

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第3553018号
(P3553018)

(45) 発行日 平成16年8月11日(2004.8.11)

(24) 登録日 平成16年5月14日(2004.5.14)

(51) Int. Cl.⁷

F I

C 1 2 N 15/09

C 1 2 N 15/00 Z N A A

C 1 2 N 9/12

C 1 2 N 9/12

C 1 2 Q 1/68

C 1 2 Q 1/68 Z

請求項の数 12 (全 69 頁)

(21) 出願番号	特願2000-536832 (P2000-536832)	(73) 特許権者	591215177
(86) (22) 出願日	平成11年3月15日 (1999.3.15)		ロシュ ダイアグノスティックス ゲーエムベーハー
(65) 公表番号	特表2002-506637 (P2002-506637A)		ドイツ連邦共和国 68298 マンハイム, サンドホファーシュトラッセ 116
(43) 公表日	平成14年3月5日 (2002.3.5)	(74) 代理人	100091096
(86) 国際出願番号	PCT/EP1999/001674		弁理士 平木 祐輔
(87) 国際公開番号	W01999/047649	(74) 代理人	100096183
(87) 国際公開日	平成11年9月23日 (1999.9.23)		弁理士 石井 貞次
審査請求日	平成12年9月13日 (2000.9.13)	(74) 代理人	100118773
(31) 優先権主張番号	198 10 879.6		弁理士 藤田 節
(32) 優先日	平成10年3月13日 (1998.3.13)	(72) 発明者	フレイ, ブルノ
(33) 優先権主張国	ドイツ (DE)		ドイツ連邦共和国 ディーラー 82377 ペンツベルク, ホシュフェルドシュトラッセ 50
前置審査			最終頁に続く

(54) 【発明の名称】 ポリメラーゼキメラ

(57) 【特許請求の範囲】

【請求項 1】

少なくとも2つの異なるポリメラーゼの機能性アミノ酸断片からなるポリメラーゼキメラであって、これらの機能性アミノ酸断片が該ポリメラーゼキメラ中で活性であり、ポリメラーゼ活性をもつドメインが第1のポリメラーゼに由来し、3'-5'エキソヌクレアーゼ活性をもつドメインが第2のポリメラーゼに由来し、そして該ポリメラーゼキメラのアミノ酸配列が配列番号10に示されるアミノ酸配列からなることを特徴とする、上記ポリメラーゼキメラ。

【請求項 2】

前記キメラがさらにRT活性をもつ、請求項1項記載のポリメラーゼキメラ。

【請求項 3】

前記キメラのアミノ酸配列にヒスチジンタグが組み込まれている、請求項1または2記載のポリメラーゼキメラ。

【請求項 4】

請求項1～3のいずれか1項記載のポリメラーゼキメラをコードするDNA。

【請求項 5】

配列番号4に示されるポリメラーゼキメラをコードするDNA。

【請求項 6】

請求項4または5記載のDNA配列を含むベクター。

【請求項 7】

10

20

請求項 6 記載のベクターを含む形質転換細胞。

【請求項 8】

請求項 1 ~ 3 のいずれか 1 項記載のポリメラーゼキメラの作製方法であって、次のステップ：

- アミノ酸配列アライメント、3次元モデル、または実験的に決定された3次元構造の助けをかりて変異体を設計すること、
 - 遺伝子工学的操作によりドメイン交換変異体を作製すること、
 - 出発ベクターに該DNA断片を連結すること、
 - 該DNA断片を担うベクターにより形質転換された宿主において該キメラを発現させること、
 - 発現されたポリメラーゼキメラを精製すること、
- を含んでなる方法。

10

【請求項 9】

PCRのための請求項 1 ~ 3 のいずれか 1 項記載のポリメラーゼキメラの使用。

【請求項 10】

DNA断片の配列を決定するための請求項 1 ~ 3 のいずれか 1 項記載のポリメラーゼキメラの使用。

【請求項 11】

RNA鋳型により開始するRT-PCRのための請求項 1 ~ 3 のいずれか 1 項記載のポリメラーゼキメラの使用。

20

【請求項 12】

請求項 1 ~ 3 のいずれか 1 項記載のポリメラーゼキメラを含有するキット。

【発明の詳細な説明】

【0001】

本発明は、特定の用途に対して有利である、天然ポリメラーゼの諸性質を合わせもつ、ドメインを表すアミノ酸断片から構成されるポリメラーゼキメラに関するものである。驚いたことに、各種の酵素からのドメインは該キメラ中で活性であり、協働作用を示すことがわかった。本発明は特に、ポリメラーゼ活性を有するドメインと3' - 5'エキソヌクレアーゼ活性を有するドメインが異なる酵素に由来するようなポリメラーゼキメラに関する。この種のキメラはRT活性をもつこともできる。さらに、本発明は、本発明によるキメラの作製方法、並びに、例えばポリメラーゼ連鎖反応において、核酸を合成するための該キメラの使用に関する。また、本発明は、本発明のポリメラーゼキメラを含有するキットに関する。

30

【0002】

Braithwaite, D. K. および Ito, J. (1993) *Nucl. Acids Res.* 21, 787 - 802 によると、DNAポリメラーゼはそれらのアミノ酸配列の類似性に従って3つの主要なファミリー(サブクラスを含む)に分類される。Joyce, C. M. および Steitz, T. A. (1994) *Annu. Rev. Biochem.* 63, 777 - 822 には、見出されたモチーフと保存アミノ酸がまとめて記載されている。原核生物においては、主に3種類のポリメラーゼ、すなわちポリメラーゼ I、II および III に区別される。これらのポリメラーゼは細胞内のそれらの機能の点で相違し、またそれらの性質の点でも相違する。DNAポリメラーゼ I は修復酵素であると考えられており、しばしば5' - 3' および3' - 5'エキソヌクレアーゼ活性をもつ。ポリメラーゼ II は損傷した鋳型鎖から開始するDNA合成を促進すると考えられており、それゆえに突然変異を防止する。ポリメラーゼ III は細胞の複製酵素であり、ヌクレオチドを高速(約30,000/分)で合成し、非常に高いプロセスビティ(連続的合成能)をもつと考えられている。ポリメラーゼ III は5' - 3'エキソヌクレアーゼ活性をもっていない。ポリメラーゼ類のその他の性質は、それらの起源、例えば熱安定性またはプロセスビティによるものである。

40

【0003】

50

ポリメラーゼの特定の性質はその用途に従うことが望ましい。例えば、熱安定性で、忠実度が高く（すなわち、ブルフリーディング活性をもつポリメラーゼ）、プロセス的で、高速合成能のポリメラーゼはPCRに適している。ジデオキシヌクレオチドとデオキシヌクレオチドをそれほど区別しない酵素は配列決定に適している。これに対して、ポリメラーゼのブルフリーディング（校正）活性、すなわち3' - 5' エキソヌクレアーゼ活性は配列決定には好ましくない。PCRなどのいくつかの用途においては、ポリメラーゼは5' - 3' エキソヌクレアーゼ活性（5' ヌクレアーゼ活性）を全くもたないか、ほとんどもたないことが望ましい。

【0004】

ポリメラーゼはまた、鋳型としてRNAを受け入れるその能力、すなわちその逆転写酵素（RT）活性の点で相違しうる。RT活性はマンガンおよび/またはマグネシウムイオンの存在に依存性でありうる。多くの場合、ポリメラーゼのRT活性はマンガンイオンに非依存性であることが望ましい。なぜならば、ポリメラーゼのリーディング精度はマンガンイオンの存在下で低下するからである。さらに、ポリメラーゼはそのプロセスビティの点でも相違し、プロセスビティは多くの用途にとって望ましい性質である。

【0005】

したがって、特定の用途に関してポリメラーゼの性質を最適化するという必要性が存在する。以前には、かかる必要性はしばしばポリメラーゼに突然変異を導入したり、ポリメラーゼの機能を欠失させたりすることによって達成されていた。

【0006】

こうして、例えば、5' - 3' エキソヌクレアーゼ活性は、突然変異の導入（Merkens, L.S. (1995) *Biochem. Biophys. Acta* 1264, 243 - 248）またはトランケーション（Jacobsen, H. (1974) *Eur. J. Biochem.* 45, 623 - 627; Barnes, W.M. (1992) *Gene* 112, 29 - 35）によって除去された。ジデオキシヌクレオチドとデオキシヌクレオチドを区別するポリメラーゼの能力は、点突然変異を導入することによって低下させた（Tabor S. および Richardson, C.C. (1995) *Proc. Natl. Acad. Sci.* 92, 6339 - 6343）。TaborとRichardsonは、活性部位ハイブリッドの構築について記載している。

【0007】

最適化された性質をもつポリメラーゼを提供する目的は、構造的および機能的に互いに独立しているドメインを交換することによりポリメラーゼのキメラを作製することで、初めて本発明により達成された。本発明においてドメインとは、そのドメインが本質的にその機能を保持するように、全ての必須中心または全ての機能的に重要なアミノ酸を含む領域として理解される。それゆえ、ドメインの一部のみ、つまりドメインの機能性断片を交換することも可能である。かくして、本発明においては、これらのドメインを機能性アミノ酸断片ということが出来る。さらに、突然変異またはトランケーションによって前記キメラを改変することも出来る。それが有利であると考えられるならば、それぞれの用途に応じてその性質をさらに最適化する突然変異を前記キメラに導入することも可能である。例えば、ポリメラーゼのジデオキシヌクレオチドとデオキシヌクレオチドとを識別する能力を低減させる突然変異を導入することが出来る。あるいはまた、プロセスビティのような望ましい性質を、突然変異の導入またはトランケーションによって強化したり、導入したりすることも出来る。また、突然変異やトランケーションの導入によって、5' ヌクレアーゼ活性などの望ましくない性質を取り除くことも可能である。

【0008】

かくして、本発明の主題は、特定の用途に応じて天然のポリメラーゼの有利な性質を合わせもつポリメラーゼキメラである。本発明によるポリメラーゼキメラは、異なる酵素の機能性アミノ酸断片（好ましくは、異なる酵素のドメインを表すもの）から構成される。本発明では、驚いたことに、異なる酵素からのドメインが該キメラ中で活性であり、ドメイ

10

20

30

40

50

ン間で協働作用を示すことが見出された。本発明はまた、最適化された性質をもつポリメラーゼキメラの一般的な作製方法に関する。それゆえ、この本発明方法は、ドメインを交換することにより酵素の任意の組合せからキメラを設計することを可能にする。さらに、ドメイン間の接触部位での相互作用を種々の方法で調和させることが好適である。これは例えばキメラの熱安定性の向上をもたらす。本発明の更なる主題は、本発明によるキメラを含む核酸合成用のキットである。

【0009】

ブルーフリーディング機能をもつ熱安定性DNAポリメラーゼは、PCRにおいて実際に用いられる機会が増えつつある。Taqポリメラーゼと熱安定性ブルーフリーディングDNAポリメラーゼ（例えば、Pfu、Pwo、Ventポリメラーゼ）の混合物を使用すると、特に長いDNA分子をうまく増幅できることがわかった。したがって、本発明の更なる主題は、Taqポリメラーゼの高いプロセスビティおよび熱安定性と、別のDNAポリメラーゼの3'-5'エキソヌクレアーゼ活性とを、1つの酵素にまとめることであった。それゆえ、本発明は特に、少なくともTaqポリメラーゼのプロセスビティに相当するプロセスビティを有し、かつ3'-5'エキソヌクレアーゼ活性（ブルーフリーディング活性）の存在ゆえに増幅反応中にポリマー鎖にヌクレオチドを組み込むときの誤差率が低い熱安定性ポリメラーゼキメラに関する。これら2つの性質を組み合わせることにより、例えば長いPCR産物、すなわち2kb以上の核酸断片を作る能力を備えたキメラを作製することが可能となる。本発明によるキメラは比較的短い断片にも適している。

10

20

【0010】

かくして、本発明は特に、2つの異なるポリメラーゼの機能性アミノ酸断片から成るポリメラーゼキメラに関し、ここで第1または第2ポリメラーゼが3'-5'エキソヌクレアーゼ活性を有し、それゆえ該ポリメラーゼキメラは5'-3'ポリメラーゼ活性と3'-5'エキソヌクレアーゼ活性を有するものである。ポリメラーゼは天然のポリメラーゼであっても、組換え体のポリメラーゼであってもよい。本発明によるポリメラーゼキメラは2種類または数種類の異なるポリメラーゼに由来する機能性アミノ酸断片から構成することができる。本発明によるポリメラーゼキメラは異なるポリメラーゼに由来する2種類または数種類の機能性アミノ酸断片から構成することができる。該断片のアミノ酸配列は天然のポリメラーゼの配列または突然変異によって改変された配列に相当するものであり得る。

30

【0011】

ポリメラーゼキメラ構築用のアミノ酸断片は、第1または第2ポリメラーゼの機能性ポリメラーゼドメインに各々一致することが好ましい。本発明の意味する機能性ポリメラーゼドメインは、活性にとって必須である全てのアミノ酸を含有する領域であり、以下ではドメインと省略するものとする。

【0012】

本発明は特に、少なくとも2つの異なるポリメラーゼからの機能性アミノ酸断片（短いドメイン中の）からなるポリメラーゼキメラに関し、ここでポリメラーゼ活性を有するドメインは1つのポリメラーゼと相同であり、3'エキソヌクレアーゼ活性を有するドメインはもう1つのポリメラーゼと相同である。さらに、このキメラは、5'エキソヌクレアーゼ活性を有するドメインが第1または第2ポリメラーゼと相同であり得る場合、5'エキソヌクレアーゼ活性をも有する可能性がある。しかし、5'エキソヌクレアーゼドメインが部分的もしくは完全に欠失するか、または点突然変異を有する可能性もある。本発明によるポリメラーゼキメラは、逆転写酵素（RT）活性をさらに有することができる。

40

【0013】

ポリメラーゼキメラのアミノ酸断片の一部がTaqポリメラーゼのアミノ酸配列の一部に一致していることがさらに好ましい。

【0014】

3'-5'エキソヌクレアーゼ活性を有するドメインまたはアミノ酸断片が該キメラに組

50

み込まれたポリメラーゼは、例えば、Pol-I型ポリメラーゼか、またはPol-II型ポリメラーゼでもあり得る。3'-5'エキソヌクレアーゼ活性を有するPol-I型ポリメラーゼの代表的なものは、例えば、大腸菌 (*Escherichia coli*) ポリメラーゼ (Ec.1)、サルモネラ (*Salmonella*) ポリメラーゼI、バチルス (*Bacillus*) ポリメラーゼI、テルモシフォン (*Thermosiphon*) ポリメラーゼIおよびテルモトガ・ネアポリタナ (*Thermotoga neapolitana*) ポリメラーゼ (Tne) である。3'-5'エキソヌクレアーゼ活性を有するPol-II型ポリメラーゼの代表的なものは、例えば、ピロコッカス・ウーゼイ (*Pyrococcus woesei*) ポリメラーゼ (Pwo)、ピロコッカス・フリオサス (*Pyrococcus furiosus*) ポリメラーゼ (Pfu)、テルモコッカス・リトラリス (*Thermococcus litoralis*) ポリメラーゼ (Tli)、ピロディクタム・アビシ (*Pyrodictum abyssii*) ポリメラーゼである。

【0015】

例として挙げた代表的なPol-I型およびPol-II型ポリメラーゼを、以下でさらに詳しく説明する。

【0016】

テルムス・アクアティカス (*Thermus aquaticus*) 由来のTaq DNAポリメラーゼ (Taqポリメラーゼ)、大腸菌DNAポリメラーゼI (*E. coli pol I*) およびテルモトガ・ネアポリタナ (*Thermotoga neapolitana*) DNAポリメラーゼ (Tneポリメラーゼ) は、Aファミリー由来の細菌のDNAポリメラーゼである。それらはpol I型のDNAポリメラーゼである。何故なら、様々な酵素活性が、大腸菌 pol Iに見られるものとかかなり類似した様式で様々なドメインに位置しているからである。ピロコッカス・ウーゼイ (*Pyrococcus woesei*) DNAポリメラーゼ (Pwoポリメラーゼ) は、テルモコッカス・リトラリス (*Thermococcus litoralis*) DNAポリメラーゼ (Vent (登録商標) ポリメラーゼ) およびピロコッカス・フリオサス (*Pyrococcus furiosus*) DNAポリメラーゼ (Pfuポリメラーゼ) と同様、Bファミリーの古細菌のDNAポリメラーゼである。

【0017】

Taqポリメラーゼは、Chien, A.ら (1976)、J. Bacteriol. 127, 1550-1557、Kaledin, A.S.ら (1980)、Biokhimiya 45, 644-651およびLawyer, F.C.ら (1989)、J. Biol. Chem. 264, 6427-6437により記載されている。それは、もともと好熱性真正細菌であるテルムス・アクアティカス (*Thermus aquaticus*) から単離され、後に大腸菌中にクローニングされた。該酵素は94 kDaの分子量を有し、モノマーで活性である。Taqポリメラーゼは、高い熱安定性 (95で40分/100で5分の半減期) および高度にプロセッシブな5'-3' DNAポリメラーゼ (ポリマー化速度: 1秒あたり75ヌクレオチド) を有するため、ポリメラーゼ連鎖反応 (PCR) に使用するのに適している。ポリメラーゼ活性とは別に、5'ヌクレアーゼ活性がLongleyら (1990)、Nucl. Acids Res. 18、7317-7322により検出された。この酵素は3'-5'エキソヌクレアーゼ活性を全く有していないため、ポリヌクレオチド鎖を連続的に伸長させるための4つのデオキシリボヌクレオチド三リン酸の取り込み中にエラーが生じ、これが遺伝子増幅を妨げる (エラー率: 2×10^{-4} エラー/塩基、Cha, R.S. およびThilly, W.G. (1993)、PCR Methods Applic. 3, 18-29)。Taqポリメラーゼの三次構造は、1995年以来公知である (Kimら、1995、Korolevら、1995)。

【0018】

大腸菌 pol I は、Kornberg, A. および Baker, T.A. (1992)

10

20

30

40

50

、DNA Replication、第2版、Freeman、New York、113
165に記載されている。この酵素は、103 kDaの分子量を有し、モノマーで活
性である。大腸菌 pol Iは、5'ヌクレアーゼ活性および5' - 3' DNAポリメラー
ゼ活性を有する。Taqポリメラーゼとは対照的に、大腸菌 pol Iはプルーフリーディ
ング機能としての3' - 5'エキソヌクレアーゼ活性をさらに有する。大腸菌 pol Iお
よびそのクレノー(Klenow)断片(Jacobsen, H.ら(1974)、E
ur. J. Biochem. 45, 623 - 627)を、Taqポリメラーゼの導入
の前にPCRに用いた。しかし、その低い熱安定性により、それらはあまり適していない
。何故なら、各サイクルでそれらを新しく添加する必要があるからである。大腸菌 pol
Iのクレノー断片の三次構造は、1983年以来公知である(Brick, P.ら(1
983)、J. Mol. Biol. 166, 453 - 456、Ollis, D.L
.ら(1985)、Nature 313, 762 - 766およびBeese, L.S
.ら(1993)、Science 260, 352 - 355)。

【0019】

Tneポリメラーゼは、好熱性真正細菌であるテルモトガ・ネアポリタナ(Thermo
toga neapolitana)から単離され、後に大腸菌中にクローニングされた
。Tneポリメラーゼのアミノ酸配列は、テルモトガ・マリチマ(Thermotoga
maritima)DNAポリメラーゼ(UITma(登録商標)ポリメラーゼ)のも
のと類似している(B. Frey博士からの個人的情報)。それは、高い熱安定性、5
'ヌクレアーゼ活性、3' - 5'エキソヌクレアーゼ活性および5' - 3' DNAポリメ
ラーゼ活性を有する。不利な点は、Taqポリメラーゼと比較してポリマー化速度が遅い
ことである。高い忠実度を要する場合、同様のアミノ酸配列を有するUITma(登録商
標)ポリメラーゼをPCRに用いる。Tneポリメラーゼの構造のうち、現在までに公知
であるのはアミノ酸配列のみである(Boehringer Mannheim)。しかし、この酵素は大腸菌 pol Iと相同であるため、三次構造は未知であるが、相同性モデ
リングは可能である。

