tyr Ile Pro Gln Leu Arg Ser Lys Asn Gly
785                 790                 795                 800

Asn Arg Val Gln Glu Gly Glu Arg Ile Ala Val Asn Thr Pro Ile Gln
                805                 810                 815

Gly Thr Ala Ala Asp Ile Ile Lys Ile Ala Met Ile Asn Ile His Asn
            820                 825                 830

Arg Leu Lys Lys Glu Asn Leu Arg Ser Lys Met Ile Leu Gln Val His
        835                 840                 845

Asp Glu Leu Val Phe Glu Val Pro Asp Asn Glu Leu Glu Ile Val Lys
    850                 855                 860

Asp Leu Val Arg Asp Glu Met Glu Asn Ala Val Lys Leu Asp Val Pro
865                 870                 875                 880

Leu Lys Val Asp Val Tyr Tyr Gly Lys Glu Trp Glu
                885                 890


<210> SEQ ID NO 34
<211> LENGTH: 893
<212> TYPE: PRT
<213> ORGANISM: Thermotoga maritima
<220> FEATURE:
<223> OTHER INFORMATION: Thermotoga maritima DNA polymerase (Tma)

<400> SEQUENCE: 34

Met Ala Arg Leu Phe Leu Phe Asp Gly Thr Ala Leu Ala Tyr Arg Ala
1               5                   10                  15

Tyr Tyr Ala Leu Asp Arg Ser Leu Ser Thr Ser Thr Gly Ile Pro Thr
```

-continued

```
            20                  25                  30

Asn Ala Thr Tyr Gly Val Ala Arg Met Leu Val Arg Phe Ile Lys Asp
        35                  40                  45

His Ile Ile Val Gly Lys Asp Tyr Val Ala Val Ala Phe Asp Lys Lys
    50                  55                  60

Ala Ala Thr Phe Arg His Lys Leu Leu Glu Thr Tyr Lys Ala Gln Arg
65                  70                  75                  80

Pro Lys Thr Pro Asp Leu Leu Ile Gln Gln Leu Pro Tyr Ile Lys Lys
                85                  90                  95

Leu Val Glu Ala Leu Gly Met Lys Val Leu Glu Val Glu Gly Tyr Glu
            100                 105                 110

Ala Asp Asp Ile Ile Ala Thr Leu Ala Val Lys Gly Leu Pro Leu Phe
        115                 120                 125

Asp Glu Ile Phe Ile Val Thr Gly Asp Lys Asp Met Leu Gln Leu Val
    130                 135                 140

Asn Glu Lys Ile Lys Val Trp Arg Ile Val Lys Gly Ile Ser Asp Leu
145                 150                 155                 160

Glu Leu Tyr Asp Ala Gln Lys Val Lys Glu Lys Tyr Gly Val Glu Pro
                165                 170                 175

Gln Gln Ile Pro Asp Leu Leu Ala Leu Thr Gly Asp Glu Ile Asp Asn
            180                 185                 190

Ile Pro Gly Val Thr Gly Ile Gly Glu Lys Thr Ala Val Gln Leu Leu
        195                 200                 205

Glu Lys Tyr Lys Asp Leu Glu Asp Ile Leu Asn His Val Arg Glu Leu
    210                 215                 220

Pro Gln Lys Val Arg Lys Ala Leu Leu Arg Asp Arg Glu Asn Ala Ile
225                 230                 235                 240

Leu Ser Lys Lys Leu Ala Ile Leu Glu Thr Asn Val Pro Ile Glu Ile
                245                 250                 255

Asn Trp Glu Glu Leu Arg Tyr Gln Gly Tyr Asp Arg Glu Lys Leu Leu
            260                 265                 270

Pro Leu Leu Lys Glu Leu Glu Phe Ala Ser Ile Met Lys Glu Leu Gln
        275                 280                 285

Leu Tyr Glu Glu Ser Glu Pro Val Gly Tyr Arg Ile Val Lys Asp Leu
    290                 295                 300

Val Glu Phe Glu Lys Leu Ile Glu Lys Leu Arg Glu Ser Pro Ser Phe
305                 310                 315                 320

Ala Ile Asp Leu Glu Thr Ser Ser Leu Asp Pro Phe Asp Cys Asp Ile
                325                 330                 335

Val Gly Ile Ser Val Ser Phe Lys Pro Lys Glu Ala Tyr Tyr Ile Pro
            340                 345                 350

Leu His His Arg Asn Ala Gln Asn Leu Asp Glu Lys Glu Val Leu Lys
        355                 360                 365

Lys Leu Lys Glu Ile Leu Glu Asp Pro Gly Ala Lys Ile Val Gly Gln
    370                 375                 380

Asn Leu Lys Phe Asp Tyr Lys Val Leu Met Val Lys Gly Val Glu Pro
385                 390                 395                 400

Val Pro Pro Tyr Phe Asp Thr Met Ile Ala Ala Tyr Leu Leu Glu Pro
                405                 410                 415

Asn Glu Lys Lys Phe Asn Leu Asp Asp Leu Ala Leu Lys Phe Leu Gly
            420                 425                 430
```

-continued

```
Tyr Lys Met Thr Ser Tyr Gln Glu Leu Met Ser Phe Ser Phe Pro Leu
        435                 440                 445

Phe Gly Phe Ser Phe Ala Asp Val Pro Val Glu Lys Ala Ala Asn Tyr
    450                 455                 460

Ser Cys Glu Asp Ala Asp Ile Thr Tyr Arg Leu Tyr Lys Thr Leu Ser
465                 470                 475                 480

Leu Lys Leu His Glu Ala Asp Leu Glu Asn Val Phe Tyr Lys Ile Glu
                485                 490                 495