【0020】

Pfuポリメラーゼは、超好熱性海生古細菌であるピロコッカス・フリオサス(Pyrr
ococcus furiosus)から単離された。それは、高い熱安定性(95で
1時間後に95%の活性)、3' - 5'エキソヌクレアーゼ活性および5' - 3' DNA
ポリメラーゼ活性を有する(Lundberg, K.S.ら(1991)、Gene 1
08, 1 - 6)。DNA合成の忠実度は、Taqポリメラーゼの約10倍高い。高い忠
実度を要する場合、PfuポリメラーゼをPCRに用いる。構造のうち、現在までに公知
であるのは、アミノ酸配列のみである。

【0021】

Pwoポリメラーゼ(PCR Applications Manual(1995)、B
oehringer Mannheim GmbH、Biochemica、28 32)
は、もともと超好熱性古細菌であるピロコッカス・ウーゼイ(Pyrrcococcus
woesei)から単離され、後に大腸菌中にクローニングされた。この酵素は、約90
kDaの分子量を有し、モノマーで活性である。Pwoポリメラーゼは、Taqポリメ
ラーゼよりも高い熱安定性(半減期は100で2時間以上)、高度にプロセッシブな5
' - 3' DNAポリメラーゼ活性およびDNA合成の忠実度を増大させる高い3' - 5'
エキソヌクレアーゼ活性を有する。この酵素は、5'ヌクレアーゼ活性を有さない。ポリ
マー化速度(1秒あたり30ヌクレオチド)は、Taqポリメラーゼのものより小さい。
高い忠実度を要する場合、この酵素をPCRに用いる。DNA合成の忠実度は、Taqポ
リメラーゼを用いる場合よりも10倍以上高い。

【0022】

Athポリメラーゼは、好熱性古細菌であるアネロセルム・テルモフィルム(Anaer
ocellum thermophilum)から単離され、後に大腸菌中にクローニン
グされた。Athポリメラーゼは、高い熱安定性を有し、安定化させる界面活性剤の非存

10

20

30

40

50

在下で、80 で30分間インキュベートした後にも元の活性の少なくとも90%の活性を依然として有する。このポリメラーゼは、マグネシウムイオンの存在下ではRT活性をも有する。Athポリメラーゼは、「Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH」、Mascheroder Weg 1b, D38124 Braunschweig DSMの受託番号8995に寄託されている。Athポリメラーゼは、5' - 3'ポリメラーゼ活性、5' - 3'エキソヌクレアーゼ活性を有するが、3' - 5'エキソヌクレアーゼ活性は有さない。

【0023】

ヒスチジンタグまたは他の精製補助剤を、ポリメラーゼキメラのアミノ酸配列中にさらに組み込んで、精製を改良することができる。 10

【0024】

ポリメラーゼの3' - 5'エキソヌクレアーゼ活性を別のポリメラーゼ(例えば、Taqポリメラーゼ)中に導入するための主な方法は4つあり、これらも本発明の主題である。

【0025】

1. Taqポリメラーゼの分子領域の交換による別のDNAポリメラーゼの3' - 5'エキソヌクレアーゼ領域の挿入

Taqポリメラーゼは、機能的および構造的に独立した(Joyce, C.M.およびSteitz, T.A. (1987)、TIBS 12, 288-292)ならびに他のDNAポリメラーゼのモデルとして役に立つ(Joyce, C.M. (1991)、Curr. Opin. Struct. Biol. 1, 123-129)ドメインからなる大腸菌pol Iと相同であるので、この手法は特に適している。交換に適したDNAポリメラーゼは、3' - 5'エキソヌクレアーゼが実証されており、そのDNA配列が公知であり、3' - 5'エキソヌクレアーゼ活性をコードする遺伝子が利用可能であるものである。モデル構造に基づく合理的タンパク質設計については、3' - 5'エキソヌクレアーゼ領域およびポリメラーゼ領域が大腸菌pol Iと相同であることがさらに有利である。3' - 5'エキソヌクレアーゼ領域は、大腸菌pol Iの構造によく合致し、Taqポリメラーゼのポリメラーゼ領域に隣接するのが好ましい。さらなる利点は、タンパク質の構造データおよび高い熱安定性が利用可能であるとともに三次構造が解明されているということである。 20 30

【0026】

従って、例えば以下のDNAポリメラーゼが適している。

【0027】

a. 大腸菌 pol I

熱安定性の他に、大腸菌pol Iは、上記の条件を全て満たす。クレノー断片の三次構造は、Brookhavenデータバンクにおいて利用可能であり、Taqポリメラーゼと同様、AファミリーのDNAポリメラーゼに属する。アミノ酸配列の同一性は、32%である。公知のドメイン構造を考慮すると、最大的一致は、2つのタンパク質のN末端およびC末端領域に認められる(5'ヌクレアーゼドメインにおいては32%の同一性、ポリメラーゼドメインにおいては49%の同一性)。より短いTaqポリメラーゼは、3' - 5'エキソヌクレアーゼドメインの領域にいくつかの欠失を有している(3' - 5'エキソヌクレアーゼドメインおよび中間ドメインにおいて14%の同一性)。大腸菌pol Iは熱不安定性であり、キメラタンパク質の2つのドメイン間の境界における相互作用はもはや最適ではないので、おそらく、タンパク質キメラもTaqポリメラーゼより熱安定性が低いであろう。これは、境界におけるアミノ酸を続いて改変することにより回復させることができる。 40

【0028】

b. 熱安定性DNAポリメラーゼ

現在PCRに用いられる3' - 5'エキソヌクレアーゼを有する熱安定性DNAポリメラーゼの中では、Pwoポリメラーゼ、Pfuポリメラーゼ、Vent(登録商標)ポリメ 50

ラーゼ、TneポリメラーゼおよびUITma（登録商標）ポリメラーゼがTaq DNAポリメラーゼとの組合わせに適しているようである。PwoポリメラーゼおよびTneポリメラーゼの遺伝子は入手可能である（Boehringer Mannheim Companyを介して）。Pfuポリメラーゼは、Stratagene Inc.から取得できる。Tneポリメラーゼは、Taqポリメラーゼおよび大腸菌pol Iとの相同性により、合理的タンパク質設計によく適する。Pfuポリメラーゼを用いる場合、アミノ酸配列のアラインメントに基づいてのみ設計が可能であり、公知の保存的アミノ酸および機能にとって必須であるモチーフを考慮に入れる。

【0029】

2. 中間ドメインにおけるTaq DNAポリメラーゼの改変

10

3' - 5' エキソヌクレアーゼ活性を挿入するためには、活性にとって必須である全てのアミノ酸を構造中に挿入する必要がある。現在の知識によると、これは特にExo I、Exo IIおよびExo IIIの3つのモチーフに当てはまる。触媒作用に必要な空間的位置に置くためには、必須モチーフを適当な方法でさらに結合しなければならない。

【0030】

ポリメラーゼ領域においてTaq DNAポリメラーゼを改変することも可能である。ポリメラーゼの新規の(de novo)設計も原理的には考えられる。

【0031】

本発明によるキメラは、

1. 5'ヌクレアーゼドメインを除去すること（タンパク質溶解的にも可能）、または後に5'ヌクレアーゼ活性を不活化すること（Merkens, L.S. (1995)、Biochem. Biophys. Acta 1264, 243-248に記載されている）、
 2. 点突然変異または断片交換により改変すること、
 3. キメラの境界での構造を最適化すること、
 4. ランダム突然変異誘発および/または他のポリメラーゼ遺伝子とのランダム組換えにより最適化すること（分子進化）、
- によりさらに最適化できる。

20

【0032】

本発明によるポリメラーゼキメラの例を以下に挙げる。

30

【0033】

- ・Taq DNAポリメラーゼ(M1-V307) 大腸菌 DNAポリメラーゼ(D355-D501) Taq DNAポリメラーゼ(A406-E832)
- ・Taq DNAポリメラーゼ(M1-P291) 大腸菌 DNAポリメラーゼ(Y327-K511) Taq DNAポリメラーゼ(L416-E832)
- ・Taq DNAポリメラーゼ(M1-P291) 大腸菌 DNAポリメラーゼ(Y327-H519) Taq DNAポリメラーゼ(E424-E832) : 点突然変異 A643G; Ile455Val 配列番号1
- ・Taq DNAポリメラーゼ(M1-P291) 大腸菌 DNAポリメラーゼ(Y327-V536) Taq DNAポリメラーゼ(L441-E832)
- ・Taq DNAポリメラーゼ(M1-P291) 大腸菌 DNAポリメラーゼ(Y327-G544) Taq DNAポリメラーゼ(V449-E832) ; 配列番号2
- ・Taq DNAポリメラーゼ(M1-P302) 大腸菌 DNAポリメラーゼ(K348-S365) Taq DNAポリメラーゼ(A319-E347) 大腸菌 DNAポリメラーゼ(N450-T505) Taq DNAポリメラーゼ(E410-E4832) ;
- ・Taq DNAポリメラーゼ(M1-V307) Tne DNAポリメラーゼ(D323-D468) Taq DNAポリメラーゼ(A406-E832)
- ・Taq DNAポリメラーゼ(M1-P291) Tne DNAポリメラーゼ(P295-I478) Taq DNAポリメラーゼ(L416-E832)

40

50

・ T a q DNAポリメラーゼ (M 1 - P 2 9 1) T n e DNAポリメラーゼ (P 2 9 5 - E 4 8 5) T a q DNAポリメラーゼ (E 4 2 4 - E 8 3 2) ; サイレント突然変異 A 1 4 4 9 C 配列番号 3

・ T a q DNAポリメラーゼ (M 1 - P 2 9 1) T n e DNAポリメラーゼ (P 2 9 5 - V 5 0 2) T a q DNAポリメラーゼ (L 4 4 1 - E 8 3 2)

・ T a q DNAポリメラーゼ (M 1 - P 2 9 1) T n e DNAポリメラーゼ (P 2 9 5 - G 5 1 0) T a q DNAポリメラーゼ (V 4 4 9 - E 8 3 2) ; サイレント突然変異 C 1 7 6 7 T 配列番号 4

・ T a q DNAポリメラーゼ (M 1 - P 3 0 2) T n e DNAポリメラーゼ (E 3 1 6 - D 3 3 3) T a q DNAポリメラーゼ (A 3 1 9 - E 3 4 7) T n e DNAポリメラーゼ (I 3 8 1 - M 3 9 4) T a q DNAポリメラーゼ (R 3 6 2 - L 3 8 0) T n e DNAポリメラーゼ (E 4 1 5 - T 4 7 2) T a q DNAポリメラーゼ (E 4 1 0 - E 8 3 2) ;

・ G 3 0 8 D / V 3 1 0 E / L 3 5 2 N / L 3 5 6 D / E 4 0 1 Y / R 3 0 5 D

・ T a q DNAポリメラーゼ (1 - 2 9 1) P f u DNAポリメラーゼ (V 1 0 0 - R 3 4 6) T a q DNAポリメラーゼ (E 4 2 4 - E 8 3 2)

・ T a q DNAポリメラーゼ (1 - 2 9 1) P f u DNAポリメラーゼ (H 1 0 3 - S 3 3 4) T a q DNAポリメラーゼ (E 4 2 4 - E 8 3 2) ; 配列番号 5

・ T a q DNAポリメラーゼ (1 - 2 9 1) P f u DNAポリメラーゼ (V 1 0 0 - F 3 8 9) T a q DNAポリメラーゼ (E 4 2 4 - E 8 3 2)

・ T a q DNAポリメラーゼ (1 - 2 9 1) P f u DNAポリメラーゼ (V 1 0 0 - F 3 8 9) T a q DNAポリメラーゼ (V 4 4 9 - E 8 3 2) ; 配列番号 6

・ T a q DNAポリメラーゼ (1 - 2 9 1) P f u DNAポリメラーゼ (M 1 - F 3 8 9) T a q DNAポリメラーゼ (V 4 4 9 - E 8 3 2)

上記のポリメラーゼキメラのうち、以下のものをさらに詳しく試験した。

【 0 0 3 4 】

・ T a q DNAポリメラーゼ (M 1 - P 2 9 1) 大腸菌 DNAポリメラーゼ (Y 3 2 7 - H 5 1 9) T a q DNAポリメラーゼ (E 4 2 4 - E 8 3 2) : 点突然変異 A 6 4 3 G ; I l e 4 5 5 V a l (T a q E c 1) 配列番号 1

・ T a q DNAポリメラーゼ (M 1 - P 2 9 1) 大腸菌 DNAポリメラーゼ (Y 3 2 7 - G 5 4 4) T a q DNAポリメラーゼ (V 4 4 9 - E 8 3 2) , (T a q E c 2) 配列番号 2

・ T a q DNAポリメラーゼ (M 1 - P 2 9 1) T n e DNAポリメラーゼ (P 2 9 5 - E 4 8 5) T a q DNAポリメラーゼ (E 4 2 4 - E 8 3 2) ; サイレント突然変異 A 1 4 4 9 C (T a q T n e 1) 配列番号 3

・ T a q DNAポリメラーゼ (M 1 - P 2 9 1) T n e DNAポリメラーゼ (P 2 9 5 - G 5 1 0) T a q DNAポリメラーゼ (V 4 4 9 - E 8 3 2) ; サイレント突然変異 C 1 7 6 7 T (T a q T n e 2) 配列番号 4

・ T a q DNAポリメラーゼ (1 - 2 9 1) P f u DNAポリメラーゼ (V 1 0 0 - R 3 4 6) T a q DNAポリメラーゼ (E 4 2 4 - E 8 3 2) , (T a q P f u 1) 配列番号 5

・ T a q DNAポリメラーゼ (1 - 2 9 1) P f u DNAポリメラーゼ (V 1 0 0 - F 3 8 9) T a q DNAポリメラーゼ (V 4 4 9 - E 8 3 2) , (T a q P f u 2) 配列番号 6

【 0 0 3 5 】

適当なDNAポリメラーゼを選択するために、DNAポリメラーゼおよびDNA結合タンパク質の利用可能な配列のうちの複数のアミノ酸配列アラインメントを、例えばプログラムGCG (D e v e r e u x ら、1984、N u c l . A c i d s R e s . 1 2、3 8 7 - 3 9 5) を用いて確立する。好適なアラインメントを見つけるためには、二次構造推定、既知構造に基づく配列アラインメント、既知モチーフおよび機能的に必須なアミノ

10

20

30

40

50

酸、ならびに系統発生的側面を考慮に入れることが必要である。該タンパク質が機能的および構造的に独立したドメインからなる場合、個々のドメインに関してまずアミノ酸配列アラインメントを確立し、その後のみそれらを完全な配列アラインメントに組み合わせることが適当である。

【0036】

三次構造が既知である相同な配列が見つかった場合、該相同タンパク質から3Dモデル構造を誘導することが可能である。プログラムBRAGI(ReicheltおよびSchomburg、1988、J. Mol. Graph. 6、161-165)を用いてモデルを作成することができる。プログラムAMBER(Weinerら、1984、J. Am. Chem. Soc. 106、765-784)を、個々の分子領域および分子全体の構造のエネルギー最小化のために用いることができ、プログラムProcheckを用いてモデルの質を調べることができる。初めのタンパク質の構造のC座標のみが利用可能である場合、例えばプログラムO(Jonesら、1991、Acta Cryst. A 47、110-119)を用いて該構造を再構築することができる。タンパク質データバンクにおいては利用不可能であるが立体写真の走査および座標のピックアップ(例えば、プログラムMagickを用いる)およびz座標の計算(例えば、プログラムstereoを用いる)により立体写真として既に公表されているC座標を取得することも可能である。アミノ酸配列アラインメント、3Dモデル、または実験的に決定した3D構造に基づいて、変異体を設計することができる。

【0037】

更に、ポリメラーゼ活性をもつドメインが逆転写酵素活性を有するキメラ変異体を作製した。適切なポリメラーゼの例としては、Ath(*Anaerocellum thermophilum*)またはTth(*Thermus thermophilum*)由来のポリメラーゼが挙げられる。3'-5'エキソヌクラーゼ活性に、例えばTneポリメラーゼまたはPfu若しくはPwoポリメラーゼなどの他のポリメラーゼ由来のドメインを挿入する。このキメラは、5'エキソヌクラーゼ活性を有するドメインが第1ポリメラーゼ並びに第2ポリメラーゼ由来でありうる場合に、更に5'-3'エキソヌクラーゼ活性を有することができる。

【0038】

組換えハイブリッドポリメラーゼHYBおよびHYBd5は、*Anaerocellum thermophilum*由来のDNAポリメラーゼのように、マグネシウムイオンの存在下およびマンガンイオンの存在下で比較的強い逆転写酵素活性を有する。図22に示すように、逆転写酵素活性に対するポリメラーゼ活性の比は、このタイプの酵素で最も一般的で公知のTthポリメラーゼを用いた場合よりも有利である。この知見をマグネシウム依存性逆転写酵素活性およびマンガン依存性逆転写酵素活性に当てはめると、*Anaerocellum*ポリメラーゼ由来のポリメラーゼドメインもまたハイブリッド酵素中で完全な活性を示すと結論づけることができる。更に変異体HYBd5は、図21に示すように3'-5'エキソヌクラーゼ活性を有する。この活性は、典型的な「ブルーフリーディング活性」について予想されるように、デオキシヌクレオシド三リン酸の存在によって阻害される。従って、テルモトガ・ネアポリタナ(*Thermotoga neapolitana*)のDNAポリメラーゼ由来のエキソヌクラーゼドメインもまたハイブリッド分子中で活性である。エキソヌクラーゼ活性を阻害する能力はまた、ハイブリッドポリメラーゼ分子の両方のドメインが相互作用し、これゆえハイブリッドポリメラーゼが機能的に天然酵素と非常に類似するというを示す。

【0039】

遺伝子工学的操作によるドメイン交換変異体の作製を、化学的に合成されたオリゴデオキシヌクレオチドで補助して、SOE法(Hortonら、(1989) Gene 77、61-68)に従うPCR突然変異誘発、またはその改良法(実施例におけるスキームを参照のこと)を用いて実施することができる。各DNA断片をアガロースゲル上で分離させ、単離し、出発ベクターにライゲートする。pTE、pTaq、pPL、Blues

10

20

30

40

50

cript など適切なプロモーターを有する pUC 誘導体を、大腸菌用の出発ベクターとして用いることができる。プラスミド DNA を、XL1-blue などの大腸菌株中に形質転換し、いくつかのクローンを選択して、それらのプラスミド DNA を単離する。Novablue、BL21(DE)、MC1000 などの他の菌株を用いることも可能である。もちろん、酵母、植物および哺乳動物細胞などの、他の生物中にクローニングすることも可能である。改変領域において、プラスミド DNA が配列決定されているクローンの前選択を、制限酵素分析によって実施する。

【0040】

標的タンパク質中の遺伝子発現を、Pbtaq などの多くのプラスミド中で IPTG で誘発することができる。多くの種々の変異体を作製する場合、万能な精製法を確立するのが適切である。例えば PCR によって、タンパク質にヒスチジンタグを結合した後に使用することができる、Ni-NTA (ニッケルニトリロ三酢酸) アガロース上でのアフィニティークロマトグラフィーは非常に適したものである。タンパク質濃度をタンパク質アッセイ ESSL (Boehringer Mannheim) を用いて測定することができ、また調製物の副活性による汚染は、市販の Taq ポリメラーゼ (Boehringer Mannheim) について記載のように測定することができる。ポリメラーゼ活性、エキソヌクラーゼ活性および熱安定性試験を行って更に変異体を特性決定し、それぞれの最適温度を決定する。キメラのポリメラーゼ活性を、例えば DNA ーゼ活性化仔ウシ胸腺 DNA への Dig-dUTP の組み込み率を測定することによって、非放射性試験系で測定することができ、あるいは、例えば M13mp9 ssDNA への ^{32}P dCTP の組み込み率を測定することによって、放射性試験系で測定することができる。キメラのポリメラーゼ活性の最適温度を決定するために、ポリメラーゼ反応を種々の温度で実施し、比活性を算出する。熱処理後の残存活性 (すなわち熱処理をしない場合の初期活性の %) を測定して熱安定性を決定する。3' - 5' エキソヌクラーゼ活性を、3' 末端で開始する DNA 鋳型鎖にアニーリングする 5' - Dig 標識化プライマーの組み込みによって示すことができる。3' ミスマッチプライマーの修正およびそれらの伸長 (ブルーフリーディング) は、制限酵素 (例えば EcoRI) の認識配列における、鋳型鎖にアニーリングする ミスマッチ 5' - Dig 標識化プライマーの伸長によって示すことができる。酵素によって ミスマッチが修正される場合のみ制限酵素による切断が可能である。PCR で変異体を用いることによって、プロセスビティ (processivity) を試験することができる。酵素が PCR で使用するのに十分に熱安定性ではない場合、伸長温度として最適な温度で、酵素を連続添加しながら PCR を実施することができる。キメラのエキソヌクラーゼ活性を、放射性試験系で測定することができる。この試験には、キメラポリメラーゼの一定量 (通常 2.5 U) を、標識化 DNA (各試験バッファー中 5 μg ^3H DNA) と共に種々の温度で 4 時間インキュベートする。任意に dNTP を種々の濃度 (0 ~ 0.2 mM) で添加してもよい。反応終了後、放射性標識化ヌクレオチドの放出を測定する。