Met Pro Leu Val Asn Val Leu Ala Arg Met Glu Leu Asn Gly Val Tyr
            500                 505                 510

Val Asp Thr Glu Phe Leu Lys Lys Leu Ser Glu Glu Tyr Gly Lys Lys
        515                 520                 525

Leu Glu Glu Leu Ala Glu Glu Ile Tyr Arg Ile Ala Gly Glu Pro Phe
    530                 535                 540

Asn Ile Asn Ser Pro Lys Gln Val Ser Arg Ile Leu Phe Glu Lys Leu
545                 550                 555                 560

Gly Ile Lys Pro Arg Gly Lys Thr Thr Lys Thr Gly Asp Tyr Ser Thr
                565                 570                 575

Arg Ile Glu Val Leu Glu Glu Leu Ala Gly Glu His Glu Ile Ile Pro
            580                 585                 590

Leu Ile Leu Glu Tyr Arg Lys Ile Gln Lys Leu Lys Ser Thr Tyr Ile
        595                 600                 605

Asp Ala Leu Pro Lys Met Val Asn Pro Lys Thr Gly Arg Ile His Ala
    610                 615                 620

Ser Phe Asn Gln Thr Gly Thr Ala Thr Gly Arg Leu Ser Ser Ser Asp
625                 630                 635                 640

Pro Asn Leu Gln Asn Leu Pro Thr Lys Ser Glu Glu Gly Lys Glu Ile
                645                 650                 655

Arg Lys Ala Ile Val Pro Gln Asp Pro Asn Trp Trp Ile Val Ser Ala
            660                 665                 670

Asp Tyr Ser Gln Ile Glu Leu Arg Ile Leu Ala His Leu Ser Gly Asp
        675                 680                 685

Glu Asn Leu Leu Arg Ala Phe Glu Glu Gly Ile Asp Val His Thr Leu
    690                 695                 700

Thr Ala Ser Arg Ile Phe Asn Val Lys Pro Glu Glu Val Thr Glu Glu
705                 710                 715                 720

Met Arg Arg Ala Gly Lys Met Val Asn Phe Ser Ile Ile Tyr Gly Val
                725                 730                 735

Thr Pro Tyr Gly Leu Ser Val Arg Leu Gly Val Pro Val Lys Glu Ala
            740                 745                 750

Glu Lys Met Ile Val Asn Tyr Phe Val Leu Tyr Pro Lys Val Arg Asp
        755                 760                 765

Tyr Ile Gln Arg Val Val Ser Glu Ala Lys Glu Lys Gly Tyr Val Arg
    770                 775                 780

Thr Leu Phe Gly Arg Lys Arg Asp Ile Pro Gln Leu Met Ala Arg Asp
785                 790                 795                 800

Arg Asn Thr Gln Ala Glu Gly Glu Arg Ile Ala Ile Asn Thr Pro Ile
                805                 810                 815

Gln Gly Thr Ala Ala Asp Ile Ile Lys Leu Ala Met Ile Glu Ile Asp
            820                 825                 830
```

-continued

```
Arg Glu Leu Lys Glu Arg Lys Met Arg Ser Lys Met Ile Ile Gln Val
        835                 840                 845

His Asp Glu Leu Val Phe Glu Val Pro Asn Glu Glu Lys Asp Ala Leu
    850                 855                 860

Val Glu Leu Val Lys Asp Arg Met Thr Asn Val Val Lys Leu Ser Val
865                 870                 875                 880

Pro Leu Glu Val Asp Val Thr Ile Gly Lys Thr Trp Ser
                885                 890


<210> SEQ ID NO 35
<211> LENGTH: 893
<212> TYPE: PRT
<213> ORGANISM: Thermotoga neopolitana
<220> FEATURE:
<223> OTHER INFORMATION: Thermotoga neopolitana DNA polymerase (Tne)

<400> SEQUENCE: 35

Met Ala Arg Leu Phe Leu Phe Asp Gly Thr Ala Leu Ala Tyr Arg Ala
1               5                   10                  15

Tyr Tyr Ala Leu Asp Arg Ser Leu Ser Thr Ser Thr Gly Ile Pro Thr
            20                  25                  30

Asn Ala Val Tyr Gly Val Ala Arg Met Leu Val Lys Phe Ile Lys Glu
        35                  40                  45

His Ile Ile Pro Glu Lys Asp Tyr Ala Ala Val Ala Phe Asp Lys Lys
    50                  55                  60

Ala Ala Thr Phe Arg His Lys Leu Leu Val Ser Asp Lys Ala Gln Arg
65                  70                  75                  80

Pro Lys Thr Pro Ala Leu Leu Val Gln Gln Leu Pro Tyr Ile Lys Arg
                85                  90                  95

Leu Ile Glu Ala Leu Gly Phe Lys Val Leu Glu Leu Glu Gly Tyr Glu
            100                 105                 110

Ala Asp Asp Ile Ile Ala Thr Leu Ala Val Arg Ala Ala Arg Phe Leu
        115                 120                 125

Met Arg Phe Ser Leu Ile Thr Gly Asp Lys Asp Met Leu Gln Leu Val
    130                 135                 140

Asn Glu Lys Ile Lys Val Trp Arg Ile Val Lys Gly Ile Ser Asp Leu
145                 150                 155                 160

Glu Leu Tyr Asp Ser Lys Lys Val Lys Glu Arg Tyr Gly Val Glu Pro
                165                 170                 175

His Gln Ile Pro Asp Leu Leu Ala Leu Thr Gly Asp Asp Ile Asp Asn
            180                 185                 190

Ile Pro Gly Val Thr Gly Ile Gly Glu Lys Thr Ala Val Gln Leu Leu
        195                 200                 205

Gly Lys Tyr Arg Asn Leu Glu Tyr Ile Leu Glu His Ala Arg Glu Leu
    210                 215                 220