【0041】

本発明の更なる主題は、上述のポリメラーゼキメラの DNA 配列である。特に、DNA 配列 (配列番号 1 ~ 6) は本発明の主題である。本発明は更に、上述のポリメラーゼキメラのアミノ酸配列に関する。特に、アミノ酸配列 (配列番号 7 ~ 12) は本発明の主題である。更に、DNA 配列 (配列番号 17) は本発明の主題である。

【0042】

上述の DNA 配列を含むベクターを更に本発明の主題とする。好ましいベクターとしては、pBTaq (プラスミド Pbtaq4_オリゴ67 (Vilibrandt (1995)、学位論文、TU Braunschweig)) である。

【0043】

大腸菌株、特にポリメラーゼキメラ遺伝子を有するベクターを含む大腸菌 XL1-blue 株を更に本発明の主題とする。以下の菌株を DSM (Deutsche Sammlung von Mikroorganismen und Zellkulturen Gm

10

20

30

40

50

b H、M a s c h e r o d e r W e g 1 b、D - 3 8 1 2 4 B r a u n s c h w e i g) に寄託した :

- ・大腸菌 X L 1 B l u e x p B T a q E c 1 : T a q E c 1 D S M 番号 1 2 0 5 3
- ・大腸菌 X L 1 B l u e x p B T a q T n e 1 : T a q T n e 1 D S M 番号 1 2 0 5 0
- ・大腸菌 X L 1 B l u e x p B T a q T n e 2 : T a q T n e 2 D S M 番号 1 2 0 5 1
- ・大腸菌 X L 1 B l u e x p B T a q P f u 1 : T a q P f u 1 D S M 番号 1 2 0 5 2

10

【 0 0 4 4 】

本発明のポリメラーゼキメラは、例えばポリメラーゼ連鎖反応など、DNA断片を増幅するのに特に適している。更なる応用として、例えばDNA断片の配列決定がある。

【 0 0 4 5 】

A t h - T n e キメラ用の好ましいベクターには以下のものがある :

大腸菌 B L 2 1 (D E 3) p l y s S x p E T H Y B R : H Y B R

大腸菌 B L 2 1 (D E 3) p l y s S x p E T H Y B R d 5 : H Y B R d 5

【 0 0 4 6 】

ポリメラーゼキメラ遺伝子を有するベクターを含む大腸菌株を更に本発明の主題とする。

以下の菌株をDSM (D e u t s c h e S a m m l u n g v o n M i k r o o r g a n i s m e n u n d Z e l l k u l t u r e n G m b H、M a s c h e r o d e r W e g 1 b、D - 3 8 1 2 9 B r a u n s c h w e i g) に寄託している : H Y B R (D S M 番号 1 2 7 2 0) ; H Y B R d 5 (D S M 番号 1 2 7 1 9) 。

20

【 0 0 4 7 】

上述のA t h - T n e キメラの作製について実施例 8 ~ 1 1 に例示している。RT活性を有する本発明のキメラは、RNAの逆転写に特に適している。

【 0 0 4 8 】

本発明の更なる主題は、少なくとも1つの本発明のポリメラーゼキメラを含むDNA断片を増幅するためのキットである。

【 0 0 4 9 】

実施例 1 : 構築とクローニング

普遍的精製法の確立

N i - N T A (ニッケル - ニトリロトリ酢酸) アガロースでのアフィニティークロマトグラフィーを用いて、ドメイン交換変異体のための精製プロトコルを標準化した。タンパク質変異体を作製する前に、H i s タグを、T a q DNAポリメラーゼに結合するかまたは該ポリメラーゼの内部に挿入することが必要であった。プラスミド P b t a q 4 _ o l i g o 6 7 (B o e h r i n g e r M a n n h e i m) 中に2つの異なったH i s タグ変異体を設計し作製した。変異体 N H i s - T a q P o l は、N - 末端のH i s タグ、H i s タグを場合により切断するためのエンテロキナーゼ切断部位およびH i s タグタンパク質を抗体 (Q u i a g e n) で検出するためのエピトープを有している。これをE c o R I 部位からP s t I 部位までのPCRにより作製した。N - 末端タンパク質の配列決定において、変異体 N H i s - T a q P o l の 2 0 個のN - 末端アミノ酸が正しいことを確認した。

40

【 0 0 5 0 】

配列 : N H i s - T a q P o l

EcoRI TaqPol からのコドン

5' G AA TTC ATG AGG GGC TCG CAT CAC CAT CAC CAT CAC GCT GCT GAC GAT GAC GAT AAA ATG AGG GGC 3'

Met Arg Gly Ser His His His His His His Ala Ala Asp Asp Asp Asp Lys Met Arg Gly

MRGS'His エピトープ [Met-Arg-Gly-Ser-(His)₆] エンテロキナーゼ [(Asp)₄-Lys-X]

配列番号 13 : 5' G AA TTC ATG AGG GGC TCG CAT CAC CAT
CAC CAT CAC GCT GCT GAC GAT GAC GAT AAA ATG AGG GGC 3'

配列番号 14 : Met Arg Gly Ser His His His His His His
Ala Ala Asp Asp Asp Asp Lys Met Arg Gly

10

【0051】

変異体 5DH_{is}-TaqPol は、Taq DNAポリメラーゼの5'ヌクレアーゼドメインのフレキシブルループ中のグリシン79とグリシン80の間にHisタグを有し、EcoRI部位からPstI部位までのPCR突然変異誘発により作製した。

【0052】

配列 : 5DH_{is}-TaqPol

配列番号 15

配列番号 16

20

5' GAG GCC TAC GGG CAT CAC CAT CAC CAT CAC GGG TAC AAG GCG 3'

GluAlaTyrGlyHisHisHisHisHisHisGlyTyrLysAla

【0053】

2つの新遺伝子の各改変領域中の、プラスミドDNAの正確さをDNA配列決定により確認した。改変遺伝子の両方を、Hisタグなしの初期タンパク質と同一の割合および同一の条件下で発現させた。それらはNi-NTAアガロースにより容易に精製でき、標準PCRにおいてHisタグのないTaqポリメラーゼと同じ挙動を示した。N-末端Hisタグを、ドメイン交換変異体を精製するために使用した。

30

【0054】

アミノ酸配列のアライメント

以下のアミノ酸配列のアライメントは、ドメイン交換変異体を設計するために行った：

1. Tne, 大腸菌 I および Taq DNAポリメラーゼ
2. Pfu, 大腸菌 I および Taq DNAポリメラーゼ
3. DNAポリメラーゼの複数のアミノ酸配列アライメント

アライメントは個々の分子領域(ドメイン)に関連してプログラムGCGを用いて確立し、完全な配列アライメントを形成するように組み立てた。その際、既知の2次構造、モチーフおよび必須アミノ酸を考慮に入れ、Taq DNAポリメラーゼの対応するドメインの配列とクレノウ断片の3'-5'エキソヌクレアーゼドメインの配列との、構造に基づいた配列アライメントを用いた(Kimら、(1995) Nature 376, 612-616中の図2d)。

40

【0055】

相同性モデリングのためのクレノウ断片の最初の構造を選択するために、その時に利用可能であった大腸菌 DNAポリメラーゼIの構造を、プログラムBragiおよびRMSfitを用いて比較した：

クレノウ断片-dCMP複合体(PDBコード:1dpi)、2.8オングストローム(1987)、クレノウ断片-dCTP複合体(PDBコード:1kfd)、3.9オングストローム(1993)およびクレノウ断片D355A-DNA複合体(PDBコード:1kln)、3.2オングストローム(1994)。クレノウ断片(PDBコード:1k

50

1 n) 構造を選択した。座標が存在しない (B r a g i プログラム) 2 つの領域中に、2 つのループを組み込み、エネルギーを最小化した (A m b e r プログラム)。タンパク質構造の質をチェックした (P r o c h e c k プログラム)。

【 0 0 5 6 】

3次元モデルの構築

アミノ酸 2 9 2 - 8 3 2 を含む T a q DNA ポリメラーゼの分子領域の 3 次元モデルを、クレノウ断片 (P D B コード: 1 k 1 n) の構造との相同関係において B r a g i プログラムを用いて構築した。このモデリングは、アミノ酸置換、挿入および欠失の導入、新ループ領域のエネルギー最小化ならびに分子全体のエネルギー最小化を含んでなるものである (A m b e r プログラム)。

10

【 0 0 5 7 】

T a q DNA ポリメラーゼの構造は、モデリング作業の時点で既に発表されていたが、タンパク質データベースにおいては利用可能ではなかった。クレノウ断片の 3' - 5' エキソヌクレアーゼドメイン (アミノ酸 2 9 2 - 4 2 3) に相当する T a q DNA ポリメラーゼの中間ドメインのモデルを作製するために、立体写真 (K i m ら、(1 9 9 5) N a t u r e 3 7 6 , 6 1 2 - 6 1 6 中の図 2 c) をスキャンし、C 座標をスクリーン上に選び出し (左および右の写真に対しての x 座標および y 座標) (M a g i c k プログラム、J o h n C r i s t y , E . I . d u P o n t D e N e m o u r s a n d C o m p a n y I n c o r p o r a t e d) 、z 座標を計算し (S t e r e o プログラム、(C o l l a b o r a t i v e C o m p u t a t i o n a l P r o j e c t , N u m b e r 4 (1 9 9 4) A c t a C r y s t . D 5 0 , 7 6 0 - 7 6 3)) 、タンパク質主鎖をポリアラニンの生成 (プログラム 0) により再構築し、アミノ酸置換を行い (B r a g i プログラム)、そして分子全体のエネルギー最小化を行った (A m b e r プログラム)。アミノ酸残基 2 9 2 - 4 2 3 (上記参照) のモデルを、ポリメラーゼドメイン (アミノ酸 4 2 4 - 8 3 2) (上記参照) のモデルに結合させ、一方で T a q DNA ポリメラーゼとクレノウ断片との構造アライメントを可能とした (K i m ら、(1 9 9 5) N a t u r e 3 7 6 , 6 1 2 - 6 1 6 中の図 2 b および 2 c) 。モデル構造の全体をエネルギー最小化し (A m b e r プログラム)、モデル構造の質をチェックした (P r o c h e c k プログラム、(L a s k o w s k i , R . , A . , ら、(1 9 9 3) J . A p p l . C r y s t . 2 6 , 2 8 3 - 2 9 1)) 。

20

30

【 0 0 5 8 】

T n e DNA ポリメラーゼ (残基 2 9 7 - 8 9 3) の 3 次元モデルを、クレノウ断片 (P D B コード: 1 k 1 n) の構造との相同関係において構築した。このモデリングは、アミノ酸置換、挿入および欠失の導入 (B r a g i プログラム)、新ループ領域のエネルギー最小化、分子全体のエネルギー最小化 (A m b e r プログラム) およびモデル構造の質のチェック (P r o c h e c k プログラム) を含んでなるものである。

【 0 0 5 9 】

2 0 種類のタンパク質変異体を設計した。それらは大腸菌 p o l I および T n e ポリメラーゼを用いた場合には 3 次元構造に基づき、P f u ポリメラーゼを用いた場合にはアミノ酸アライメントに基づいた。

40

【 0 0 6 0 】

遺伝子操作によるドメイン交換変異体の作製

N - 末端 H i s タグを P C R により挿入し、ドメイン交換変異体を、化学合成オリゴデオキシヌクレオチドを用いて改変 S O E 法 (H o r t o n ら、(1 9 8 9) G e n e 7 7 , 6 1 - 6 8) によりスキーム中に示されているように作製した。それぞれの DNA 断片をアガロースゲル上で分離し、Q I A q u i c k ゲル抽出キット (Q i a g e n 社) を付随するプロトコルに従って使用して単離し、後続の P C R 反応 I から I V において使用した。また、P C R 反応 V の場合は、認識配列がフランキングプライマー中に位置する 2 種の制限酵素 (E c o R I および P s t I) によってそれらを再切断した。DNA 断片のライゲーション、およびコンピテント X L 1 B l u e 大腸菌細胞の作製と該細胞のエレ

50

クトロレーションによる形質転換を、Villbrandt (1995, Dissertation, TU Braunschweig) に記載された通りに行った。いくつかのクローンを選び出し、それらのプラスミドDNAをQIAprep Spin プラスミドキット (Qiagen社) を付随するプロトコルに従って用いて単離した。微生物学的実施方法、および液体培地またはプレート培地の調製のための配合、並びにグリセリン培養物の確立は、Sambrookらのハンドブック (1989, Molecular cloning - a laboratory manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York) に記載されたように実施した。ドメイン交換変異体は、初期タンパク質と同じ割合で発現させた。

10

【0061】

実施例2：精製（1つのキメラについて）ドメイン交換変異体の精製

全てのドメイン交換変異体を、同一のプロトコルにより大腸菌 (*Escherichia coli*) XL1-Blue から単離した。発酵は、37 °C で16時間、100 mg/ml のアンピシリン、12.5 mg/ml テトラサイクリン、1 mM IPTG を含む LB 培地中で、1リッタースケールで行った。細胞を遠心分離し、20 ml の溶菌緩衝液 (50 mM Tris-HCl, pH 8.5, 10 mM 2-メルカプトエタノール、1 mM PMSF) 中に取り上げ、-70 °C で少なくとも16時間以上凍結し、10分間超音波処理した。細胞破片を遠心分離により除き、滅菌濾過した上清をカラム容量が3.5 ml (半径0.65 cm、高さ2.7 cm) のNi-NTA (ニッケル-ニトリロ酢酸) アガロースカラム (Qiagen) にアプライした。カラムを40 ml のバッファーA (20 mM Tris-HCl, pH 8.5, 100 mM KCl, 20 mM イミダゾール、10 mM 2-メルカプトエタノール、10% (v/v) グリセロール) で洗浄し、続いて10 ml のバッファーB (20 mM Tris-HCl, pH 8.5, 1 M KCl, 20 mM イミダゾール、10 mM 2-メルカプトエタノール、10% (v/v) グリセロール) で洗浄し、そして再び10 ml のバッファーAで洗浄した。15 ml のバッファーC (20 mM Tris-HCl, pH 8.5, 100 mM KCl, 100 mM イミダゾール、10 mM 2-メルカプトエタノール、10% (v/v) グリセロール) で溶出させた。流速は毎分0.5 ml で、画分のサイズは、洗浄画分の場合が10 ml、溶出画分の場合が1 ml とした。プールした画分を保存バッファー (20 mM Tris-HCl, pH 8.0, 100 mM KCl, 0.1 mM EDTA, 1 mM DTT, 0.5% Tween 20, 50% グリセロール) に対して透析し、200 µg/ml のゼラチン、および最終濃度が0.5% となるようにNonident P40を添加した。このタンパク質溶液を-20 °C で保存した。

20

30

【0062】

Ni-NTAアガロースでのドメイン交換変異体TaqEc1の精製の分析を、図7に示す。

【0063】

タンパク質濃度の測定

OD₂₈₀の測定と、タンパク質アッセイESL (Boehringer Mannheim) とによりタンパク質濃度を測定した。図8は、タンパク質純度の測定を示す：SDS-PAGE, Phast system (10-15%) : 銀染色。

40

【0064】

実施例3：キメラのポリメラーゼ活性の至適温度

キメラのポリメラーゼ活性を、非放射性試験システムにおいて測定した。放射性試験システムを使用して数値を調整した。DNase-活性化ウシ胸腺DNAへのDig-dUTPの取込み速度を非放射性試験システムにおいて測定した。50 µl の試験混合物は、5 µl のバッファー混合物 (500 mM Tris-HCl, 150 mM (NH₄)₂SO₄, 100 mM KCl, 70 mM MgCl₂, 100 mM 2-メルカプトエタノー

50

沈殿性 DNA 中に取り込むために必要な酵素量として定義した。標準値を決定するために、全混合物の 2 μ l アリコート乾燥フィルター上に分注して乾燥させた。ブランク値は酵素を含まないサンプルを、同様にインキュベートし洗浄することによっても測定した。

【0067】

至適温度は非放射性 DNA ポリメラーゼ試験を様々な温度で行って決定した。

【0068】

様々な温度での比活性

酵素	温度 (°C)					
	25	37	50	60	72	80
TaqPol (BM)	0.0	0.0	5764.4	8489.1	50000.0	57986.1
NHis-TaqPol	0.0	0.0	5616.1	12165.2	60843.7	74784.4
TaqEc1	704.9	10353.4	50066.5	41034.4	2677.5	1016.2
TaqTne1	0.0	2559.4	15967.0	18900.4	1100.0	0.0
TaqTne2	747.2	5180.2	23549.6	30627.3	64139.1	28727.4

10

【0069】

実施例 4 : キメラのポリメラーゼ活性の熱安定性

熱安定性は、反応混合物を 80 と 95 で 1 分間、3 分間または 6 分間加熱し、続いて残存活性を、非放射性 DNA ポリメラーゼ試験により計測して測定した (図 9 参照)。

【0070】

表 : Taq DNA ポリメラーゼ (TaqPol)、His タグを有する Taq DNA ポリメラーゼ (NHis-TaqPol)、および 3 種類のドメイン交換変異体 (TaqEc1、TaqTne1、TaqTne2) の熱処理後の 72 での残存活性 (熱処理なしの初期活性に対するパーセント)

(Dig-dUTP の DNase - 活性化ウシ胸腺 DNA への取込み)

酵素	1分間	3分間	6分間	1分間	3分間	6分間
	80°C	80°C	80°C	95°C	95°C	95°C
TaqPol	100	100	100	100	100	100
NHis-TaqPol	100	100	100	100	100	100
TaqEc1	0	0	0	0	0	0
TaqTne1	16	0	0	0	0	0
TaqTne2	100	100	100	92	0	0

20

30

【0071】

実施例 5 : 酵素の連続添加を伴う PCR

PCR において酵素を連続的に添加してポリメラーゼキメラを試験した。伸長反応を 72 (図 10) と 55 (図 11) で行った。反応容量 100 μ l の各反応混合物は、DNA または p-plasmid DNA (BM Co.) を 1 ng、各プライマー (25-mer) を 1 μ M、各 dNTP を 200 μ M、および MgCl₂ 含有標準 PCR バッファー (Boehringer Mannheim) を含んでいた。反応条件は下記の通りである。

【0072】

72 での伸長反応の場合 : 94 で 1 分 / 50 で 30 秒 / 72 で 1 分 / / 25 サイクル、PCR 反応前に 94 で 2 分、PCR 反応後に 72 で 7 分。0.5 μ l のドメイン交換変異体を各サイクルごとに 50 で添加。

【0073】

55 での伸長反応の場合 : 95 で 1 分 / 50 で 30 秒 / 55 で 1 分 / / 25 サイクル、PCR 反応前に 95 で 2 分、PCR 反応後に 55 で 7 分。0.5 μ l のドメイン交換変異体を各サイクルごとに 50 で添加。