Pro Gln Arg Val Arg Lys Ala Leu Leu Arg Asp Arg Glu Val Ala Ile
225                 230                 235                 240

Leu Ser Lys Lys Leu Ala Thr Leu Val Thr Asn Ala Pro Val Glu Val
                245                 250                 255

Asp Trp Glu Glu Met Lys Tyr Arg Gly Tyr Asp Lys Arg Lys Leu Leu
            260                 265                 270

Pro Ile Leu Lys Glu Leu Glu Phe Ala Ser Ile Met Lys Glu Leu Gln
        275                 280                 285
```

-continued

```
Leu Tyr Glu Glu Ala Glu Pro Thr Gly Tyr Glu Ile Val Lys Asp His
    290                 295                 300

Lys Thr Phe Glu Asp Leu Ile Glu Lys Leu Lys Glu Val Pro Ser Phe
305                 310                 315                 320

Ala Leu Asp Leu Glu Thr Ser Ser Leu Asp Pro Phe Asn Cys Glu Ile
                325                 330                 335

Val Gly Ile Ser Val Ser Phe Lys Pro Lys Thr Ala Tyr Tyr Ile Pro
            340                 345                 350

Leu His His Arg Asn Ala His Asn Leu Asp Glu Thr Leu Val Leu Ser
        355                 360                 365

Lys Leu Lys Glu Ile Leu Glu Asp Pro Ser Ser Lys Ile Val Gly Gln
    370                 375                 380

Asn Leu Lys Tyr Asp Tyr Lys Val Leu Met Val Lys Gly Ile Ser Pro
385                 390                 395                 400

Val Tyr Pro His Phe Asp Thr Met Ile Ala Ala Tyr Leu Leu Glu Pro
                405                 410                 415

Asn Glu Lys Lys Phe Asn Leu Glu Asp Leu Ser Leu Lys Phe Leu Gly
            420                 425                 430

Tyr Lys Met Thr Ser Tyr Gln Glu Leu Met Ser Phe Ser Ser Pro Leu
        435                 440                 445

Phe Gly Phe Ser Phe Ala Asp Val Pro Val Asp Lys Ala Ala Glu Tyr
    450                 455                 460

Ser Cys Glu Asp Ala Asp Ile Thr Tyr Arg Leu Tyr Lys Ile Leu Ser
465                 470                 475                 480

Met Lys Leu His Glu Ala Glu Leu Glu Asn Val Phe Tyr Arg Ile Glu
                485                 490                 495

Met Pro Leu Val Asn Val Leu Ala Arg Met Glu Phe Asn Trp Val Tyr
            500                 505                 510

Val Asp Thr Glu Phe Leu Lys Lys Leu Ser Glu Glu Tyr Gly Lys Lys
        515                 520                 525

Leu Glu Glu Leu Ala Glu Lys Ile Tyr Gln Ile Ala Gly Glu Pro Phe
    530                 535                 540

Asn Ile Asn Ser Pro Lys Gln Val Ser Asn Ile Leu Phe Glu Lys Leu
545                 550                 555                 560

Gly Ile Lys Pro Arg Gly Lys Thr Thr Lys Thr Gly Asp Tyr Ser Thr
                565                 570                 575

Arg Ile Glu Val Leu Glu Glu Ile Ala Asn Glu His Glu Ile Val Pro
            580                 585                 590

Leu Ile Leu Glu Phe Arg Lys Ile Leu Lys Leu Lys Ser Thr Tyr Ile
        595                 600                 605

Asp Thr Leu Pro Lys Leu Val Asn Pro Lys Thr Gly Arg Phe His Ala
    610                 615                 620

Ser Phe His Gln Thr Gly Thr Ala Thr Gly Arg Leu Ser Ser Ser Asp
625                 630                 635                 640

Pro Asn Leu Gln Asn Leu Pro Thr Lys Ser Glu Glu Gly Lys Glu Ile
                645                 650                 655

Arg Lys Ala Ile Val Pro Gln Asp Pro Asp Trp Trp Ile Val Ser Ala
            660                 665                 670

Asp Tyr Ser Gln Ile Glu Leu Arg Ile Leu Ala His Leu Ser Gly Asp
        675                 680                 685

Glu Asn Leu Val Lys Ala Phe Glu Glu Gly Ile Asp Val His Thr Leu
```

-continued

```
    690                 695                 700

Thr Ala Ser Arg Ile Tyr Asn Val Lys Pro Glu Glu Val Asn Glu Glu
705                 710                 715                 720

Met Arg Arg Val Gly Lys Met Val Asn Phe Ser Ile Ile Tyr Gly Val
                725                 730                 735

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

Glu Lys Met Ile Ile Ser Tyr Phe Thr Leu Tyr Pro Lys Val Arg Ser
        755                 760                 765

Tyr Ile Gln Gln Val Val Ala Glu Ala Lys Glu Lys Gly Tyr Val Arg
    770                 775                 780

Thr Leu Phe Gly Arg Lys Arg Asp Ile Pro Gln Leu Met Ala Arg Asp
785                 790                 795                 800

Lys Asn Thr Gln Ser Glu Gly Glu Arg Ile Ala Ile Asn Thr Pro Ile
                805                 810                 815

Gln Gly Thr Ala Ala Asp Ile Ile Lys Leu Ala Met Ile Asp Ile Asp
            820                 825                 830

Glu Glu Leu Arg Lys Arg Asn Met Lys Ser Arg Met Ile Ile Gln Val
        835                 840                 845

His Asp Glu Leu Val Phe Glu Val Pro Asp Glu Glu Lys Glu Glu Leu
    850                 855                 860

Val Asp Leu Val Lys Asn Lys Met Thr Asn Val Val Lys Leu Ser Val
865                 870                 875                 880

Pro Leu Glu Val Asp Ile Ser Ile Gly Lys Ser Trp Ser
                885                 890


<210> SEQ ID NO 36
<211> LENGTH: 876
<212> TYPE: PRT
<213> ORGANISM: Bacillus stearothermophilus
<220> FEATURE:
<223> OTHER INFORMATION: Bacillus stearothermophilus DNA polymerase
      (Bst)

<400> SEQUENCE: 36

Met Lys Asn Lys Leu Val Leu Ile Asp Gly Asn Ser Val Ala Tyr Arg
1               5                   10                  15

Ala Phe Phe Ala Leu Pro Leu Leu His Asn Asp Lys Gly Ile His Thr
            20                  25                  30

Asn Ala Val Tyr Gly Phe Thr Met Met Leu Asn Lys Ile Leu Ala Glu
        35                  40                  45

Glu Gln Pro Thr His Ile Leu Val Ala Phe Asp Ala Gly Lys Thr Thr
    50                  55                  60

Phe Arg His Glu Thr Phe Gln Asp Tyr Lys Gly Gly Arg Gln Gln Thr
65                  70                  75                  80

Pro Pro Glu Leu Ser Glu Gln Phe Pro Leu Leu Arg Glu Leu Leu Lys
                85                  90                  95

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

Ile Ile Gly Thr Met Ala Ala Arg Ala Glu Arg Glu Gly Phe Ala Val
        115                 120                 125

Lys Val Ile Ser Gly Asp Arg Asp Leu Thr Gln Leu Ala Ser Pro Gln
    130                 135                 140
```

-continued

```
Val Thr Val Glu Ile Thr Lys Lys Gly Ile Thr Asp Ile Glu Ser Tyr
145                 150                 155                 160

Thr Pro Glu Thr Val Val Glu Lys Tyr Gly Leu Thr Pro Glu Gln Ile
                165                 170                 175

Val Asp Leu Lys Gly Leu Met Gly Asp Lys Ser Asp Asn Ile Pro Gly
            180                 185                 190

Val Pro Gly Ile Gly Glu Lys Thr Ala Val Lys Leu Leu Lys Gln Phe
        195                 200                 205

Gly Thr Val Glu Asn Val Leu Ala Ser Ile Asp Glu Ile Lys Gly Glu
    210                 215                 220

Lys Leu Lys Glu Asn Leu Arg Gln Tyr Arg Asp Leu Ala Leu Leu Ser
225                 230                 235                 240

Lys Gln Leu Ala Ala Ile Cys Arg Asp Ala Pro Val Glu Leu Thr Leu
                245                 250                 255

Asp Asp Ile Val Tyr Lys Gly Glu Asp Arg Glu Lys Val Val Ala Leu
            260                 265                 270

Phe Gln Glu Leu Gly Phe Gln Ser Phe Leu Asp Lys Met Ala Val Gln
        275                 280                 285

Thr Asp Glu Gly Glu Lys Pro Leu Ala Gly Met Asp Phe Ala Ile Ala
    290                 295                 300

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

Glu Val Val Gly Asp Asn Tyr His His Ala Pro Ile Val Gly Ile Ala
                325                 330                 335

Leu Ala Asn Glu Arg Gly Arg Phe Phe Leu Arg Pro Glu Thr Ala Leu
            340                 345                 350

Ala Asp Pro Lys Phe Leu Ala Trp Leu Gly Asp Glu Thr Lys Lys Lys
        355                 360                 365

Thr Met Phe Asp Ser Lys Arg Ala Ala Val Ala Leu Lys Trp Lys Gly
    370                 375                 380

Ile Glu Leu Arg Gly Val Val Phe Asp Leu Leu Leu Ala Ala Tyr Leu
385                 390                 395                 400

Leu Asp Pro Ala Gln Ala Ala Gly Asp Val Ala Ala Val Ala Lys Met
                405                 410                 415

His Gln Tyr Glu Ala Val Arg Ser Asp Glu Ala Val Tyr Gly Lys Gly
            420                 425                 430

Ala Lys Arg Thr Val Pro Asp Glu Pro Thr Leu Ala Glu His Leu Ala
        435                 440                 445

Arg Lys Ala Ala Ala Ile Trp Ala Leu Glu Glu Pro Leu Met Asp Glu
    450                 455                 460

Leu Arg Arg Asn Glu Gln Asp Arg Leu Leu Thr Glu Leu Glu Gln Pro
465                 470                 475                 480

Leu Ala Gly Ile Leu Ala Asn Met Glu Phe Thr Gly Val Lys Val Asp
                485                 490                 495

Thr Lys Arg Leu Glu Gln Met Gly Ala Glu Leu Thr Glu Gln Leu Gln
            500                 505                 510

Ala Val Glu Arg Arg Ile Tyr Glu Leu Ala Gly Gln Glu Phe Asn Ile
        515                 520                 525

Asn Ser Pro Lys Gln Leu Gly Thr Val Leu Phe Asp Lys Leu Gln Leu
    530                 535                 540

Pro Val Leu Lys Lys Thr Lys Thr Gly Tyr Ser Thr Ser Ala Asp Val
```

-continued

```
545                 550                 555                 560

Leu Glu Lys Leu Ala Pro His His Glu Ile Val Glu His Ile Leu His
                565                 570                 575

Tyr Arg Gln Leu Gly Lys Leu Gln Ser Thr Tyr Ile Glu Gly Leu Leu
            580                 585                 590

Lys Val Val His Pro Val Thr Gly Lys Val His Thr Met Phe Asn Gln
        595                 600                 605

Ala Leu Thr Gln Thr Gly Arg Leu Ser Ser Val Glu Pro Asn Leu Gln
    610                 615                 620