40

50

【0074】

実施例6：3'-5'エキソヌクレアーゼ試験 - TaqEc1変異体

サンプルを、ヌクレオチド不存在下で、DNA鋳型鎖とアニールする5'-Dig標識プライマーと共にインキュベートした。10 μ lの試験混合物は、1 μ lバッファー(100mM Tris-HCl, 15mM MgCl₂, 500mM KCl, 0.1mg/mlゼラチン, pH 8.3)、1 μ l酵素TaqEc1(500ユニット/ μ l)、1pmol鋳型鎖(50mer, スキーム参照のこと)および500fmolの5'-Dig標識プライマーP1(マッチプライマー, 23mer, スキーム参照のこと)もしくはP2(ミスマッチプライマー, 23mer, スキーム参照のこと)を含んでいた。これらの反応混合物は、50 $^{\circ}$ で、様々なインキュベート時間にてインキュベートした。DNA断片は、12.5%アクリルアミドゲル(SequaGel Kit, Medco Company)にて分離し、接触プロットングによってナイロン膜(Boehringer Mannheim)上に転写した。該ナイロン膜を以下のように処理した：100ml バッファー1(0.1M マレイン酸, 0.15M NaCl, pH 7.5中、1%ブロッキング試薬(Boehringer Mannheim))中での30分間のインキュベート；100ml バッファー2(抗Dig-AP Fabフラグメント抗体(Boehringer Mannheim)をバッファー1中に1:10000で希釈したもの)中での30分間のインキュベート；各回毎135mlのバッファー3(0.3% Tween 20を含むバッファー1)での30分間の洗浄を3回；50ml バッファー4(0.1M Tris-HCl, 0.1M NaCl, 50mM MgCl₂, pH 9.5)中での5分間のインキュベート；50ml バッファー5(CPD star(Boehringer Mannheim)をバッファー4中に1:10000で希釈したもの)中での5分間のインキュベート。上記ナイロン膜をWatman紙上で乾燥させ、化学発光検出のために、化学発光フィルム(Boehringer Mannheim)に30~60分間曝露した。3'-5'エキソヌクレアーゼが存在すれば、プライマーの3'末端での分解が可視化される(図を参照のこと)。Hisタグを有するTaqポリメラーゼ(NHis-TaqPol)を陰性対照として使用し、またUITma DNAポリメラーゼを陽性対照として使用した。両対照酵素に対しては、反応混合物を72 $^{\circ}$ にてインキュベートした。UITma DNAポリメラーゼに対しては、メーカーの反応バッファーを用いた。図12および13は、3'-5'エキソヌクレアーゼ試験の変異体TaqEc1を示す。

【0075】

実施例7：3'-ミスマッチプライマーの修正および伸長 - TaqEc1変異体(3'-ミスマッチプライマー修正アッセイ)

鋳型鎖にアニールするDig標識プライマー(50mer, スキーム参照のこと)を4つの異なる実験において伸長させた。該プライマーは、制限酵素EcoRIの認識配列にアニールするマッチプライマー(P1, 23mer, スキーム参照のこと)および2種類の異なるミスマッチプライマー(P2, P3, 23mer, スキーム参照のこと)であった。20 μ lの試験混合物は、1 μ lバッファー(100mM Tris-HCl, 15mM MgCl₂, 500mM KCl, 0.1mg/mlゼラチン, pH 8.3)、1 μ l酵素TaqEc1(500ユニット/ μ l)、dATP, dCTP, dGTP, dTTPを各々10 μ M、1pmol鋳型鎖、および500fmolの5'-Dig標識プライマーP1(マッチプライマー)、P2(ミスマッチプライマー)もしくはP3(ミスマッチプライマー)を含んでいた。これらの反応混合物は50 $^{\circ}$ で60分間インキュベートした後、5分間で95 $^{\circ}$ まで加熱した。10 μ lアリコートを取り、これを10ユニットのEcoRIによって37 $^{\circ}$ で30分間切断した。そのDNA断片を12.5%アクリルアミドゲル(SequaGel Kit, Medco Company)にて分離し、接触プロットングによりナイロン膜(Boehringer Mannheim)上に転写した。該ナイロン膜は上記のように処理し、化学発光フィルム(Boehringer Mannheim)に30~60分間曝露した。マッチプライマーを

用いた場合、EcoRIを用いた消化によって28bpおよび18bpの断片が生じた。ミスマッチプライマーでは、ミスマッチヌクレオチドがマッチヌクレオチドによって置換される場合のみ、そのような結果を生じた(図14を参照のこと)。

【0076】

実施例8：Ath PolとTne Polのハイブリッドポリメラーゼ遺伝子についての組換えDNAポリメラーゼ設計の改変

コンピューター予測

キメラポリメラーゼ遺伝子の構造は、前記ポリメラーゼと大腸菌POLI遺伝子 - この配列は、Brookhavenのデータバンク中の解明された三次元構造(クレノウ断片についてのもの)と最も高度に一致している - との配列アライメント(Thompson, J. D. Higgins, D. G. およびGibson, T. J. Nucleic Acids Research, 1994, 22: 4673 - 4680)から導き出した。そのペアアライメントは約40%の一致を示し、したがって1KLN構造がおそらく最も可能性の高い基本型であるとみなすことができる。一方の構造から他方の構造へのスムーズな移行を確実にするためには、マルチプルアライメントの観点からみて、3種のタンパク質全てと高度な類似性を有する部位に交点を配置するべきである。それゆえ、交点はポリメラーゼドメインと3' - 5' エキソヌクレアーゼドメインの間とするべきである(図17、18)。

【0077】

ハイブリッドポリメラーゼ遺伝子および発現ベクターの構築

コンピューター予測およびシミュレーションは、ハイブリッド遺伝子の構築のための基礎として役立つ。ATH POLドメインおよびTNE EXOドメインを得るための手法として、図18に示した構造を有する2つのプライマー対を使用したPCR増幅およびサブクローニングを用いた。これらのプライマーは、図2B、Cに示されるように各々の遺伝子のN末端およびC末端ならびに遺伝子中央の連結配列に特異的な配列を有している。ATHUPおよびTNELOWプライマーの12塩基の重複部分は、後のハイブリッド遺伝子の再構築用に設計し、さらに、ポリメラーゼドメインに対する更なる改変のために使用しうる明確なSalI制限部位に挿入したものである。TNEUPおよびATHLOWプライマーの5'側配列のオーバーハングは、必要な断片を後に発現ベクターにサブクローニングするためのNcoIおよびHindIIIの認識部位をコードしている。

【0078】

しかしながら、この方法の適用では、サブクローン化領域の広範囲な配列決定が必要になる。そのため更なる構築物を作製し、その遺伝子間のスプライス結合部を別の位置に、すなわち、当初の結合位置よりもさらに42アミノ酸下流から、より高度な類似性を有するポリメラーゼ間の領域に移した。新しい設計の利点は、提示したスプライス結合部を含むTNEポリメラーゼ配列内に明確なBamHI配列が存在することである。ハイブリッド遺伝子を構築するために、ATHポリメラーゼ配列中に、BamHI配列を組み込み、後に特異的突然変異誘発により遺伝子の部品を組立てるために用いた。この新規化合物のアミノ酸配列およびヌクレオチド配列は図19に示す。

【0079】

ハイブリッドポリメラーゼ遺伝子は、複数回のサブクローニング、特異的突然変異誘発、および配列決定の各工程により、図20に記載の通りに構築した。

【0080】

PCR増幅により得られた断片全てについて、その末端から始めて、後のサブクローニング工程にて用いられる明確な制限部位までを配列決定した。増幅の正確さを確保するために、PCR反応はVentポリメラーゼ(New England Biolabs)を用いて行った。特異的突然変異誘発は、「Quick Change」法(Stratagene)を用いて行った。

【0081】

実施例9：大腸菌におけるハイブリッドポリメラーゼ遺伝子の発現

プラスミド pETHYBR および pETHYBRd5 は、Novogene 社製の大腸菌 BL21 (DE3) plySS 株中へ形質転換して、T7 ポリメラーゼの発現を引き起こした。

【0082】

このハイブリッド POL 遺伝子の発現は、活性化 DNA アッセイを用いて DNA ポリメラーゼ活性を測定することにより組換え株の抽出物においてモニターした。以下の条件を用いた。

【0083】

(1) 組換え大腸菌株を、100 mcg/ml のアンピシリン + 30 mcg/ml のクロラムフェニコールを含む LB 培地 (BL21 (DE3) plySS 中の pETHYBR および pETHYBRd5 に対して) または 100 mcg/ml のアンピシリン + 30 mcg/ml のカナマイシンを含む 20 ml の LB 培地 (JM109 / pSB1611 中の pARHYBd5 に対して) 中で培養した。

【0084】

(2) 前記培養物は、光学密度 OD₅₅₀ が約 0.6 ~ 0.7 に達するまで 37 °C で振とうした。次いで該培養物を 25 ~ 28 °C に冷却し、IPTG を最終濃度 1 mM になるよう加えた。続いてインキュベーションを 25 ~ 30 °C で継続した。4 時間のインキュベート後の非誘導培養物の密度は、2 種類の pET ベクターに関して OD₅₅₀ が約 2.2 であり、誘導培養物については約 1.5 であった。

【0085】

(3) BL21 (DE3) plySS 株のタンパク質抽出物は、前記培養物の 5 ml アリコートペレット化することにより調製した。次いで該細胞ペレットを、40 mM Tris-HCl, pH 8.0, 0.1 mM EDTA, 7 mM 2-メルカプトエタノール, 0.2 mM PMSF, 0.1% Triton X-100 を含む停止バッファー 100 µl 中に再懸濁した。細胞抽出物は、液体窒素 / 温水バス中での細胞懸濁液の凍結と解凍を 2 サイクル行うことによって調製した。続いて KCl 溶液を最終濃度 0.75 M になるよう加え、誘導培養物および非誘導培養物の抽出物を 72 °C で 15 分間加熱し、ペレット化し、そのポリメラーゼ活性を測定するために用いた。この活性測定は、活性化 DNA アッセイ (100 mcg/ml 活性化 DNA, 3 mM MgSO₄, 50 mM Tris-HCl, pH 8.9, 0.1% Triton X-100, 70 µM dA-P33, 5 - 10 µCi/ml) において、2 µl の加熱した細胞抽出物を用いて体積 20 µl にて行った。

【0086】

結果は以下の表に示す：

組換え株の抽出物における相対的 DNA ポリメラーゼ活性 (標識の組み込み%, 3 回の独立した測定 of 平均)

株	BL21 (DE3) plyS			
プラスミド	pETHYBR		pETHYBRd5	
IPTG	-	+	-	+
TCA 不溶性 r/a	5	40	2	85

これらのデータは、ハイブリッドポリメラーゼ遺伝子はどちらの種類も pET ベクター系を用いて発現されうることを示している。

【0087】

組換えハイブリッドポリメラーゼの特性づけ

熱安定性

組換えポリメラーゼの熱安定性は、大腸菌株の抽出物を95℃で様々な時間(10、30、60、120分間)加熱することによって測定した。完全型および切詰め型ハイブリッドポリメラーゼは十分に安定ではない(95℃にて10分間インキュベートした後、100%が不活性である)ことが分かった。組換えポリメラーゼの発現の程度を、10% SDS PAGEにおいて加熱した細胞抽出物を解析することにより評価した。誘導培養物と非誘導培養物との間に目に見える差違が認められなかったので、ハイブリッドポリメラーゼの産生は全可溶性タンパク質の1%を上回ることはない結論づけられるであろう。

【0088】

ブルーフリーディング活性

pETHYBRd5に由来する、例えばクレノウ断片などの組換えDNAポリメラーゼのブルーフリーディング活性を、古細菌のDNAに対して用いられるのと同じのプロトコルに従って試験した。組換え酵素はブルーフリーディング活性を有することが分かった。

【0089】

逆転写酵素活性

以下の反応混合物を用いて、逆転写酵素活性を測定した：1μg ポリdA-(dT)₁₅、330μM TTP、0.36μM ジゴキシゲニン-dUTP、200μg/ml BSA、10mM Tris-HCl、pH8.5、20mM KCl。反応混合物中のMgCl₂の濃度は0.5~10mMの間で変化させた。DTEを濃度10mMにて加えた。

【0090】

2μlの組換えDNAポリメラーゼ(pETHYBRd5に由来するもの、例えばクレノウ断片)を反応混合物に加え、50℃で15分間インキュベートした。Mn²⁺を含むTth DNAポリメラーゼを陽性対照として加えた。反応を停止させた後、該混合物を正に荷電したナイロン膜(BM)にアプライした。組み込まれたジゴキシゲニンを1995年のBMプロトコルにより検出した。

【0091】

この組換え酵素(クレノウ断片)は逆転写酵素活性を有することが分かった(図22)。その活性はMn²⁺(最適濃度1mM)の存在に依存する。さらにMg²⁺の存在はさらなる刺激効果を有した(Mg²⁺の最適濃度4mM)。

【0092】

実施例11：キメラポリメラーゼ遺伝子の構築(図20を参照のこと)

制限配列に対する略語 - B - BamHI, Bsp - BspHI, H - HindIII, N - NcoI, R - EcoRI, S - SalI, Sn - SnaI, X - XhoI, Xm - XmaI

1. ベクターpTrcHISB中に完全なポリメラーゼ遺伝子を含むpARHis10プラスミド、ならびにプライマーATHUPおよびATHLOWを用いたATHPOLドメインのPCR増幅、およびpSK+Bluescriptプラスミド中へのサブクローニング pBSAT。この挿入物をフランキングプライマーから配列決定したところ、プライマー合成の際の誤りが原因でATHUPプライマー配列中の1塩基が欠失していることが分かった。

【0093】

2. 1535位にBamHI配列を組み込むための、「Quick change」法(Stratagene)による、プライマーm1およびm2を用いたプラスミドpARHis10の特異的突然変異誘発 pARHis10mut。

【0094】

3. 鋳型であるpTNEC2プラスミドに対しプライマーTNEUPおよびTNELOWを用いたTNE EXOドメインのPCR増幅、およびSmaI切断puC19プラスミド中へのサブクローニング 組込みの方向性が異なるpTEX1およびpTEX2。

【0095】

4. 「LONG」EXOドメインを含む、pTNEC2プラスミドに由来する1444b

10

20

30

40

50

pのXhoI - BamHI断片の、XhoI - BamHI切断プラスミドpTEX1中へのサブクローニング pTEXL。

【0096】

5. 完全なATHポリメラーゼ遺伝子を、2553bpのBamHI - HindIII断片としてBamHI - HindIII切断pTEXL中に組み込む pTEXLATF。

【0097】

6. pTEXLATFプラスミドのXmaI - SnaI断片の、組み込みBamHI配列を含むpARHis10mutプラスミド由来の1094bpのXmaI - SnaI断片による置換 pTEXLATF*。

【0098】

7. pTEXLATF*に由来する4214bpのNcoI - HindIII断片の、NcoI - HindIII切断pET21dベクター中への組み込み pETNAT。

【0099】

8. pETNATプラスミドに由来する、ATHポリメラーゼのN末端ドメインをコードする1535bpのBamHI断片の欠失。これによってTNE EXOLおよびATH POL配列のインフレームでの結合がなされる pETHYBR。

【0100】

9. pETHYBRの1661bpのNcoI - BamHI断片の、pETNATに由来する829bpのBspHI - BamHI断片による置換。これにより、開始コドンとしてTNEポリメラーゼのMet284が使用され、5' - 3'エキソヌクレアーゼ活性を有すると仮定されるN末端ドメインの欠失が引き起こされる pETHYBRd5。

【配列表】

10

20

SEQUENZPROTOKOLL

(1) ALLGEMEINE ANGABEN:

(i) ANMELDER:

(A) NAME: Boehringer Mannheim GmbH
 (B) STRASSE: Sandhoferstr. 116
 (C) ORT: Mannheim
 (E) LAND: DE
 (F) POSTLEITZAHL: 68305
 (G) TELEFON: 06217595482
 (H) TELEFAX: 06217594457

10

(ii) BEZEICHNUNG DER ERFINDUNG: Polymerasenchimaeren

(iii) ANZAHL DER SEQUENZEN: 14

(iv) COMPUTER-LESBARE FASSUNG:

(A) DATENTRÄGER: Floppy disk
 (B) COMPUTER: IBM PC compatible
 (C) BETRIEBSSYSTEM: PC-DOS/MS-DOS
 (D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPA)

(2) ANGABEN ZU SEQ ID NO: 1:

20

(i) SEQUENZKENNZEICHEN:

(A) LÄNGE: 2733 Basenpaare
 (B) ART: Nucleotid
 (C) STRANGFORM: Einzelstrang
 (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Genom-DNA

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 1:

```

ATGAGGGGCT CGCATCACCA TCACCATCAC GCTGCTGACG ATGACGATAA AATGAGGGGC   60
ATGCTACCGC TATTTGAGCC CAAGGGCCCG GTCCTCCTGG TCGACGGCCA CCACCTGGCC  120
TACCGCACCT TCCACGCCCT GAAGGGCCTC ACCACCAGCC GGGGGGAGCC GGTGCAGGCG  180
GTCTACGGCT TCGCCAAGAG CCTCCTCAAG GCCCTCAAGG AGGACGGGGA CGCGGTGATC  240
GTGGTCTTTG ACGCCAAGGC CCCCTCCTTC CGCCACGAGG CCTACGGGGG GTACAAGGCG  300
GGCCGGGCCC CCACGCCGGA GGA CTTTCCC CGGCAACTCG CCCTCATCAA GGAGCTGGTG  360
GACCTCCTGG GGCTGGCGCG CCTCGAGGTC CCGGGCTACG AGGCGGACGA CGTCCTGGCC  420

```

30

AGCCTGGCCA AGAAGGCGGA AAAGGAGGGC TACGAGGTCC GCATCCTCAC CGCCGACAAA 480
 GACCTTTACC AGCTCCTTTC CGACCGCATC CACGTCCTCC ACCCCGAGGG GTACCTCATC 540
 ACCCCGGCCT GGCTTTGGGA AAAGTACGGC CTGAGGCCCG ACCAGTGGGC CGACTACCGG 600
 GCCCTGACCG GGGACGAGTC CGACAACCTT CCCGGGGTCA AGGGCATCGG GGAGAAGACG 660
 GCGAGGAAGC TTCTGGAGGA GTGGGGGAGC CTGGAAGCCC TCCTCAAGAA CCTGGACCGG 720
 CTGAAGCCCG CCATCCGGGA GAAGATCCTG GCCCACATGG ACGATCTGAA GCTCTCCTGG 780
 GACCTGGCCA AGGTGCGCAC CGACCTGCCC CTGGAGGTGG ACTTCGCCAA AAGGCGGGAG 840
 CCCGACCGGG AGAGGCTTAG GGCCTTTCTG GAGAGGCTTG AGTTTGGCAG CCTCCTCCAC 900
 GAGTTCGGCC TTCTGGAAAG CCCCTATGAC AACTACGTCA CCATCCTTGA TGAAGAAACA 960
 CTGAAAGCGT GGATTGCGAA GCTGGAAAAA GCGCCGGTAT TTGCATTTGA TACCGAAACC 1020
 GACAGCCTTG ATAACATCTC TGCTAACCTG GTCGGGCTTT CTTTTGCTAT CGAGCCAGGC 1080
 GTAGCGGCAT ATATTCCGGT TGCTCATGAT TATCTTGATG CGCCCGATCA AATCTCTCGC 1140
 GAGCGTGAC TCGAGTTGCT AAAACCGCTG CTGGAAGATG AAAAGGCGCT GAAGGTCGGG 1200
 CAAAACCTGA AATACGATCG CGGTATTCTG GCGAACTACG GCATTGAACT GCGTGGGATT 1260
 GCGTTTGATA CCATGCTGGA GTCCTACATT CTCAATAGCG TTGCCGGGCG TCACGATATG 1320
 GACAGCCTCG CGGAACGTTG GTTGAAGCAC AAAACCATCA CTTTTGAAGA GATTGCTGGT 1380
 AAAGGCAAAA ATCAACTGAC CTTTAACCAG ATTGCCCTCG AAGAAGCCGG ACGTTACGCC 1440
 GCCGAAGATG CAGATGTCAC CTTGCAGTTG CATCTGAAAA TGTGGCCGGA TCTGCAAAAA 1500
 CACGAGAGGC TCCTTTGGCT TTACCGGGAG GTGGAGAGGC CCCTTTCCGC TGTCTGGCC 1560
 CACATGGAGG CCACGGGGGT GCGCCTGGAC GTGGCCTATC TCAGGGCCTT GTCCCTGGAG 1620
 GTGGCCGAGG AGGTCGCCCC CCTCGAGGCC GAGGTCTTCC GCCTGGCCCG CCACCCCTTC 1680
 AACCTCAACT CCCGGGACCA GCTGGAAAGG GTCCTCTTTG ACGAGCTAGG GCTTCCCGCC 1740
 ATCGGCAAGA CGGAGAAGAC CGGCAAGCGC TCCACCAGCG CCGCCGTCCT GGAGGCCCTC 1800
 CGCGAGGCC ACCCCATCGT GGAGAAGATC CTGCAGTACC GGGAGCTCAC CAAGCTGAAG 1860
 AGCACCTACA TTGACCCCTT GCCGGACCTC ATCCACCCCA GGACGGGCCG CCTCCACACC 1920

10

20

30

CGCTTCAACC AGACGGCCAC GGCCACGGGC AGGCTAAGTA GCTCCGATCC CAACCTCCAG 1980
AACATCCCCG TCCGCACCCC GCTTGGGCAG AGGATCCGCC GGGCCTTCAT CGCCGAGGAG 2040
GGGTGGCTAT TGGTGGCCCT GGA CTATAGC CAGATAGAGC TCAGGGTGCT GGCCACCTC 2100
TCCGGCGACG AGAACCTGAT CCGGGTCTTC CAGGAGGGGC GGGACATCCA CACGGAGACC 2160
GCCAGCTGGA TGTTCGGCGT CCCCCGGGAG GCCGTGGACC CCCTGATGCG CGGGCGGCC 2220
AAGACCATCA ACTTCGGGGT CCTCTACGGC ATGTCGGCCC ACCGCCTCTC CCAGGAGCTA 2280
GCCATCCCTT ACAGGAGGC CCAGGCCTTC ATTGAGCGCT ACTTTCAGAG CTTCCCCAAG 2340
GTGCGGGCCT GGATTGAGAA GACCCTGGAG GAGGGCAGGA GGCGGGGGTA CGTGGAGACC 2400
CTCTTCGGCC GCCGCCGCTA CGTGCCAGAC CTAGAGGCC GGGTGAAGAG CGTGCGGGAG 2460
GCGGCCGAGC GCATGGCCTT CAACATGCC GTCCAGGGCA CCGCCGCCGA CCTCATGAAG 2520
CTGGCTATGG TGAAGCTCTT CCCCAGGCTG GAGGAAATGG GGGCCAGGAT GCTCCTTCAG 2580
GTCCACGACG AGCTGGTCTT CGAGGCCCCA AAAGAGAGGG CGGAGGCCGT GGCCCGCTG 2640
GCCAAGGAGG TCATGGAGGG GGTGTATCCC CTGGCCGTGC CCCTGGAGGT GGAGGTGGGG 2700
ATAGGGGAGG ACTGGCTCTC CGCCAAGGAG TGA 2733

10

20

(2) ANGABEN ZU SEQ ID NO: 2:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 2733 Basenpaare
- (B) ART: Nucleotid
- (C) STRANGFORM: Einzelstrang
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Genom-DNA

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 2:

30

ATGAGGGGCT CGCATACCA TCACCATCAC GCTGCTGACG ATGACGATAA AATGAGGGGC 60
ATGCTACCGC TATTTGAGCC CAAGGGCCGG GTCCTCCTGG TCGACGGCCA CCACCTGGCC 120
TACCGCACCT TCCACGCCCT GAAGGGCCTC ACCACCAGCC GGGGGGAGCC GGTGCAGGCG 180
GTCTACGGCT TCGCCAAGAG CCTCCTCAAG GCCCTCAAGG AGGACGGGGA CGCGGTGATC 240
GTGGTCTTTG ACGCCAAGGC CCCCTCCTTC CGCCACGAGG CCTACGGGGG GTACAAGGCG 300

GGCCGGGCCC CCACGCCGGA GGACTTTCCT CGGCAACTCG CCCTCATCAA GGAGCTGGTG 360
GACCTCCTGG GGCTGGCGCG CCTCGAGGTC CCGGGCTACG AGGCGGACGA CCTCCTGGCC 420
AGCCTGGCCA AGAAGGCGGA AAAGGAGGGC TACGAGGTCC GCATCCTCAC CGCCGACAAA 480
GACCTTTACC AGCTCCTTTC CGACCGCATC CACGTCTCTC ACCCCGAGGG GTACCTCATC 540
ACCCCGCCCT GGCTTTGGGA AAAGTACGGC CTGAGGCCCG ACCAGTGGGC CGACTACCGG 600
GCCCTGACCG GGGACGAGTC CGACAACCTT CCCGGGGTCA AGGGCATCGG GGAGAAGACG 660
GCGAGGAAGC TTCTGGAGGA GTGGGGGAGC CTGGAAGCCC TCCTCAAGAA CCTGGACCGG 720
CTGAAGCCCG CCATCCGGGA GAAGATCCTG GCCCACATGG ACGATCTGAA GCTCTCCTGG 780
GACCTGGCCA AGGTGCGCAC CGACCTGCCC CTGGAGGTGG ACTTCGCCAA AAGGCGGGAG 840
CCCGACCGGG AGAGGCTTAG GGCCTTTCTG GAGAGGCTTG AGTTTGGCAG CCTCCTCCAC 900
GAGTTCGGCC TTCTGGAAAG CCCCTATGAC AACTACGTCA CCATCCTTGA TGAAGAAACA 960
CTGAAAGCGT GGATTGCGAA GCTGGAAAAA GCGCCGGTAT TTGCATTIGA TACCGAAACC 1020
GACAGCCTTG ATAACATCTC TGCTAACCTG GTCGGGCTTF CTTTTGCTAT CGAGCCAGGC 1080
GTAGCGGCAT ATATTCCGGT TGCTCATGAT TATCTTGATG CGCCCGATCA AATCTCTCGC 1140
GAGCGTGAC TCGAGTTGCT AAAACCGCTG CTGGAAGATG AAAAGGCGCT GAAGGTCGGG 1200
CAAAACCTGA AATACGATCG CGGTATTCTG GCGAACTACG GCATTGAACT GCGTGGGATT 1260
GCGTTTGATA CCATGCTGGA GTCCTACATF CTCAATAGCG TTGCCGGGCG TCACGATATG 1320
GACAGCCTCG CGGAACGTTG GTTGAAGCAC AAAACCATCA CTTTGAAGA GATTGCTGGT 1380
AAAGGCAAAA ATCAACTGAC CTTTAACCAG ATTGCCCTCG AAGAAGCCGG ACCTTACGCC 1440
GCCGAAGATG CAGATGTCAC CTTGCAGTTG CATCTGAAAA TGTGGCCGGA TCTGCAAAAA 1500
CACAAAGGGC CGTTGAACGT CTTGAGAAAT ATCGAAATGC CGCTGGTGCC GGTGCTTTCA 1560
CGCATTGAAC GTAACGGTGT GCGCCTGGAC GTGGCCTATC TCAGGGCCTT GTCCCTGGAG 1620
GTGGCCGAGG AGATCGCCCG CCTCGAGGCC GAGGTCTTCC GCCTGGCCGG CCACCCCTTC 1680
AACCTCAACT CCCGGGACCA GCTGGAAAGG GTCCTCTTTG ACGAGCTAGG GCTTCCCGCC 1740
ATCGGCAAGA CGGAGAAGAC CGGCAAGCGC TCCACCAGCG CCGCCGTCCT GGAGGCCCTC 1800

10

20

30

CGCGAGGCC ACCCCATCGT GGAGAAGATC CTGCAGTACC GGGAGCTCAC CAAGCTGAAG 1860
 AGCACCTACA TTGACCCCTT GCCGGACCTC ATCCACCCCA GGACGGGCCG CCTCCACACC 1920
 CGCTTCAACC AGACGGCCAC GGCCACGGGC AGGCTAAGTA GCTCCGATCC CAACCTCCAG 1980
 AACATCCCCG TCCGCACCCC GCTTGGGCAG AGGATCCGCC GGGCCTTCAT CGCCGAGGAG 2040
 GGGTGGCTAT TGGTGGCCCT GGACTATAGC CAGATAGAGC TCAGGGTGCT GGGCCACCTC 2100
 TCCGGCGACG AGAACCTGAT CCGGGTCTTC CAGGAGGGGC GGGACATCCA CACGGAGACC 2160
 GCCAGCTGGA TGTTCGGCCT CCCCCGGGAG GCCGTGGACC CCCTGATGCG CCGGGCGGCC 2220
 AAGACCATCA ACTTCGGGGT CCTCTACGGC ATGTTCGGCC ACCGCCTCTC CCAGGAGCTA 2280
 GCCATCCCTT ACGAGGAGGC CCAGGCCTTC ATTGAGCGCT ACTTTCAGAG CTTCCCCAAG 2340
 GTGCGGGCCT GGATTGAGAA GACCCTGGAG GAGGGCAGGA GGCGGGGTA CGTGGAGACC 2400
 CTCTTCGGCC GCCGCCGCTA CGTGCCAGAC CTAGAGGCC GGGTGAAGAG CGTGCGGGAG 2460
 GCGGCCGAGC GCATGGCCTT CAACATGCC GTCCAGGGCA CCGCCGCCGA CCTCATGAAG 2520
 CTGGCTATGG TGAAGCTCTT CCCCAGGCTG GAGGAAATGG GGGCCAGGAT GCTCCTCAG 2580
 GTCCACGACG AGCTGGTCCT CGAGGCCCA AAAGAGAGGG CGGAGGCCGT GGCCCGGCTG 2640
 GCCAAGGAGG TCATGGAGGG GGTGTATCCC CTGGCCGTGC CCCTGGAGGT GGAGGTGGGG 2700
 ATAGGGGAGG ACTGGCTCTC CGCCAAGGAG TGA 2733

10

20

(2) ANGABEN ZU SEQ ID NO: 3:

- (i) SEQUENZKENNZEICHEN:
 (A) LÄNGE: 2727 Basenpaare
 (B) ART: Nucleotid
 (C) STRANGFORM: Einzelstrang
 (D) TOPOLOGIE: linear

30

(ii) ART DES MOLEKÜLS: Genom-DNA

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 3:

ATGAGGGGCT CGCATCACCA TCACCATCAC GCTGCTGACG ATGACGATAA AATGAGGGGC 60
 ATGCTACCGC TATTTGAGCC CAAGGGCCGG GTCTCTCTGG TCGACGGCCA CCACCTGGCC 120
 TACCGCACCT TCCACGCCCT GAAGGGCCTC ACCACCAGCC GGGGGGAGCC GGTGCAGGCG 180

GTCTACGGCT TCGCCAAGAG CCTCCTCAAG GCCCTCAAGG AGGACGGGGA CGCGGTGATC 240
 GTGGTCTTTG ACGCCAAGGC CCCCTCCTTC CGCCACGAGG CCTACGGGGG GTACAAGGCG 300
 GGCCGGGCCC CCACGCCGGA GGACTTTCCC CGGCAACTCG CCCTCATCAA GGAGCTGGTG 360
 GACCTCCTGG GGCTGGCGCG CCTCGAGGTC CCGGGCTACG AGGCGGACGA CGTCCTGGCC 420
 AGCCTGGCCA AGAAGGCGGA AAAGGAGGGC TACGAGGTCC GCATCCTCAC CGCCGACAAA 480
 GACCTTTACC AGCTCCTTTC CGACCGCATC CACGTCCTCC ACCCCGAGGG GTACCTCATC 540
 ACCCCGGCCT GGCTTTGGGA AAAGTACGGC CTGAGGCCCG ACCAGTGGGC CGACTACCGG 600
 GCCCTGACCG GGGACGAGTC CGACAACCTT CCCGGGGTCA AGGGCATCGG GGAGAAGACG 660
 GCGAGGAAGC TTCTGGAGGA GTGGGGGAGC CTGGAAGCCC TCCTCAAGAA CCTGGACCGG 720
 CTGAAGCCCG CCATCCGGGA GAAGATCCTG GCCCACATGG ACGATCTGAA GCTCTCCTGG 780
 GACCTGGCCA AGGTGCGCAC CGACCTGCCC CTGGAGGTGG ACTTCGCCAA AAGGCGGGAG 840
 CCCGACCGGG AGAGGCTTAG GGCCTTCTG GAGAGGCTTG AGTTTGGCAG CCTCCTCCAC 900
 GAGTTCGGCC TTCTGGAAAG CCCCCCGTT GGATACAGAA TAGTGAAAGA CCTGGTGGAA 960
 TTTGAAAAAC TCATAGAGAA ACTGAGAGAA TCCCCTTCGT TCGCCATAGA TCTTGAGACG 1020
 TCTTCCCTCG ATCCTTTCGA CTGCGACATT GTCGGTATCT CTGTGTCTTT CAAACCAAAG 1080
 GAAGCGTACT ACATACCACT CCATCATAGA AACGCCAGA ACCTGGATGA AAAAGAAGTT 1140
 CTGAAAAAGC TAAAAGAAAT CCTGGAGGAC CCCGGAGCAA AGATCGTTGG TCAGAAATTG 1200
 AAATTCGATT ACAAGGTGTT GATGGTAAAG GGTGTTGAAC CTGTCCCTCC TCACTTCGAC 1260
 ACGATGATAG CGGCTTACCT TCTTGAGCCG AACGAAAAGA AGTTCAATCT GGACGATCTC 1320
 GCATTGAAAT TTCTTGATA CAAAATGACC TCTTACCAGG AACTCATGTC CTTCTCTTCT 1380
 CCGCTGTTTG GTTTCAGTTT TGCCGATGTT CCTGTAGAAA AAGCAGCGAA CTATTCCTGT 1440
 GAAGATGCCG ACATCACCTA CAGACTCTAC AAGATCCTGA GCTTAAAACT CCACGAGGAG 1500
 AGGCTCCTTT GGCTTTACCG GGAGGTGGAG AGGCCCTTT CCGCTGTCCCT GGCCACATG 1560
 GAGGCCACGG GGGTGCCTT GGACGTGGCC TATCTCAGGG CTTTGTCCCT GGAGGTGGCC 1620
 GAGGAGATCG CCCGCTCGA GGCCGAGGTC TTCCGCTGG CCGGCCACCC CTTCAACCTC 1680

10

20

30

AACTCCCGGG ACCAGCTGGA AAGGGTCCTC TTTGACGAGC TAGGGCTTCC CGCCATCGGC 1740
 AAGACGGAGA AGACCGGCAA GCGCTCCACC AGCGCCGCCG TCCTGGAGGC CCTCCGCGAG 1800
 GCCCACCCEA TCGTGGAGAA GATCCTGCAG TACCGGGAGC TCACCAAGCT GAAGAGCACC 1860
 TACATTGACC CCTTGCCGGA CCTCATCCAC CCCAGGACGG GCCGCCTCCA CACCCGCTTC 1920
 AACCAGACGG CCACGGCCAC GGGCAGGCTA AGTAGCTCCG ATCCCAACCT CCAGAACATC 1980
 CCCGTCCGCA CCCCGCTTGG GCAGAGGATC CCGCGGGCCT TCATCGCCGA GGAGGGGTGG 2040
 CTATTGGTGG CCCTGGACTA TAGCCAGATA GAGCTCAGGG TGCTGGCCCA CCTCTCCGGC 2100
 GACGAGAACC TGATCCGGGT CTTCCAGGAG GGGCGGGACA TCCACACGGA GACCGCCAGC 2160
 TGGATGTTTC GCGTCCCCCG GGAGGCCGTG GACCCCTGA TCGCCCGGC GGCCAAGACC 2220
 ATCAACTTCG GGGTCCTCTA CGGCATGTCG GCCCACCGCC TCTCCAGGA GCTAGCCATC 2280
 CCTTACGAGG AGGCCCAGGC CTTTATTGAG CGCTACTTTC AGAGCTTCCC CAAGGTGCGG 2340
 GCCTGGATTG AGAAGACCCT GGAGGAGGGC AGGAGGCGGG GGTACCTGGA GACCCTCTTC 2400
 GGCCGCCGCC GCTACGTGCC AGACCTAGAG GCCCGGGTGA AGAGCGTGCG GGAGGCCGCC 2460
 GAGCGCATGG CCTTCAACAT GCCCGTCCAG GGCACCGCCG CCGACCTCAT GAAGCTGGCT 2520
 ATGGTGAAGC TCTTCCCCAG GCTGGAGGAA ATGGGGGCCA GGATGCTCCT TCAGGTCCAC 2580
 GACGAGCTGG TCCTCGAGGC CCCAAAAGAG AGGGCGGAGG CCGTGGCCCG GCTGGCCAAG 2640
 GAGGTCATGG AGGGGGTGTA TCCCCTGGCC GTGCCCTGG AGGTGGAGGT GGGGATAGGG 2700
 GAGGACTGGC TCTCCGCCAA GGAGTGA 2727

10

20

(2) ANGABEN ZU SEQ ID NO: 4:

- (i) SEQUENZKENNZEICHEN:
- (A) LÄNGE: 2727 Basenpaare
 - (B) ART: Nucleotid
 - (C) STRANGFORM: Einzelstrang
 - (D) TOPOLOGIE: linear

30

(ii) ART DES MOLEKÜLS: Genom-DNA

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 4:

ATGAGGGGCT CGCATACCA TCACCATCAC GCTGCTGACG ATGACGATAA AATGAGGGGC 60

ATGCTACCGC TATTTGAGCC CAAGGGCCCG GTCCTCCTGG TCGACGGCCA CCACCTGGCC 120
 TACCGCACCT TCCACGCCCT GAAGGGCCTC ACCACCAGCC GGGGGGAGCC GGTGCAGGCC 180
 GTCTACGGCT TCGCCAAGAG CCTCCTCAAG GCCCTCAAGG AGGACGGGGA CGCGGTGATC 240
 GTGGTCTTTG ACGCCAAGGC CCCCTCCTTC CGCCACGAGG CCTACGGGGG GTACAAGGCG 300
 GGCCGGGCCC CCACGCCGGA GGACTTTCCC CGGCAACTCG CCCTCATCAA GGAGCTGGTG 360
 GACCTCCTGG GGCTGGCGCG CCTCGAGGTC CCGGGCTACG AGGCGGACGA CGTCCCTGGCC 420
 AGCCTGGCCA AGAAGGCCGA AAAGGAGGGC TACGAGGTCC GCATCCTCAC CGCCGACAAA 480
 GACCTTTACC AGCTCCTTTC CGACCCCATC CACGTCTCC ACCCCGAGGG GTACCTCATC 540
 ACCCCGGCCT GGCTTTGGGA AAAGTACGGC CTGAGGCCCG ACCAGTGGGC CGACTACCGG 600
 GCCCTGACCG GGGACGAGTC CGACAACCTT CCCGGGGTCA AGGGCATCCG GGAGAAGACG 660
 GCGAGGAAGC TTCTGGAGGA GTGGGGGAGC CTGGAAGCCC TCCTCAAGAA CCTGGACCGG 720
 CTGAAGCCCG CCATCCGGGA GAAGATCCTG GCCCACATGG ACGATCTGAA GCTCTCCTGG 780
 GACCTGGCCA AGGTGCGCAC CGACCTGCCC CTGGAGGTGG ACTTCGCCAA AAGGCGGGAG 840
 CCCGACCGGG AGAGGCTTAG GGCCTTTCTG GAGAGGCTTG AGTTTGGCAG CCTCCTCCAC 900
 GAGTTCGGCC TTCTGGAAAG CCCCCCGTT GGATACAGAA TAGTGAAAGA CCTGGTGGAA 960
 TTTGAAAAC TCATAGAGAA ACTGAGAGAA TCCCCTTCGT TCGCCATAGA TCTTGAGACG 1020
 TCTTCCCTCG ATCCTTTCGA CTGCGACATT GTCGGTATCT CTGTGTCTTT CAAACCAAAG 1080
 GAAGCGTACT ACATACCACT CCATCATAGA AACGCCCAGA ACCTGGATGA AAAAGAAGTT 1140
 CTGAAAAGC TAAAAGAAAT CCTGGAGGAC CCCGGAGCAA AGATCGTTGG TCAGAATTTG 1200
 AAATTGATT ACAAGGTGTI GATGGTAAAG GGTGTTGAAC CTGTCCCTCC TCACTTCGAC 1260
 ACGATGATAG CGGCTTACCT TCTTGAGCCG AACGAAAAGA AGTTCAATCT GGACGATCTC 1320
 GCATTGAAAT TTCTTGATA CAAAATGACC TCTTACCAGG AACTCATGTC CTTCTCTTCT 1380
 CCGCTGTTTG GTTTCAGTTF TGCCGATGTT CCTGIAGAAA AAGCAGCGAA CTATTCCTGT 1440
 GAAGATGCAG ACATCACCTA CAGACTCTAC AAGATCCTGA GCTTAAACT CCACGAGGCA 1500
 GATCTGGAGA ACGTGTCTA CAAGATAGAA ATGCCTCTTG TGAGCGTGCT TGCACGGATG 1560