Asn Ile Pro Ile Arg Leu Glu Glu Gly Arg Lys Ile Arg Gln Ala Phe
625                 630                 635                 640

Val Pro Ser Glu Pro Asp Trp Leu Ile Phe Ala Ala Asp Tyr Ser Gln
                645                 650                 655

Ile Glu Leu Arg Val Leu Ala His Ile Ala Glu Asp Asp Asn Leu Ile
            660                 665                 670

Glu Ala Phe Arg Arg Gly Leu Asp Ile His Thr Lys Thr Ala Met Asp
        675                 680                 685

Ile Phe His Val Ser Glu Glu Asp Val Thr Ala Asn Met Arg Arg Gln
    690                 695                 700

Ala Lys Ala Val Asn Phe Gly Ile Val Tyr Gly Ile Ser Asp Tyr Gly
705                 710                 715                 720

Leu Ala Gln Asn Leu Asn Ile Thr Arg Lys Glu Ala Ala Glu Phe Ile
                725                 730                 735

Glu Arg Tyr Phe Ala Ser Phe Pro Gly Val Lys Gln Tyr Met Asp Asn
            740                 745                 750

Ile Val Gln Glu Ala Lys Gln Lys Gly Tyr Val Thr Thr Leu Leu His
        755                 760                 765

Arg Arg Arg Tyr Leu Pro Asp Ile Thr Ser Arg Asn Phe Asn Val Arg
    770                 775                 780

Ser Phe Ala Glu Arg Thr Ala Met Asn Thr Pro Ile Gln Gly Ser Ala
785                 790                 795                 800

Ala Asp Ile Ile Lys Lys Ala Met Ile Asp Leu Ser Val Arg Leu Arg
                805                 810                 815

Glu Glu Arg Leu Gln Ala Arg Leu Leu Leu Gln Val His Asp Glu Leu
            820                 825                 830

Ile Leu Glu Ala Pro Lys Glu Glu Ile Glu Arg Leu Cys Arg Leu Val
        835                 840                 845

Pro Glu Val Met Glu Gln Ala Val Ala Leu Arg Val Pro Leu Lys Val
    850                 855                 860

Asp Tyr His Tyr Gly Pro Thr Trp Tyr Asp Ala Lys
865                 870                 875


<210> SEQ ID NO 37
<211> LENGTH: 877
<212> TYPE: PRT
<213> ORGANISM: Bacillus caldotenax
<220> FEATURE:
<223> OTHER INFORMATION: Bacillus caldotenax DNA polymerase (Bca)

<400> SEQUENCE: 37

Met Lys Lys Lys Leu Val Leu Ile Asp Gly Ser Ser Val Ala Tyr Arg
1               5                   10                  15

Ala Phe Phe Ala Leu Pro Leu Leu His Asn Asp Lys Gly Ile His Thr
```

-continued

```
            20                  25                  30

Asn Ala Val Tyr Gly Phe Thr Met Met Leu Asn Lys Ile Leu Ala Glu
        35                  40                  45

Glu Glu Pro Thr His Met Leu Val Ala Phe Asp Ala Gly Lys Thr Thr
    50                  55                  60

Phe Arg His Glu Ala Phe Gln Glu Tyr Lys Gly Gly Arg Gln Gln Thr
65                  70                  75                  80

Pro Pro Glu Leu Ser Glu Gln Phe Pro Leu Leu Arg Glu Leu Leu Arg
                85                  90                  95

Ala Tyr Arg Ile Pro Ala Tyr Glu Leu Glu Asn Tyr Glu Ala Asp Asp
            100                 105                 110

Ile Ile Gly Thr Leu Ala Ala Arg Ala Glu Gln Glu Gly Phe Glu Val
        115                 120                 125

Lys Val Ile Ser Gly Asp Arg Asp Leu Thr Gln Leu Ala Ser Pro His
    130                 135                 140

Val Thr Val Asp Ile Thr Lys Lys Gly Ile Thr Asp Ile Glu Pro Tyr
145                 150                 155                 160

Thr Pro Glu Ala Val Arg Glu Lys Tyr Gly Leu Thr Pro Glu Gln Ile
                165                 170                 175

Val Asp Leu Lys Gly Leu Met Gly Asp Lys Ser Asp Asn Ile Pro Gly
            180                 185                 190

Val Pro Gly Ile Gly Glu Lys Thr Ala Val Lys Leu Leu Arg Gln Phe
        195                 200                 205

Gly Thr Val Glu Asn Val Leu Ala Ser Ile Asp Glu Ile Lys Gly Glu
    210                 215                 220

Lys Leu Lys Glu Thr Leu Arg Gln His Arg Glu Met Ala Leu Leu Ser
225                 230                 235                 240

Lys Lys Leu Ala Ala Ile Arg Arg Asp Ala Pro Val Glu Leu Ser Leu
                245                 250                 255

Asp Asp Ile Ala Tyr Gln Gly Glu Asp Arg Glu Lys Val Val Ala Leu
            260                 265                 270

Phe Lys Glu Leu Gly Phe Gln Ser Phe Leu Glu Lys Met Glu Ser Pro
        275                 280                 285

Ser Ser Glu Glu Glu Lys Pro Leu Ala Lys Met Ala Phe Thr Leu Ala
    290                 295                 300

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

Glu Val Val Glu Glu Asn Tyr His Asp Ala Pro Ile Val Gly Ile Ala
                325                 330                 335

Val Val Asn Glu His Gly Arg Phe Phe Leu Arg Pro Glu Thr Ala Leu
            340                 345                 350

Ala Asp Pro Gln Phe Val Ala Trp Leu Gly Asp Glu Thr Lys Lys Lys
        355                 360                 365

Ser Met Phe Asp Ser Lys Arg Ala Ala Val Ala Leu Lys Trp Lys Gly
    370                 375                 380

Ile Glu Leu Cys Gly Val Ser Phe Asp Leu Leu Leu Ala Ala Tyr Leu
385                 390                 395                 400

Leu Asp Pro Ala Gln Gly Val Asp Asp Val Ala Ala Ala Ala Lys Met
                405                 410                 415

Lys Gln Tyr Glu Ala Val Arg Pro Asp Glu Ala Val Tyr Gly Lys Gly
            420                 425                 430
```

-continued

```
Ala Lys Arg Ala Val Pro Asp Glu Pro Val Leu Ala Glu His Leu Val
        435                 440                 445

Arg Lys Ala Ala Ala Ile Trp Ala Leu Glu Arg Pro Phe Leu Asp Glu
    450                 455                 460

Leu Arg Arg Asn Glu Gln Asp Arg Leu Leu Val Glu Leu Glu Gln Pro
465                 470                 475                 480

Leu Ser Ser Ile Leu Ala Glu Met Glu Phe Ala Gly Val Lys Val Asp
                485                 490                 495

Thr Lys Arg Leu Glu Gln Met Gly Glu Glu Leu Ala Glu Gln Leu Arg
            500                 505                 510

Thr Val Glu Gln Arg Ile Tyr Glu Leu Ala Gly Gln Glu Phe Asn Ile
        515                 520                 525

Asn Ser Pro Lys Gln Leu Gly Val Ile Leu Phe Glu Lys Leu Gln Leu
    530                 535                 540

Pro Val Leu Lys Lys Ser Lys Thr Gly Tyr Ser Thr Ser Ala Asp Val
545                 550                 555                 560

Leu Glu Lys Leu Ala Pro Tyr His Glu Ile Val Glu Asn Ile Leu Gln
                565                 570                 575

His Tyr Arg Gln Leu Gly Lys Leu Gln Ser Thr Tyr Ile Glu Gly Leu
            580                 585                 590

Leu Lys Val Val Arg Pro Asp Thr Lys Lys Val His Thr Ile Phe Asn
        595                 600                 605

Gln Ala Leu Thr Gln Thr Gly Arg Leu Ser Ser Thr Glu Pro Asn Leu
    610                 615                 620

Gln Asn Ile Pro Ile Arg Leu Glu Glu Gly Arg Lys Ile Arg Gln Ala
625                 630                 635                 640

Phe Val Pro Ser Glu Ser Asp Trp Leu Ile Phe Ala Ala Asp Tyr Ser
                645                 650                 655

Gln Ile Glu Leu Arg Val Leu Ala His Ile Ala Glu Asp Asp Asn Leu
            660                 665                 670

Met Glu Ala Phe Arg Arg Asp Leu Asp Ile His Thr Lys Thr Ala Met
        675                 680                 685

Asp Ile Phe Gln Val Ser Glu Asp Glu Val Thr Pro Asn Met Arg Arg
    690                 695                 700

Gln Ala Lys Ala Val Asn Phe Gly Ile Val Tyr Gly Ile Ser Asp Tyr
705                 710                 715                 720

Gly Leu Ala Gln Asn Leu Asn Ile Ser Arg Lys Glu Ala Ala Glu Phe
                725                 730                 735

Ile Glu Arg Tyr Phe Glu Ser Phe Pro Gly Val Lys Arg Tyr Met Glu
            740                 745                 750

Asn Ile Val Gln Glu Ala Lys Gln Lys Gly Tyr Val Thr Thr Leu Leu
        755                 760                 765

His Arg Arg Arg Tyr Leu Pro Asp Ile Thr Ser Arg Asn Phe Asn Val
    770                 775                 780

Arg Ser Phe Ala Glu Arg Met Ala Met Asn Thr Pro Ile Gln Gly Ser
785                 790                 795                 800

Ala Ala Asp Ile Ile Lys Lys Ala Met Ile Asp Leu Asn Ala Arg Leu
                805                 810                 815

Lys Glu Glu Arg Leu Gln Ala Arg Leu Leu Leu Gln Val His Asp Glu
            820                 825                 830
```

-continued

```
Leu Ile Leu Glu Ala Pro Lys Glu Glu Met Glu Arg Leu Cys Arg Leu
        835                 840                 845

Val Pro Glu Val Met Glu Gln Ala Val Thr Leu Arg Val Pro Leu Lys
    850                 855                 860

Val Asp Tyr His Tyr Gly Ser Thr Trp Tyr Asp Ala Lys
865                 870                 875


<210> SEQ ID NO 38
<211> LENGTH: 13
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic polymerase motif corresponding to the
      D580X mutation of Z05
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (7)...(7)
<223> OTHER INFORMATION: Xaa = Ser or Thr
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (8)...(8)
<223> OTHER INFORMATION: Xaa = any amino acid other than Asp or Glu

<400> SEQUENCE: 38

Thr Gly Arg Leu Ser Ser Xaa Xaa Pro Asn Leu Gln Asn
1               5                   10


<210> SEQ ID NO 39
<211> LENGTH: 831
<212> TYPE: PRT
<213> ORGANISM: Carboxydothermus hydrogenoformans
<220> FEATURE:
<223> OTHER INFORMATION: Carboxydothermus hydrogenoformans DNA
      polymerase (Chy)

<400> SEQUENCE: 39

Met Gly Lys Val Val Leu Val Asp Gly Asn Ser Leu Leu His Arg Ala
1               5                   10                  15

Phe Phe Ala Leu Pro Pro Leu Lys Thr Thr Lys Gly Glu Pro Thr Gly
            20                  25                  30

Ala Val Tyr Glu Phe Leu Thr Met Leu Phe Arg Val Ile Lys Asp Glu
        35                  40                  45