10

20

30

GAACTGAACG GTGTGCGCCT GGACGTGGCC TATCTCAGGG CCTTGTCCCT GGAGGTGGCC 1620
 GAGGAGATCG CCCGCCTCGA GGCCGAGGTC TTCCGCCTGG CCGGCCACCC CTCAACCTC 1680
 AACTCCCGGG ACCAGCTGGA AAGGGTCTCT TTTGACGAGC TAGGGCTTCC CGCCATCGGC 1740
 AAGACGGAGA AGACCGGCAA GCGCTCTACC AGCGCCGCCG TCCTGGAGGC CCTCCGCGAG 1800
 GCCCACCCCA TCGTGGAGAA GATCCTGCAG TACCGGGAGC TCACCAAGCT GAAGAGCACC 1860
 TACATTGACC CCTTGCCGGA CCTCATCCAC CCCAGGACGG GCCGCCTCCA CACCCGCTTC 1920
 AACCAGACGG CCACGGCCAC GGGCAGGCTA AGTAGCTCCG ATCCCAACCT CCAGAACATC 1980
 CCCGTCCGCA CCCCGCTTGG GCAGAGGATC CGCCGGGCCT TCATCGCCGA GGAGGGGTGG 2040
 CTATTGGTGG CCCTGGACTA TAGCCAGATA GAGCTCAGGG TGCTGGCCCA CCTCTCCGGC 2100
 GACGAGAACC TGATCCGGGT CTTCCAGGAG GGGCGGGACA TCCACACGGA GACCCCGAGC 2160
 TGGATGTTCG GCGTCCCCCG GGAGGCCGTG GACCCCTGA TCGCCCGGGC GGCCAAGACC 2220
 ATCAACTTCG GGGTCTCTA CGGCATGTCG GCCCACCGCC TCTCCAGGA GCTAGCCATC 2280
 CCTTACGAGG AGGCCAGGC CTTCAATTGAG CGCTACTTTC AGAGCTTCCC CAAGGTGCGG 2340
 GCCTGGATTG AGAAGACCCT GGAGGAGGGC AGGAGGCGGG GGTACGTGGA GACCCTCTTC 2400
 GGCCGCCGCC GCTACGTGCC AGACCTAGAG GCCCGGGTGA AGAGCGTGCG GGAGGCGGCC 2460
 GAGCGCATGG CCTTCAACAT GCCCGTCCAG GGCACCGCCG CCGACCTCAT GAAGCTGGCT 2520
 ATGGTGAAGC TCTTCCCCAG GCTGGAGGAA ATGGGGGCCA GGATGCTCCT TCAGGTCCAC 2580
 GACGAGCTGG TCCTCGAGGC CCCAAAAGAG AGGGCGGAGG CCGTGGCCCC GCTGGCCAAG 2640
 GAGGTCAATG AGGGGGTGTA TCCCCTGGCC GTGCCCTGG AGGTGGAGGT GGGGATAGGG 2700
 GAGGACTGGC TCTCCGCCAA GGAGTGA 2727

10

20

30

(2) ANGABEN ZU SEQ ID NO: 5:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 2850 Basenpaare
- (B) ART: Nucleotid
- (C) STRANGFORM: Einzelstrang
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Genom-DNA

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 5:

ATGAGGGGCT CGCATCACCA TCACCATCAC GCTGCTGACG ATGACGATAA AATGAGGGGC 60
 ATGCTACCGC TATTTGAGCC CAAGGGCCGG GTCCTCCTGG TCGACGGCCA CCACCTGGCC 120
 TACCGCACCT TCCACGCCCT GAAGGGCCTC ACCACCAGCC GGGGGGAGCC GGTGCAGCGC 180
 GTCTACGGCT TCGCCAAGAG CCTCCTCAAG GCCCTCAAGG AGGACGGGGA CGCGGTGATC 240
 GTGGTCTTTG ACGCCAAGGC CCCCTCCTTC CGCCACGAGG CCTACGGGGG GTACAAGGCG 300
 GGCCGGGCCC CCACGCCGGA GGACTTTCCC CGGCAACTCG CCCTCATCAA GGAGCTGGTG 360
 GACCTCCTGG GGCTGGCGCG CCTCGAGGTC CCGGGCTACG AGGCGGACGA CGTCCTGGCC 420
 AGCCTGGCCA AGAAGGCGGA AAAGGAGGGC TACGAGGTCC GCATCCTCAC CGCCGACAAA 480
 GACCTTTACC AGCTCCTTTC CGACCGCATC CACGTCCTCC ACCCCGAGGG GTACCTCATC 540
 ACCCCGGCCT GGCTTTGGGA AAAGTACGGC CTGAGGCCCG ACCAGTGGGC CGACTACCGG 600
 GCCCTGACCG GGGACGAGTC CGACAACCTT CCCGGGGTCA AGGCATCGG GGAGAAGACG 660
 GCGAGGAAGC TTCTGGAGGA GTGGGGGAGC CTGGAAGCCC TCCTCAAGAA CCTGGACCGG 720
 CTGAAGCCCG CCATCCGGGA GAAGATCCTG GCCCACATGG ACGATCTGAA GCTCTCCTGG 780
 GACCTGGCCA AGGTGCGCAC CGACCTGCCC CTGGAGGTGG ACTTCGCCAA AAGGCGGGAG 840
 CCCGACCGGG AGAGGCTTAG GGCCTTCTG GAGAGGCTTG AGTTTGGCAG CCTCCTCCAC 900
 GAGTTCGGCC TTCTGGAAAG CCCCCATCCA GCAGTTGTGG ACATCTTCGA ATACGATATT 960
 CCATTTGCAA AGAGATACCT CATCGACAAA GGCCTAATAC CAATGGAGGG GGAAGAAGAG 1020
 CTAAAGATTC TTGCCTTCGA TATAGAAACC CTCTATCACG AAGGAGAAGA GTTTGGAAAA 1080
 GGCCCAATTA TAATGATTAG TTATGCAGAT GAAAATGAAG CAAAGGTGAT TACTTGGAAA 1140
 AACATAGATC TTCCATACGT TGAGGTTGTA TCAAGCGAGA GAGAGATGAT AAAGAGATTT 1200
 CTCAGGATTA TCAGGGAGAA GGATCCTGAC ATTATAGTTA CTTATAATGG AGACTCATTC 1260
 GACTTCCCAT ATTTAGCGAA AAGGGCAGAA AACTTGGGA TTAAATTAAC CATTGGAAGA 1320
 GATGGAAGCG AGCCCAAGAT GCAGAGAATA GGCGATATGA CGGCTGTAGA AGTCAAGGGA 1380
 AGAATACATT TCGACTTGTA TCATGTAATA ACAAGGACAA TAAATCTCCC AACATACACA 1440

10

20

30

CTAGAGGCTG TATATGAAGC AATTTTGGGA AAGCCAAAGG AGAAGGTATA CGCCGACGAG 1500
 ATAGCAAAAAG CCTGGGAAAAG TGGAGAGAAC CTTGAGAGAG TTCCCAAATA CTCGATGGAA 1560
 GATGCAAAGG CAACTTATGA ACTCGGGAAA GAATTCCTTC CAATGGAAAT TCAGCTTTCA 1620
 GAGAGGCTCC TTTGGCTTTA CCGGGAGGTG GAGAGGCCCC TTTCCGCTGT CCTGGCCCAC 1680
 ATGGAGGCCA CGGGGGTGCG CCTGGACGTG GCCTATCTCA GGGCCTTGTC CCTGGAGGTG 1740
 GCCGAGGAGA TCGCCCGCCT CGAGGCCGAG GTCTTCCGCC TGGCCGGCCA CCCCTTCAAC 1800
 CTCAACTCCC GGGACCAGCT GGAAAGGGTC CTCTTTGACG AGCTAGGGCT TCCCCCATT 1860
 GGCAAGACCG AGAAGACCGG CAAGCGCTCC ACCAGCGCCG CCGTCCTGGA GGCCCTCCGC 1920
 GAGGCCACC CCATCGTGA GAAGATCCTG CAGTACCGGG AGCTCACCAA GCTGAAGAGC 1980
 ACCTACATTG ACCCCTTGCC GGACCTCATC CACCCCAGGA CGGGCCGCT CCACACCCGC 2040
 TTCAACCAGA CGGCCACGGC CACGGGCAGG CTAAGTAGCT CCGATCCCAA CCTCCAGAAC 2100
 ATCCCCGTCC GCACCCCGCT TGGGCAGAGG ATCCGCCGGG CCTTCATCGC CGAGGAGGGG 2160
 TGGCTATTGG TGGCCCTGGA CTATAGCCAG ATAGAGCTCA GGGTGCTGGC CCACCTCTCC 2220
 GCGGACGAGA ACCTGATCCG GGTCTTCCAG GAGGGGCGGG ACATCCACAC GGAGACCGCC 2280
 AGCTGGATGT TCGGCGTCCC CCGGGAGGCC GTGGACCCCC TGATGCGCCG GCGGGCCAAG 2340
 ACCATCAACT TCGGGGTCTT CTACGGCATG TCGGCCACC GCCTCTCCA GGAGCTAGCC 2400
 ATCCCTTACG AGGAGGCCCA GGCCTTCATT GAGCGTACT TTCAGAGCTT CCCCAAGGTG 2460
 CGGGCCTGGA TTGAGAAGAC CCTGGAGGAG GGCAGGAGGC GGGGGTACGT GGAGACCCTC 2520
 TTCGGCCGCC GCCGCTACGT GCCAGACCTA GAGGCCCGGG TGAAGAGCGT GCGGGAGGCG 2580
 GCCGAGCGCA TGGCCTTCAA CATGCCCGTC CAGGGCACCG CCGCCGACCT CATGAAGCTG 2640
 GCTATGGTGA AGCTCTTCCC CAGGCTGGAG GAAATGGGGG CCAGGATGCT CCTTCAGGTC 2700
 CACGACGAGC TGGTCTCGA GGCCCCAAA GAGAGGGCGG AGGCCGTGGC CCGGCTGGCC 2760
 AAGGAGGTCA TGGAGGGGGT GTATCCCCTG GCCGTGCCCC TGGAGGTGGA GGTGGGGATA 2820
 GGGGAGGACT GGCTCTCCGC CAAGGAGTGA 2850

10

20

30

(2) ANGABEN ZU SEQ ID NO: 6:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 2949 Basenpaare
- (B) ART: Nucleotid
- (C) STRANGFORM: Einzelstrang
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Genom-DNA

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 6:

```

ATGAGGGGCT CGCATCACCA TCACCATCAC GCTGCTGACG ATGACGATAA AATGAGGGGC   60
ATGCTACCGC TATTTGAGCC CAAGGGCCGG GTCCTCCTGG TCGACGGCCA CCACCTGGCC  120
TACCGCACCT TCCACGCCCT GAAGGGCCTC ACCACCAGCC GGGGGGAGCC GGTGCAGGGC  180
GTCTACGGCT TCGCCAAGAG CCTCCTCAAG GCCCTCAAGG AGGACGGGGA CGCGGTGATC  240
GTGGTCTTTG ACGCCAAGGC CCCCTCCTTC CGCCACGAGG CCTACGGGGG GTACAAGGCG  300
GGCCGGGCCC CCACGCCGGA GGACTTTCCC CGGCAACTCG CCCTCATCAA GGAGCTGGTG  360
GACCTCCTGG GGCTGGCGCG CCTCGAGGTC CCGGGCTACG AGGCGGACGA CGTCCTGGCC  420
AGCCTGGCCA AGAAGGCGGA AAAGGAGGGC TACGAGGTCC GCATCCTCAC CGCCGACAAA  480
GACCTTTACC AGCTCCTTTC CGACCGCATC CACGTCCTCC ACCCCGAGGG GTACCTCATC  540
ACCCCGGCCT GGCTTTGGGA AAAGTACGGC CTGAGGCCCG ACCAGTGGGC CGACTACCGG  600
GCCCTGACCG GGGACGAGTC CGACAACCTT CCCGGGGTCA AGGGCATCGG GGAGAAGACC  660
GCGAGGAAGC TTCTGGAGGA GTGGGGGAGC CTGGAAGCCC TCCTCAAGAA CCTGGACCGG  720
CTGAAGCCCG CCATCCGGGA GAAGATCCTG GCCCACATGG ACGATCTGAA GCTCTCCTGG  780
GACCTGGCCA AGGTGCGCAC CGACCTGCCC CTGGAGGTGG ACTTCGCCAA AAGGCGGGAG  840
CCCGACCGGG AGAGGCTTAG GGCCTTTCTG GAGAGGCTTG AGTTTGGCAG CCTCCTCCAC  900
GAGTTCGGCC TTCTGGAAAG CCCCGTTAGA GAACATCCAG CAGTTGTGGA CATCTTCGAA  960
TACGATATTC CATTTGCAAA GAGATACCTC ATCGACAAAG GCCTAATACC AATGGAGGGG 1020
GAAGAAGAGC TAAAGATTCT TGCCTTCGAT ATAGAAACCC TCTATCACGA AGGAGAAGAG 1080
TTTGAAAAG GCCCAATTAT AATGATTAGT TATGCAGATG AAAATGAAGC AAAGGTGATT 1140

```

10

20

30

ACTTGAAAA ACATAGATCT TCCATACGTT GAGGTTGTAT CAAGCGAGAG AGAGATGATA 1200
 AAGAGATTC TCAGGATTAT CAGGGAGAAG GATCCTGACA TTATAGTTAC TTATAATGGA 1260
 GACTCATTCG ACTTCCCATA TTTAGCGAAA AGGGCAGAAA AACTTGGGAT TAAATTAACC 1320
 ATTGGAAGAG ATGGAAGCGA GCCCAAGATG CAGAGAATAG GCGATATGAC GGCTGTAGAA 1380
 GTCAAGGGAA GAATACATTT CGACTTGTAT CATGTAATAA CAAGGACAAT AAATCTCCCA 1440
 ACATACACAC TAGAGGCTGT ATATGAAGCA ATTTTTGGAA AGCCAAAGGA GAAGGTATAC 1500
 GCCGACGAGA TAGCAAAAGC CTGGGAAAGT GGAGAGAACC TTGAGAGAGT TGCCAAATAC 1560
 TCGATGGAAG ATGCAAAGGC AACTTATGAA CTCGGGAAAG AATTCTTCC AATGGAAATT 1620
 CAGCTTTCAA GATTAGTTGG ACAACCTTTA TGGGATGTTT CAAGGTCAAG CACAGGGAAC 1680
 CTTGTAGAGT GGTCTTACT TAGGAAAGCC TACGAAAGAA ACGAAGTAGC TCCAAACAAG 1740
 CCAAGTGAAG AGGAGTATCA AAGAAGGCTC AGGGAGAGCT ACACAGGTGG ATTCGTGCGC 1800
 CTGGACGTGG CCTATCTCAG GGCCTTGTCC CTGGAGGTGG CCGAGGAGAT CGCCCGCTC 1860
 GAGGCCGAGG TCTTCCGCCT GGCCGGCCAC CCCTTCAACC TCAACTCCCG GGACCAGCTG 1920
 GAAAGGGTCC TCTTTGACGA GCTAGGGCTT CCCGCCATCG GCAAGACGGA GAAGACCGGC 1980
 AAGCGCTCCA CCAGCGCCGC CGTCCTGGAG GCCCTCCGCG AGGCCACCC CATCGTGGAG 2040
 AAGATCCTGC AGTACCGGGA GCTCACCAAG CTGAAGAGCA CCTACATTGA CCCCTTGCCG 2100
 GACCTCATCC ACCCCAGGAC GGGCCGCTC CACACCCGCT TCAACCAGAC GGCCACGGCC 2160
 ACGGCAGGC TAAGTAGCTC CGATCCCAAC CTCCAGAACA TCCCGTCCG CACCCCGCTT 2220
 GGGCAGAGGA TCCGCCGGC CTTCATCGCC GAGGAGGGT GGCTATTGGT GGCCCTGGAC 2280
 TATAGCCAGA TAGAGCTCAG GGTGCTGGCC CACCTCTCCG GCGACGAGAA CCTGATCCGG 2340
 GTCTTCCAGG AGGGGCGGGA CATCCACACG GAGACCGCCA GCTGGATGTT CGGCGTCCCC 2400
 CGGGAGGCCG TGGACCCCT GATGCGCCG GCGGCCAAGA CCATCAACTT CGGGTCTCTC 2460
 TACGGCATGT CGGCCACCG CCTCTCCCAG GAGCTAGCCA TCCCTTACGA GGAGGCCAG 2520
 GCCTTCATG AGCGCTACTT TCAGAGCTTC CCCAAGGTGC GGGCCTGGAT TGAGAAGACC 2580
 CTGGAGGAGG GCAGGAGCG GGGGTACGTG GAGACCCTCT TCGGCCGCG CCGCTACGTG 2640

10

20

30

CCAGACCTAG AGGCCCGGGT GAAGAGCGTG CGGGAGGCGG CCGAGCGCAT GGCCTTCAAC 2700
 ATGCCCGTCC AGGGCACCCG CCGCGACCTC ATGAAGCTGG CTATGGTGAA GCTCTTCCCC 2760
 AGGCTGGAGG AAATGGGGGC CAGGATGCTC CTTCAGGTCC ACGACGAGCT GGTCTCTGAG 2820
 GCCCCAAAAG AGAGGGCGGA GGCCGTGGCC CGGCTGGCCA AGGAGGTCAT GGAGGGGGTG 2880
 TATCCCCTGG CCGTGCCCTT GGAGGTGGAG GTGGGGATAG GGGAGGACTG GCTCTCCGCC 2940
 AAGGAGTGA 2949

(2) ANGABEN ZU SEQ ID NO: 7:

10

- (i) SEQUENZKENNZEICHEN:
 (A) LÄNGE: 910 Aminosäuren
 (B) ART: Aminosäure
 (C) STRANGFORM: Einzelstrang
 (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Protein

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 7:

Met Arg Gly Ser His His His His His His Ala Ala Asp Asp Asp Asp
 1 5 10 15
 Lys Met Arg Gly Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu
 20 25 30
 Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe His Ala Leu Lys
 35 40 45
 Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe
 50 55 60
 Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Asp Ala Val Ile
 65 70 75 80
 Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Gly
 85 90 95
 Gly Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln
 100 105 110
 Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Leu Ala Arg Leu
 115 120 125

20

30

Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Ser Leu Ala Lys
 130 135 140
 Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Lys
 145 150 155 160
 Asp Leu Tyr Gln Leu Leu Ser Asp Arg Ile His Val Leu His Pro Glu
 165 170 175
 Gly Tyr Leu Ile Thr Pro Ala Trp Leu Trp Glu Lys Tyr Gly Leu Arg
 180 185 190
 Pro Asp Gln Trp Ala Asp Tyr Arg Ala Leu Thr Gly Asp Glu Ser Asp
 195 200 205
 Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Arg Lys Leu
 210 215 220
 Leu Glu Glu Trp Gly Ser Leu Glu Ala Leu Leu Lys Asn Leu Asp Arg
 225 230 235 240
 Leu Lys Pro Ala Ile Arg Glu Lys Ile Leu Ala His Met Asp Asp Leu
 245 250 255
 Lys Leu Ser Trp Asp Leu Ala Lys Val Arg Thr Asp Leu Pro Leu Glu
 260 265 270
 Val Asp Phe Ala Lys Arg Arg Glu Pro Asp Arg Glu Arg Leu Arg Ala
 275 280 285
 Phe Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly Leu
 290 295 300
 Leu Glu Ser Pro Tyr Asp Asn Tyr Val Thr Ile Leu Asp Glu Glu Thr
 305 310 315 320
 Leu Lys Ala Trp Ile Ala Lys Leu Glu Lys Ala Pro Val Phe Ala Phe
 325 330 335
 Asp Thr Glu Thr Asp Ser Leu Asp Asn Ile Ser Ala Asn Leu Val Gly
 340 345 350
 Leu Ser Phe Ala Ile Glu Pro Gly Val Ala Ala Tyr Ile Pro Val Ala
 355 360 365
 His Asp Tyr Leu Asp Ala Pro Asp Gln Ile Ser Arg Glu Arg Ala Leu
 370 375 380
 Glu Leu Leu Lys Pro Leu Leu Glu Asp Glu Lys Ala Leu Lys Val Gly
 385 390 395 400