Lys Pro Glu Tyr Leu Ala Val Ala Phe Asp Ile Ser Arg Lys Thr Phe
    50                  55                  60

Arg Thr Glu Gln Phe Thr Ala Tyr Lys Gly His Arg Lys Glu Ala Pro
65                  70                  75                  80

Asp Glu Leu Val Pro Gln Phe Ala Leu Val Arg Glu Val Leu Lys Val
                85                  90                  95

Leu Asn Val Pro Tyr Ile Glu Leu Asp Gly Tyr Glu Ala Asp Asp Ile
            100                 105                 110

Ile Gly His Leu Ser Arg Ala Phe Ala Gly Gln Gly His Glu Val Val
        115                 120                 125

Ile Tyr Thr Ala Asp Arg Asp Met Leu Gln Leu Val Asp Glu Lys Thr
    130                 135                 140

Val Val Tyr Leu Thr Lys Lys Gly Ile Thr Glu Leu Val Lys Met Asp
145                 150                 155                 160

Leu Ala Ala Ile Leu Glu Asn Tyr Gly Leu Lys Pro Lys Gln Leu Val
                165                 170                 175

Asp Val Lys Gly Leu Met Gly Asp Pro Ser Asp Asn Ile Pro Gly Val
```

-continued

```
            180                 185                 190

Pro Gly Ile Gly Glu Lys Thr Ala Leu Asp Leu Ile Lys Thr Tyr Gly
        195                 200                 205

Ser Val Glu Glu Val Leu Ala Arg Lys Asp Glu Leu Lys Pro Lys Leu
    210                 215                 220

Arg Glu Lys Leu Ala Glu His Glu Asn Leu Ala Lys Ile Ser Lys Gln
225                 230                 235                 240

Leu Ala Thr Ile Leu Arg Glu Ile Pro Leu Glu Ile Ser Leu Glu Asp
                245                 250                 255

Leu Lys Val Lys Glu Pro Asn Tyr Glu Glu Val Ala Lys Leu Phe Leu
            260                 265                 270

His Leu Glu Phe Lys Ser Phe Leu Lys Glu Ile Glu Pro Lys Ile Lys
        275                 280                 285

Lys Glu Tyr Gln Glu Gly Lys Asp Leu Val Gln Val Glu Thr Val Glu
    290                 295                 300

Thr Glu Gly Gln Ile Ala Val Val Phe Ser Asp Gly Phe Tyr Val Asp
305                 310                 315                 320

Asp Gly Glu Lys Thr Lys Phe Tyr Ser Leu Asp Arg Leu Asn Glu Ile
                325                 330                 335

Glu Glu Ile Phe Arg Asn Lys Lys Ile Ile Thr Asp Asp Ala Lys Gly
            340                 345                 350

Ile Tyr His Val Cys Leu Glu Lys Gly Leu Thr Phe Pro Glu Val Cys
        355                 360                 365

Phe Asp Ala Arg Ile Ala Ala Tyr Val Leu Asn Pro Ala Asp Gln Asn
    370                 375                 380

Pro Gly Leu Lys Gly Leu Tyr Leu Lys Tyr Asp Leu Pro Val Tyr Glu
385                 390                 395                 400

Asp Val Ser Leu Asn Ile Arg Gly Leu Phe Tyr Leu Lys Lys Glu Met
                405                 410                 415

Met Arg Lys Ile Phe Glu Gln Glu Gln Glu Arg Leu Phe Tyr Glu Ile
            420                 425                 430

Glu Leu Pro Leu Thr Pro Val Leu Ala Gln Met Glu His Thr Gly Ile
        435                 440                 445

Gln Val Asp Arg Glu Ala Leu Lys Glu Met Ser Leu Glu Leu Gly Glu
    450                 455                 460

Gln Ile Glu Glu Leu Ile Arg Glu Ile Tyr Val Leu Ala Gly Glu Glu
465                 470                 475                 480

Phe Asn Leu Asn Ser Pro Arg Gln Leu Gly Val Ile Leu Phe Glu Lys
                485                 490                 495

Leu Gly Leu Pro Val Ile Lys Lys Thr Lys Thr Gly Tyr Ser Thr Asp
            500                 505                 510

Ala Glu Val Leu Glu Glu Leu Leu Pro Phe His Glu Ile Ile Gly Lys
        515                 520                 525

Ile Leu Asn Tyr Arg Gln Leu Met Lys Leu Lys Ser Thr Tyr Thr Asp
    530                 535                 540

Gly Leu Met Pro Leu Ile Asn Glu Arg Thr Gly Lys Leu His Thr Thr
545                 550                 555                 560

Phe Asn Gln Thr Gly Thr Leu Thr Gly Arg Leu Ala Ser Ser Glu Pro
                565                 570                 575

Asn Leu Gln Asn Ile Pro Ile Arg Leu Glu Leu Gly Arg Lys Leu Arg
            580                 585                 590
```

-continued

```
Lys Met Phe Ile Pro Ser Pro Gly Tyr Asp Tyr Ile Val Ser Ala Asp
        595                 600                 605

Tyr Ser Gln Ile Glu Leu Arg Leu Leu Ala His Phe Ser Glu Glu Pro
    610                 615                 620

Lys Leu Ile Glu Ala Tyr Gln Lys Gly Glu Asp Ile His Arg Lys Thr
625                 630                 635                 640

Ala Ser Glu Val Phe Gly Val Ser Leu Glu Glu Val Thr Pro Glu Met
                645                 650                 655

Arg Ala His Ala Lys Ser Val Asn Phe Gly Ile Val Tyr Gly Ile Ser
            660                 665                 670

Asp Phe Gly Leu Gly Arg Asp Leu Lys Ile Pro Arg Glu Val Ala Gly
        675                 680                 685

Lys Tyr Ile Lys Asn Tyr Phe Ala Asn Tyr Pro Lys Val Arg Glu Tyr
    690                 695                 700

Leu Asp Glu Leu Val Arg Thr Ala Arg Glu Lys Gly Tyr Val Thr Thr
705                 710                 715                 720

Leu Phe Gly Arg Arg Arg Tyr Ile Pro Glu Leu Ser Ser Lys Asn Arg
                725                 730                 735

Thr Val Gln Gly Phe Gly Glu Arg Thr Ala Met Asn Thr Pro Leu Gln
            740                 745                 750

Gly Ser Ala Ala Asp Ile Ile Lys Leu Ala Met Ile Asn Val Glu Lys
        755                 760                 765

Glu Leu Lys Ala Arg Lys Leu Lys Ser Arg Leu Leu Leu Ser Val His
    770                 775                 780

Asp Glu Leu Val Leu Glu Val Pro Ala Glu Glu Leu Glu Glu Val Lys
785                 790                 795                 800

Ala Leu Val Lys Gly Val Met Glu Ser Val Val Glu Leu Lys Val Pro
                805                 810                 815

Leu Ile Ala Glu Val Gly Ala Gly Lys Asn Trp Tyr Glu Ala Lys
            820                 825                 830
```

What is claimed is:

1. A DNA polymerase having reverse transcriptase efficiency compared with a control DNA polymerase, wherein the DNA polymerase comprises a motif in the polymerase domain comprising

X_1-X_2-X_3-F-X_4-X_5-X_6-X_7-D-X_8-H-T-X_9-T-A-X_{10}-X_{11} (SEQ ID NO:8); wherein:

X_1 is I, L, V, Q or M;

X_2 is R, K, Q or E;

X_3 is V or A;

X_4 is Q, R, E, K, or V;

X_5 is E or R;

X_6 is G or D;

X_7 is K, R, I, L, or A;

X_8 is any amino acid other than I or V;

X_9 is Q, E, L, I, R, or K;

X_{10} is S, A, or M; and

X_{11} is W, R, K, Q or D.

2. The DNA polymerase of claim **1**, wherein X_8 is selected from G, A, W, P, S, T, F, Y, C, N, Q, D, E, K, R, L, M, or H.

3. The DNA polymerase of claim **1**, wherein the DNA polymerase comprises a motif in the polymerase domain comprising

I-R-V-F-X4-E-G-X7-D-X_8-H-T-X9-T-A-X10-W (SEQ ID NO:9); wherein:

X_4 is Q or R;

X_7 is K or R;

X_8 is any amino acid other than I;

X_9 is Q or E; and

X_{10} is S or A.

4. The DNA polymerase of claim **1**, wherein X_8 is any amino acid other than I.

5. The DNA polymerase of claim **1**, wherein X_8 is F.

6. The DNA polymerase of claim **1**, further comprising the amino acid motif X_1-X_2-X_3-X_4-X_5-X_6-X_7-X_8-X_9-X_{10}-X_{11}-X_{12}-X_{13}-G-Y-V-X_{14}-T-L) (SEQ ID NO:29); wherein

X_1 is A, D, S, E, R or Q;

X_2 is W or Y;

X_3 is any amino acid other than I, L, or M;

X_4 is E, A, Q, K, N or D;

X_5 is K, G, R, Q, H or N;

X_6 is T, V, M or I;

X_7 is L, V or K;

X_8 is E, S, A, D or Q;

X_9 is E or F;

X_{10} is G or A;

X_{11} is R or K;

X_{12} is K, R, E, T or Q;
X_{13} is R, K or H; and
X_{14} is E, R or T.

7. The DNA polymerase of claim **1**, further comprising a motif comprising

T-G-R-L-S-S-X_{b7}-X_{b8}-P-N-L-Q-N (SEQ ID NO:38);
wherein X_{b7} is S or T; and X_{b8} is any amino acid other than D or E.

8. The DNA polymerase of claim **7**, wherein the amino acid corresponding to position X_{b8} is selected from the group consisting of L, G, T, Q, A, S, N, R, and K.

9. The DNA polymerase of claim **8**, wherein the amino acid corresponding to position X_{b8} is G.

10. The DNA polymerase of claim **1**, wherein the DNA polymerase comprises an amino acid sequence having at least 90% sequence identity to a sequence selected from the group consisting of SEQ ID NOS: 1, 2, 3, 4, 5, 6, 7, 32, 33, 34, 35, 36, and 37.

11. The DNA polymerase of claim **1**, wherein the control DNA polymerase has the same amino acid sequence as the DNA polymerase except that the amino acid sequence of the control DNA polymerase at position X_8 is I or V.

12. A recombinant nucleic acid encoding the DNA polymerase according to claim **1**.

13. A method for conducting primer extension, comprising: contacting a DNA polymerase as in claim **1** with a primer, a polynucleotide template, and nucleoside triphosphates under conditions suitable for extension of the primer, thereby producing an extended primer.

14. The method of claim **13**, wherein the template is RNA.

15. The method of claim **13**, wherein the primer extension method comprises a polymerase chain reaction (PCR).

16. A kit for producing an extended primer, comprising: at least one container providing a DNA polymerase as in claim **1**.

17. The kit according to claim **16**, further comprising one or more additional containers selected from the group consisting of:

(a) a container providing a primer hybridizable, under primer extension conditions, to a predetermined polynucleotide template;
(b) a container providing nucleoside triphosphates; and
(c) a container providing a buffer suitable for primer extension.

18. A reaction mixture comprising a DNA polymerase as in claim **1**, at least one primer, a polynucleotide template, and nucleoside triphosphates.

19. The reaction mixture of claim **18**, wherein the polynucleotide template is RNA.

20. The reaction mixture of claim **18**, further comprising a second thermostable DNA polymerase.

* * * * *