10

20

30

Gln Asn Leu Lys Tyr Asp Arg Gly Ile Leu Ala Asn Tyr Gly Ile Glu
405 410 415

Leu Arg Gly Ile Ala Phe Asp Thr Met Leu Glu Ser Tyr Ile Leu Asn
420 425 430

Ser Val Ala Gly Arg His Asp Met Asp Ser Leu Ala Glu Arg Trp Leu
435 440 445

Lys His Lys Thr Ile Thr Phe Glu Glu Ile Ala Gly Lys Gly Lys Asn
450 455 460

Gln Leu Thr Phe Asn Gln Ile Ala Leu Glu Glu Ala Gly Arg Tyr Ala
465 470 475 480

Ala Glu Asp Ala Asp Val Thr Leu Gln Leu His Leu Lys Met Trp Pro
485 490 495

Asp Leu Gln Lys His Glu Arg Leu Leu Trp Leu Tyr Arg Glu Val Glu
500 505 510

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

Leu Asp Val Ala Tyr Leu Arg Ala Leu Ser Leu Glu Val Ala Glu Glu
530 535 540

Val Ala Arg Leu Glu Ala Glu Val Phe Arg Leu Ala Gly His Pro Phe
545 550 555 560

Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg Val Leu Phe Asp Glu Leu
565 570 575

Gly Leu Pro Ala Ile Gly Lys Thr Glu Lys Thr Gly Lys Arg Ser Thr
580 585 590

Ser Ala Ala Val Leu Glu Ala Leu Arg Glu Ala His Pro Ile Val Glu
595 600 605

Lys Ile Leu Gln Tyr Arg Glu Leu Thr Lys Leu Lys Ser Thr Tyr Ile
610 615 620

Asp Pro Leu Pro Asp Leu Ile His Pro Arg Thr Gly Arg Leu His Thr
625 630 635 640

Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly Arg Leu Ser Ser Ser Asp
645 650 655

Pro Asn Leu Gln Asn Ile Pro Val Arg Thr Pro Leu Gly Gln Arg Ile
660 665 670

10

20

30

Arg Arg Ala Phe Ile Ala Glu Glu Gly Trp Leu Leu Val Ala Leu Asp
 675 680 685
 Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala His Leu Ser Gly Asp Glu
 690 695 700
 Asn Leu Ile Arg Val Phe Gln Glu Gly Arg Asp Ile His Thr Glu Thr
 705 710 715 720
 Ala Ser Trp Met Phe Gly Val Pro Arg Glu Ala Val Asp Pro Leu Met
 725 730 735
 Arg Arg Ala Ala Lys Thr Ile Asn Phe Gly Val Leu Tyr Gly Met Ser
 740 745 750 10
 Ala His Arg Leu Ser Gln Glu Leu Ala Ile Pro Tyr Glu Glu Ala Gln
 755 760 765
 Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys Val Arg Ala Trp
 770 775 780
 Ile Glu Lys Thr Leu Glu Glu Gly Arg Arg Arg Gly Tyr Val Glu Thr
 785 790 795
 Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Glu Ala Arg Val Lys
 805 810 815 20
 Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met Pro Val Gln
 820 825 830
 Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys Leu Phe Pro
 835 840 845
 Arg Leu Glu Glu Met Gly Ala Arg Met Leu Leu Gln Val His Asp Glu
 850 855 860
 Leu Val Leu Glu Ala Pro Lys Glu Arg Ala Glu Ala Val Ala Arg Leu
 865 870 875 880
 Ala Lys Glu Val Met Glu Gly Val Tyr Pro Leu Ala Val Pro Leu Glu
 885 890 895 30
 Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys Glu
 900 905 910

(2) ANGABEN ZU SEQ ID NO: 8:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 910 Aminosäuren
- (B) ART: Aminosäure
- (C) STRANGFORM: Einzelstrang
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Protein

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 8:

Met	Arg	Gly	Ser	His	His	His	His	His	His	Ala	Ala	Asp	Asp	Asp	Asp	10
1			5						10					15		
Lys	Met	Arg	Gly	Met	Leu	Pro	Leu	Phe	Glu	Pro	Lys	Gly	Arg	Val	Leu	
			20					25					30			
Leu	Val	Asp	Gly	His	His	Leu	Ala	Tyr	Arg	Thr	Phe	His	Ala	Leu	Lys	
		35					40					45				
Gly	Leu	Thr	Thr	Ser	Arg	Gly	Glu	Pro	Val	Gln	Ala	Val	Tyr	Gly	Phe	
	50					55					60					
Ala	Lys	Ser	Leu	Leu	Lys	Ala	Leu	Lys	Glu	Asp	Gly	Asp	Ala	Val	Ile	20
65				70						75					80	
Val	Val	Phe	Asp	Ala	Lys	Ala	Pro	Ser	Phe	Arg	His	Glu	Ala	Tyr	Gly	
				85					90					95		
Gly	Tyr	Lys	Ala	Gly	Arg	Ala	Pro	Thr	Pro	Glu	Asp	Phe	Pro	Arg	Gln	
			100					105					110			
Leu	Ala	Leu	Ile	Lys	Glu	Leu	Val	Asp	Leu	Leu	Gly	Leu	Ala	Arg	Leu	
		115					120					125				
Glu	Val	Pro	Gly	Tyr	Glu	Ala	Asp	Asp	Val	Leu	Ala	Ser	Leu	Ala	Lys	
	130					135					140					
Lys	Ala	Glu	Lys	Glu	Gly	Tyr	Glu	Val	Arg	Ile	Leu	Thr	Ala	Asp	Lys	30
145				150						155					160	
Asp	Leu	Tyr	Gln	Leu	Leu	Ser	Asp	Arg	Ile	His	Val	Leu	His	Pro	Glu	
			165					170						175		
Gly	Tyr	Leu	Ile	Thr	Pro	Ala	Trp	Leu	Trp	Glu	Lys	Tyr	Gly	Leu	Arg	
			180					185					190			
Pro	Asp	Gln	Trp	Ala	Asp	Tyr	Arg	Ala	Leu	Thr	Gly	Asp	Glu	Ser	Asp	
		195					200					205				

Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Arg Lys Leu
 210 215 220
 Leu Glu Glu Trp Gly Ser Leu Glu Ala Leu Leu Lys Asn Leu Asp Arg
 225 230 235 240
 Leu Lys Pro Ala Ile Arg Glu Lys Ile Leu Ala His Met Asp Asp Leu
 245 250 255
 Lys Leu Ser Trp Asp Leu Ala Lys Val Arg Thr Asp Leu Pro Leu Glu
 260 265 270
 Val Asp Phe Ala Lys Arg Arg Glu Pro Asp Arg Glu Arg Leu Arg Ala
 275 280 285
 Phe Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly Leu
 290 295 300
 Leu Glu Ser Pro Tyr Asp Asn Tyr Val Thr Ile Leu Asp Glu Glu Thr
 305 310 315 320
 Leu Lys Ala Trp Ile Ala Lys Leu Glu Lys Ala Pro Val Phe Ala Phe
 325 330 335
 Asp Thr Glu Thr Asp Ser Leu Asp Asn Ile Ser Ala Asn Leu Val Gly
 340 345 350
 Leu Ser Phe Ala Ile Glu Pro Gly Val Ala Ala Tyr Ile Pro Val Ala
 355 360 365
 His Asp Tyr Leu Asp Ala Pro Asp Gln Ile Ser Arg Glu Arg Ala Leu
 370 375 380
 Glu Leu Leu Lys Pro Leu Leu Glu Asp Glu Lys Ala Leu Lys Val Gly
 385 390 395 400
 Gln Asn Leu Lys Tyr Asp Arg Gly Ile Leu Ala Asn Tyr Gly Ile Glu
 405 410 415
 Leu Arg Gly Ile Ala Phe Asp Thr Met Leu Glu Ser Tyr Ile Leu Asn
 420 425 430
 Ser Val Ala Gly Arg His Asp Met Asp Ser Leu Ala Glu Arg Trp Leu
 435 440 445
 Lys His Lys Thr Ile Thr Phe Glu Glu Ile Ala Gly Lys Gly Lys Asn
 450 455 460
 Gln Leu Thr Phe Asn Gln Ile Ala Leu Glu Glu Ala Gly Arg Tyr Ala
 465 470 475 480

10

20

30

Ala Glu Asp Ala Asp Val Thr Leu Gln Leu His Leu Lys Met Trp Pro
485 490 495

Asp Leu Gln Lys His Lys Gly Pro Leu Asn Val Phe Glu Asn Ile Glu
500 505 510

Met Pro Leu Val Pro Val Leu Ser Arg Ile Glu Arg Asn Gly Val Arg
515 520 525

Leu Asp Val Ala Tyr Leu Arg Ala Leu Ser Leu Glu Val Ala Glu Glu
530 535 540

Ile Ala Arg Leu Glu Ala Glu Val Phe Arg Leu Ala Gly His Pro Phe
545 550 555 560

Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg Val Leu Phe Asp Glu Leu
565 570 575

Gly Leu Pro Ala Ile Gly Lys Thr Glu Lys Thr Gly Lys Arg Ser Thr
580 585 590

Ser Ala Ala Val Leu Glu Ala Leu Arg Glu Ala His Pro Ile Val Glu
595 600 605

Lys Ile Leu Gln Tyr Arg Glu Leu Thr Lys Leu Lys Ser Thr Tyr Ile
610 615 620

Asp Pro Leu Pro Asp Leu Ile His Pro Arg Thr Gly Arg Leu His Thr
625 630 635 640

Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly Arg Leu Ser Ser Ser Asp
645 650 655

Pro Asn Leu Gln Asn Ile Pro Val Arg Thr Pro Leu Gly Gln Arg Ile
660 665 670

Arg Arg Ala Phe Ile Ala Glu Glu Gly Trp Leu Leu Val Ala Leu Asp
675 680 685

Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala His Leu Ser Gly Asp Glu
690 695 700

Asn Leu Ile Arg Val Phe Gln Glu Gly Arg Asp Ile His Thr Glu Thr
705 710 715 720

Ala Ser Trp Met Phe Gly Val Pro Arg Glu Ala Val Asp Pro Leu Met
725 730 735

Arg Arg Ala Ala Lys Thr Ile Asn Phe Gly Val Leu Tyr Gly Met Ser
740 745 750

10

20

30

Ala His Arg Leu Ser Gln Glu Leu Ala Ile Pro Tyr Glu Glu Ala Gln
755 760 765

Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys Val Arg Ala Trp
770 775 780

Ile Glu Lys Thr Leu Glu Glu Gly Arg Arg Arg Gly Tyr Val Glu Thr
785 790 795 800

Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Glu Ala Arg Val Lys
805 810 815

Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met Pro Val Gln
820 825 830 10

Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys Leu Phe Pro
835 840 845

Arg Leu Glu Glu Met Gly Ala Arg Met Leu Leu Gln Val His Asp Glu
850 855 860

Leu Val Leu Glu Ala Pro Lys Glu Arg Ala Glu Ala Val Ala Arg Leu
865 870 875 880

Ala Lys Glu Val Met Glu Gly Val Tyr Pro Leu Ala Val Pro Leu Glu
885 890 895 20

Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys Glu
900 905 910

(2) ANGABEN ZU SEQ ID NO: 9:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 908 Aminosäuren
- (B) ART: Aminosäure
- (C) STRANGFORM: Einzelstrang
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Protein 30

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 9:

Met Arg Gly Ser His His His His His His Ala Ala Asp Asp Asp Asp
1 5 10 15

Lys Met Arg Gly Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu
20 25 30

Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe His Ala Leu Lys
 35 40 45
 Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe
 50 55 60
 Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Asp Ala Val Ile
 65 70 75 80
 Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Gly
 85 90 95
 Gly Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln
 100 105 110
 Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Leu Ala Arg Leu
 115 120 125
 Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Ser Leu Ala Lys
 130 135 140
 Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Lys
 145 150 155 160
 Asp Leu Tyr Gln Leu Leu Ser Asp Arg Ile His Val Leu His Pro Glu
 165 170 175
 Gly Tyr Leu Ile Thr Pro Ala Trp Leu Trp Glu Lys Tyr Gly Leu Arg
 180 185 190
 Pro Asp Gln Trp Ala Asp Tyr Arg Ala Leu Thr Gly Asp Glu Ser Asp
 195 200 205
 Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Arg Lys Leu
 210 215 220
 Leu Glu Glu Trp Gly Ser Leu Glu Ala Leu Leu Lys Asn Leu Asp Arg
 225 230 235 240
 Leu Lys Pro Ala Ile Arg Glu Lys Ile Leu Ala His Met Asp Asp Leu
 245 250 255
 Lys Leu Ser Trp Asp Leu Ala Lys Val Arg Thr Asp Leu Pro Leu Glu
 260 265 270
 Val Asp Phe Ala Lys Arg Arg Glu Pro Asp Arg Glu Arg Leu Arg Ala
 275 280 285
 Phe Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly Leu
 290 295 300

10

20

30

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

10

20

30

Pro Ala Ile Gly Lys Thr Glu Lys Thr Gly Lys Arg Ser Thr Ser Ala
 580 585 590
 Ala Val Leu Glu Ala Leu Arg Glu Ala His Pro Ile Val Glu Lys Ile
 595 600 605
 Leu Gln Tyr Arg Glu Leu Thr Lys Leu Lys Ser Thr Tyr Ile Asp Pro
 610 615 620
 Leu Pro Asp Leu Ile His Pro Arg Thr Gly Arg Leu His Thr Arg Phe
 625 630 635 640
 Asn Gln Thr Ala Thr Ala Thr Gly Arg Leu Ser Ser Ser Asp Pro Asn
 645 650 655
 Leu Gln Asn Ile Pro Val Arg Thr Pro Leu Gly Gln Arg Ile Arg Arg
 660 665 670
 Ala Phe Ile Ala Glu Glu Gly Trp Leu Leu Val Ala Leu Asp Tyr Ser
 675 680 685
 Gln Ile Glu Leu Arg Val Leu Ala His Leu Ser Gly Asp Glu Asn Leu
 690 695 700
 Ile Arg Val Phe Gln Glu Gly Arg Asp Ile His Thr Glu Thr Ala Ser
 705 710 715 720
 Trp Met Phe Gly Val Pro Arg Glu Ala Val Asp Pro Leu Met Arg Arg
 725 730 735
 Ala Ala Lys Thr Ile Asn Phe Gly Val Leu Tyr Gly Met Ser Ala His
 740 745 750
 Arg Leu Ser Gln Glu Leu Ala Ile Pro Tyr Glu Glu Ala Gln Ala Phe
 755 760 765
 Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys Val Arg Ala Trp Ile Glu
 770 775 780
 Lys Thr Leu Glu Glu Gly Arg Arg Arg Gly Tyr Val Glu Thr Leu Phe
 785 790 795 800
 Gly Arg Arg Arg Tyr Val Pro Asp Leu Glu Ala Arg Val Lys Ser Val
 805 810 815
 Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met Pro Val Gln Gly Thr
 820 825 830
 Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys Leu Phe Pro Arg Leu
 835 840 845

10

20

30

Glu Glu Met Gly Ala Arg Met Leu Leu Gln Val His Asp Glu Leu Val
 850 855 860

Leu Glu Ala Pro Lys Glu Arg Ala Glu Ala Val Ala Arg Leu Ala Lys
 865 870 875 880

Glu Val Met Glu Gly Val Tyr Pro Leu Ala Val Pro Leu Glu Val Glu
 885 890 895

Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys Glu
 900 905

(2) ANGABEN ZU SEQ ID NO: 10:

10

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 908 Aminosäuren
- (B) ART: Aminosäure
- (C) STRANGFORM: Einzelstrang
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Protein

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 10:

Met Arg Gly Ser His His His His His His Ala Ala Asp Asp Asp Asp
 1 5 10 15 20

Lys Met Arg Gly Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu
 20 25 30

Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe His Ala Leu Lys
 35 40 45

Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe
 50 55 60

Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Asp Ala Val Ile
 65 70 75 80

Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Gly
 85 90 95

Gly Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln
 100 105 110

Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Leu Ala Arg Leu
 115 120 125

20

30

Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Ser Leu Ala Lys
 130 135 140
 Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Lys
 145 150 155 160
 Asp Leu Tyr Gln Leu Leu Ser Asp Arg Ile His Val Leu His Pro Glu
 165 170 175
 Gly Tyr Leu Ile Thr Pro Ala Trp Leu Trp Glu Lys Tyr Gly Leu Arg
 180 185 190
 Pro Asp Gln Trp Ala Asp Tyr Arg Ala Leu Thr Gly Asp Glu Ser Asp
 195 200 205
 Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Arg Lys Leu
 210 215 220
 Leu Glu Glu Trp Gly Ser Leu Glu Ala Leu Leu Lys Asn Leu Asp Arg
 225 230 235 240
 Leu Lys Pro Ala Ile Arg Glu Lys Ile Leu Ala His Met Asp Asp Leu
 245 250 255
 Lys Leu Ser Trp Asp Leu Ala Lys Val Arg Thr Asp Leu Pro Leu Glu
 260 265 270
 Val Asp Phe Ala Lys Arg Arg Glu Pro Asp Arg Glu Arg Leu Arg Ala
 275 280 285
 Phe Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly Leu
 290 295 300
 Leu Glu Ser Pro Pro Val Gly Tyr Arg Ile Val Lys Asp Leu Val Glu
 305 310 315 320
 Phe Glu Lys Leu Ile Glu Lys Leu Arg Glu Ser Pro Ser Phe Ala Ile
 325 330 335
 Asp Leu Glu Thr Ser Ser Leu Asp Pro Phe Asp Cys Asp Ile Val Gly
 340 345 350
 Ile Ser Val Ser Phe Lys Pro Lys Glu Ala Tyr Tyr Ile Pro Leu His
 355 360 365
 His Arg Asn Ala Gln Asn Leu Asp Glu Lys Glu Val Leu Lys Lys Leu
 370 375 380
 Lys Glu Ile Leu Glu Asp Pro Gly Ala Lys Ile Val Gly Gln Asn Leu
 385 390 395 400

10

20

30

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

10

20

30

Ala Phe Ile Ala Glu Glu Gly Trp Leu Leu Val Ala Leu Asp Tyr Ser
675 680 685

Gln Ile Glu Leu Arg Val Leu Ala His Leu Ser Gly Asp Glu Asn Leu
690 695 700

Ile Arg Val Phe Gln Glu Gly Arg Asp Ile His Thr Glu Thr Ala Ser
705 710 715 720

Trp Met Phe Gly Val Pro Arg Glu Ala Val Asp Pro Leu Met Arg Arg
725 730 735

Ala Ala Lys Thr Ile Asn Phe Gly Val Leu Tyr Gly Met Ser Ala His
740 745 750 10

Arg Leu Ser Gln Glu Leu Ala Ile Pro Tyr Glu Glu Ala Gln Ala Phe
755 760 765

Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys Val Arg Ala Trp Ile Glu
770 775 780

Lys Thr Leu Glu Glu Gly Arg Arg Arg Gly Tyr Val Glu Thr Leu Phe
785 790 795 800

Gly Arg Arg Arg Tyr Val Pro Asp Leu Glu Ala Arg Val Lys Ser Val
805 810 815 20

Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met Pro Val Gln Gly Thr
820 825 830

Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys Leu Phe Pro Arg Leu
835 840 845

Glu Glu Met Gly Ala Arg Met Leu Leu Gln Val His Asp Glu Leu Val
850 855 860

Leu Glu Ala Pro Lys Glu Arg Ala Glu Ala Val Ala Arg Leu Ala Lys
865 870 875 880

Glu Val Met Glu Gly Val Tyr Pro Leu Ala Val Pro Leu Glu Val Glu
885 890 895 30

Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys Glu
900 905

Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Arg Lys Leu
 210 215 220
 Leu Glu Glu Trp Gly Ser Leu Glu Ala Leu Leu Lys Asn Leu Asp Arg
 225 230 235 240
 Leu Lys Pro Ala Ile Arg Glu Lys Ile Leu Ala His Met Asp Asp Leu
 245 250 255
 Lys Leu Ser Trp Asp Leu Ala Lys Val Arg Thr Asp Leu Pro Leu Glu
 260 265 270
 Val Asp Phe Ala Lys Arg Arg Glu Pro Asp Arg Glu Arg Leu Arg Ala
 275 280 285
 Phe Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly Leu
 290 295 300
 Leu Glu Ser Pro His Pro Ala Val Val Asp Ile Phe Glu Tyr Asp Ile
 305 310 315 320
 Pro Phe Ala Lys Arg Tyr Leu Ile Asp Lys Gly Leu Ile Pro Met Glu
 325 330 335
 Gly Glu Glu Glu Leu Lys Ile Leu Ala Phe Asp Ile Glu Thr Leu Tyr
 340 345 350
 His Glu Gly Glu Glu Phe Gly Lys Gly Pro Ile Ile Met Ile Ser Tyr
 355 360 365
 Ala Asp Glu Asn Glu Ala Lys Val Ile Thr Trp Lys Asn Ile Asp Leu
 370 375 380
 Pro Tyr Val Glu Val Val Ser Ser Glu Arg Glu Met Ile Lys Arg Phe
 385 390 395 400
 Leu Arg Ile Ile Arg Glu Lys Asp Pro Asp Ile Ile Val Thr Tyr Asn
 405 410 415
 Gly Asp Ser Phe Asp Phe Pro Tyr Leu Ala Lys Arg Ala Glu Lys Leu
 420 425 430
 Gly Ile Lys Leu Thr Ile Gly Arg Asp Gly Ser Glu Pro Lys Met Gln
 435 440 445
 Arg Ile Gly Asp Met Thr Ala Val Glu Val Lys Gly Arg Ile His Phe
 450 455 460
 Asp Leu Tyr His Val Ile Thr Arg Thr Ile Asn Leu Pro Thr Tyr Thr
 465 470 475 480

10

20

30

Leu Glu Ala Val Tyr Glu Ala Ile Phe Gly Lys Pro Lys Glu Lys Val
 485 490 495
 Tyr Ala Asp Glu Ile Ala Lys Ala Trp Glu Ser Gly Glu Asn Leu Glu
 500 505 510
 Arg Val Ala Lys Tyr Ser Met Glu Asp Ala Lys Ala Thr Tyr Glu Leu
 515 520 525
 Gly Lys Glu Phe Leu Pro Met Glu Ile Gln Leu Ser Glu Arg Leu Leu
 530 535 540
 Trp Leu Tyr Arg Glu Val Glu Arg Pro Leu Ser Ala Val Leu Ala His
 545 550 555 560
 Met Glu Ala Thr Gly Val Arg Leu Asp Val Ala Tyr Leu Arg Ala Leu
 565 570 575
 Ser Leu Glu Val Ala Glu Glu Ile Ala Arg Leu Glu Ala Glu Val Phe
 580 585 590
 Arg Leu Ala Gly His Pro Phe Asn Leu Asn Ser Arg Asp Gln Leu Glu
 595 600 605
 Arg Val Leu Phe Asp Glu Leu Gly Leu Pro Ala Ile Gly Lys Thr Glu
 610 615 620
 Lys Thr Gly Lys Arg Ser Thr Ser Ala Ala Val Leu Glu Ala Leu Arg
 625 630 635 640
 Glu Ala His Pro Ile Val Glu Lys Ile Leu Gln Tyr Arg Glu Leu Thr
 645 650 655
 Lys Leu Lys Ser Thr Tyr Ile Asp Pro Leu Pro Asp Leu Ile His Pro
 660 665 670
 Arg Thr Gly Arg Leu His Thr Arg Phe Asn Gln Thr Ala Thr Ala Thr
 675 680 685
 Gly Arg Leu Ser Ser Ser Asp Pro Asn Leu Gln Asn Ile Pro Val Arg
 690 695 700
 Thr Pro Leu Gly Gln Arg Ile Arg Arg Ala Phe Ile Ala Glu Glu Gly
 705 710 715 720
 Trp Leu Leu Val Ala Leu Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu
 725 730 735
 Ala His Leu Ser Gly Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly
 740 745 750

10

20

30

Arg Asp Ile His Thr Glu Thr Ala Ser Trp Met Phe Gly Val Pro Arg
 755 760 765
 Glu Ala Val Asp Pro Leu Met Arg Arg Ala Ala Lys Thr Ile Asn Phe
 770 775 780
 Gly Val Leu Tyr Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala
 785 790 795 800
 Ile Pro Tyr Glu Glu Ala Gln Ala Phe Ile Glu Arg Tyr Phe Gln Ser
 805 810 815
 Phe Pro Lys Val Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg
 820 825 830
 Arg Arg Gly Tyr Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro
 835 840 845
 Asp Leu Glu Ala Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met
 850 855 860
 Ala Phe Asn Met Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu
 865 870 875 880
 Ala Met Val Lys Leu Phe Pro Arg Leu Glu Glu Met Gly Ala Arg Met
 885 890 895
 Leu Leu Gln Val His Asp Glu Leu Val Leu Glu Ala Pro Lys Glu Arg
 900 905 910
 Ala Glu Ala Val Ala Arg Leu Ala Lys Glu Val Met Glu Gly Val Tyr
 915 920 925
 Pro Leu Ala Val Pro Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp
 930 935 940
 Leu Ser Ala Lys Glu
 945

10

20

30

(2) ANGABEN ZU SEQ ID NO: 12:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 982 Aminosäuren
- (B) ART: Aminosäure
- (C) STRANGFORM: Einzelstrang
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Protein

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 12:

Met Arg Gly Ser His His His His His Ala Ala Asp Asp Asp Asp
 1 5 10 15
 Lys Met Arg Gly Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu
 20 25 30
 Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe His Ala Leu Lys
 35 40 45
 Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe
 50 55 60
 Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Asp Ala Val Ile
 65 70 75 80
 Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Gly
 85 90 95
 Gly Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln
 100 105 110
 Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Leu Ala Arg Leu
 115 120 125
 Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Ser Leu Ala Lys
 130 135 140
 Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Lys
 145 150 155 160
 Asp Leu Tyr Gln Leu Leu Ser Asp Arg Ile His Val Leu His Pro Glu
 165 170 175
 Gly Tyr Leu Ile Thr Pro Ala Trp Leu Trp Glu Lys Tyr Gly Leu Arg
 180 185 190
 Pro Asp Gln Trp Ala Asp Tyr Arg Ala Leu Thr Gly Asp Glu Ser Asp
 195 200 205
 Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Arg Lys Leu
 210 215 220
 Leu Glu Glu Trp Gly Ser Leu Glu Ala Leu Leu Lys Asn Leu Asp Arg
 225 230 235 240
 Leu Lys Pro Ala Ile Arg Glu Lys Ile Leu Ala His Met Asp Asp Leu
 245 250 255

10

20

30

Lys Leu Ser Trp Asp Leu Ala Lys Val Arg Thr Asp Leu Pro Leu Glu
 260 265 270
 Val Asp Phe Ala Lys Arg Arg Glu Pro Asp Arg Glu Arg Leu Arg Ala
 275 280 285
 Phe Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly Leu
 290 295 300
 Leu Glu Ser Pro Val Arg Glu His Pro Ala Val Val Asp Ile Phe Glu
 305 310 315 320
 Tyr Asp Ile Pro Phe Ala Lys Arg Tyr Leu Ile Asp Lys Gly Leu Ile
 325 330 335
 Pro Met Glu Gly Glu Glu Glu Leu Lys Ile Leu Ala Phe Asp Ile Glu
 340 345 350
 Thr Leu Tyr His Glu Gly Glu Glu Phe Gly Lys Gly Pro Ile Ile Met
 355 360 365
 Ile Ser Tyr Ala Asp Glu Asn Glu Ala Lys Val Ile Thr Trp Lys Asn
 370 375 380
 Ile Asp Leu Pro Tyr Val Glu Val Val Ser Ser Glu Arg Glu Met Ile
 385 390 395 400
 Lys Arg Phe Leu Arg Ile Ile Arg Glu Lys Asp Pro Asp Ile Ile Val
 405 410 415
 Thr Tyr Asn Gly Asp Ser Phe Asp Phe Pro Tyr Leu Ala Lys Arg Ala
 420 425 430
 Glu Lys Leu Gly Ile Lys Leu Thr Ile Gly Arg Asp Gly Ser Glu Pro
 435 440 445
 Lys Met Gln Arg Ile Gly Asp Met Thr Ala Val Glu Val Lys Gly Arg
 450 455 460
 Ile His Phe Asp Leu Tyr His Val Ile Thr Arg Thr Ile Asn Leu Pro
 465 470 475 480
 Thr Tyr Thr Leu Glu Ala Val Tyr Glu Ala Ile Phe Gly Lys Pro Lys
 485 490 495
 Glu Lys Val Tyr Ala Asp Glu Ile Ala Lys Ala Trp Glu Ser Gly Glu
 500 505 510
 Asn Leu Glu Arg Val Ala Lys Tyr Ser Met Glu Asp Ala Lys Ala Thr
 515 520 525

10

20

30

Tyr Glu Leu Gly Lys Glu Phe Leu Pro Met Glu Ile Gln Leu Ser Arg
 530 535 540
 Leu Val Gly Gln Pro Leu Trp Asp Val Ser Arg Ser Ser Thr Gly Asn
 545 550 555 560
 Leu Val Glu Trp Phe Leu Leu Arg Lys Ala Tyr Glu Arg Asn Glu Val
 565 570 575
 Ala Pro Asn Lys Pro Ser Glu Glu Glu Tyr Gln Arg Arg Leu Arg Glu
 580 585 590
 Ser Tyr Thr Gly Gly Phe Val Arg Leu Asp Val Ala Tyr Leu Arg Ala
 595 600 605
 Leu Ser Leu Glu Val Ala Glu Glu Ile Ala Arg Leu Glu Ala Glu Val
 610 615 620
 Phe Arg Leu Ala Gly His Pro Phe Asn Leu Asn Ser Arg Asp Gln Leu
 625 630 635 640
 Glu Arg Val Leu Phe Asp Glu Leu Gly Leu Pro Ala Ile Gly Lys Thr
 645 650 655
 Glu Lys Thr Gly Lys Arg Ser Thr Ser Ala Ala Val Leu Glu Ala Leu
 660 665 670
 Arg Glu Ala His Pro Ile Val Glu Lys Ile Leu Gln Tyr Arg Glu Leu
 675 680 685
 Thr Lys Leu Lys Ser Thr Tyr Ile Asp Pro Leu Pro Asp Leu Ile His
 690 695 700
 Pro Arg Thr Gly Arg Leu His Thr Arg Phe Asn Gln Thr Ala Thr Ala
 705 710 715 720
 Thr Gly Arg Leu Ser Ser Ser Asp Pro Asn Leu Gln Asn Ile Pro Val
 725 730 735
 Arg Thr Pro Leu Gly Gln Arg Ile Arg Arg Ala Phe Ile Ala Glu Glu
 740 745 750
 Gly Trp Leu Leu Val Ala Leu Asp Tyr Ser Gln Ile Glu Leu Arg Val
 755 760 765
 Leu Ala His Leu Ser Gly Asp Glu Asn Leu Ile Arg Val Phe Gln Glu
 770 775 780
 Gly Arg Asp Ile His Thr Glu Thr Ala Ser Trp Met Phe Gly Val Pro
 785 790 795 800

10

20

30

Arg Glu Ala Val Asp Pro Leu Met Arg Arg Ala Ala Lys Thr Ile Asn
805 810 815

Phe Gly Val Leu Tyr Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu
820 825 830

Ala Ile Pro Tyr Glu Glu Ala Gln Ala Phe Ile Glu Arg Tyr Phe Gln
835 840 845

Ser Phe Pro Lys Val Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly
850 855 860

Arg Arg Arg Gly Tyr Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val
865 870 875 880

Pro Asp Leu Glu Ala Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg
885 890 895

Met Ala Phe Asn Met Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys
900 905 910

Leu Ala Met Val Lys Leu Phe Pro Arg Leu Glu Glu Met Gly Ala Arg
915 920 925

Met Leu Leu Gln Val His Asp Glu Leu Val Leu Glu Ala Pro Lys Glu
930 935 940

Arg Ala Glu Ala Val Ala Arg Leu Ala Lys Glu Val Met Glu Gly Val
945 950 955 960

Tyr Pro Leu Ala Val Pro Leu Glu Val Glu Val Gly Ile Gly Glu Asp
965 970 975

Trp Leu Ser Ala Lys Glu
980

10

20

(2) ANGABEN ZU SEQ ID NO: 13:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 66 Basenpaare
- (B) ART: Nucleotid
- (C) STRANGFORM: Einzelstrang
- (D) TOPOLOGIE: linear

30

(ii) ART DES MOLEKÜLS: Sonstige Nucleinsäure

- (A) BESCHREIBUNG: /desc = "oligonucleotide"

(ix) MERKMAL:

- (A) NAME/SCHLÜSSEL: CDS
(B) LAGE:1..66

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 13:

GAA TTC ATG AGG GGC TCG CAT CAC CAT CAC CAT CAC GCT GCT GAC GAT
48

GAC GAT AAA ATG AGG GGC
66

(2) ANGABEN ZU SEQ ID NO: 14:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 20 Aminosäuren
(B) ART: Aminosäure
(D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Protein

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 14:

Met Arg Gly Ser His His His His His Ala Ala Asp Asp Asp Asp
1 5 10 15

Lys Met Arg Gly
20

【図面の簡単な説明】

【図1】Taq DNAポリメラーゼ(M1-P291)大腸菌DNAポリメラーゼ(Y327-H519)Taq DNAポリメラーゼ(E424-E832)のDNA配列:点突然変異A643G;Ile455Val(配列番号1);および対応するアミノ酸配列(配列番号7)。

【図2】Taq DNAポリメラーゼ(M1-P291)大腸菌DNAポリメラーゼ(Y327-G544)Taq DNAポリメラーゼ(V449-E832)のDNA配列(配列番号2);および対応するアミノ酸配列(配列番号8)。

【図3】Taq DNAポリメラーゼ(M1-P291)Tne DNAポリメラーゼ(P295-E485)Taq DNAポリメラーゼ(E424-E832)のDNA配列;サイレント突然変異A1449C(配列番号3);および対応するアミノ酸配列(配列番号9)。

【図4】Taq DNAポリメラーゼ(M1-P291)Tne DNAポリメラーゼ(P295-G510)Taq DNAポリメラーゼ(V449-E832)のDNA配列;サイレント突然変異C1767T(配列番号4);および対応するアミノ酸配列(配列番号10)。

【図5】Taq DNAポリメラーゼ(1-291)Pfu DNAポリメラーゼ(H103-S334)Taq DNAポリメラーゼ(E424-E832)のDNA配列(配列番号5);および対応するアミノ酸配列(配列番号11)。

【図6】Taq DNAポリメラーゼ(1-291)Pfu DNAポリメラーゼ(V100-F389)Taq DNAポリメラーゼ(-V449-E832)のDNA配列(配列番号6);および対応するアミノ酸配列(配列番号12)。

【図7】Ni-NTAアガロース上でのドメイン交換変異体TaqEc1の精製。クーマシーブルーで染色した8%ポリアクリルアミドゲル上での分析。

レーン1、8:広範囲タンパク質分子量マーカー(200kDa、116.25kDa、97.4kDa、66.2kDa、45kDa、31kDa)

レーン2:可溶性タンパク質

レーン3:カラム素通り画分

10

20

30

40

50

レーン 4 : 洗浄画分 バッファー B

レーン 5 : 洗浄画分 バッファー A

レーン 6、7 : 溶出画分 バッファー C

タンパク質収量 (OD₂₈₀) 約 7 mg。

【図 8】タンパク質純度の測定: SDS-PAGE、Phastシステム (10~15%) : 銀染色 MW: タンパク質分子量マーカー; NHIS-TaqPol: N末端にヒスチジンタグを有する Taq DNAポリメラーゼ; TaqEc1、TaqTne1、TaqTne2: ドメイン交換変異体。

【図 9】種々の温度におけるドメイン交換変異体の比活性。

【図 10】72 での伸長と酵素を連続添加する PCRでのドメイン交換変異体の試験。 10

DNA (左): 標的配列の大きさ = 500 bp

プラスミド pa (右): 標的配列の大きさ = 250 bp

レーン 1: Taq DNAポリメラーゼ (BM Co.), 100 ng, 5 ユニット

レーン 2: ドメイン交換変異体 TaqEc1, 500 ng, 1.25 ユニット/サイクル

レーン 3: ドメイン交換変異体 TaqTne1, 50 ng, 3.6 ユニット/サイクル

レーン 4: ドメイン交換変異体 TaqTne2, 50 ng, 3.5 ユニット/サイクル

III: DNA長標準 III (BM Co.)

VI: DNA長標準 VI (BM Co.)。

結果: ドメイン交換変異体 TaqTne2 を使用した場合、正しい大きさの PCR 産物が形成された。 20

【図 11】55 での伸長と酵素を連続添加する PCRにおけるドメイン交換変異体の試験。 DNA (左): 標的配列の大きさ = 500 bp

プラスミド pa (右): 標的配列の大きさ = 250 bp

レーン 1: ドメイン交換変異体 TaqEc1, 500 ng, 6 ユニット/サイクル

レーン 2: ドメイン交換変異体 TaqTne1, 50 ng, 7.5 ユニット/サイクル

III: DNA長標準 III (BM Co.)

VI: DNA長標準 VI (BM Co.)。

結果: ドメイン交換変異体 TaqEc1 を使用した場合、正しい大きさの PCR 産物が形成された。 30

【図 12】3' - 5' エキソヌクレアーゼ試験変異体 TaqEc1, 72 におけるインキュベーション、プライマー P1。

【図 13】3' - 5' エキソヌクレアーゼ試験変異体 TaqEc1, 50 におけるインキュベーション、プライマー P1 (左)、プライマー P2 (右)。

【図 14】3' ミスマッチプライマーの修正およびそれらの伸長 - 変異体 TaqEc1 (3' ミスマッチプライマー修正アッセイ)

(-) : 制限酵素消化なし

(+) : EcoRIでの制限酵素消化。

【図 15】概略図。

3' 末端におけるプライマーの分解 (3' - 5' エキソヌクレアーゼアッセイ) および 3' ミスマッチプライマーの修正およびそれらの伸長 (3' ミスマッチプライマー修正アッセイ)。 40

【図 16】概略図: 簡易化したフローチャート、3' 末端におけるプライマーの分解並びに 3' ミスマッチプライマーの修正および伸長。

【図 17】Ath、Tne、PolIポリメラーゼ遺伝子およびポリメラーゼキメラの推定遺伝子の CLUSTAL W (1.5) 複数配列アライメント。キメラ配列の Tne 由来部分に下線を付した。

【図 18】A: Tne-Exo および Athポリメラーゼドメインの PCR 増幅に用いたプライマーの構造。

B: 選択された交差点を示した 2つのポリメラーゼのアミノ酸配列アライメントの一部 50

。 C : ハイブリッドポリメラーゼ遺伝子の構築のために設計されたプライマーのヌクレオチド配列および位置。標的配列に相補的ではないプライマーの配列は、小文字で示す。T N E L O WおよびA T H U Pプライマーにおける相補的な「重複」配列には二重線を付した

。 【図19】 A : 2つのポリメラーゼのドメインと共にスプライシングに用いた相同領域を示す、A t hおよびT n eアミノ酸配列のアライメントの一部分。

B : 2つのポリメラーゼのスプライシング領域中のヌクレオチド配列およびアミノ酸配列。図は、T n e DNA配列中の単一のB a m H I切断部位、およびB a m H I切断部位をA t hポリメラーゼに導入するために構築された2つのオリゴ配列を示す。

10

【図20】ポリメラーゼキメラ遺伝子の構築（実施例8も参照のこと）。

【図21】組換えDNAポリメラーゼの3' - 5'エキソヌクレアーゼ活性。

1 : H i n d I I Iで加水分解した ファージのDNA

2 : H i n d I I Iで加水分解した ファージのDNA、d N T P、および組換えDNAポリメラーゼ

3 : H i n d I I Iで加水分解した ファージのDNA、d N T Pなし、組換えDNAポリメラーゼあり

4 : H i n d I I Iで加水分解した ファージのDNA。

【図22】組換えポリメラーゼH Y BおよびH Y B d 5の逆転写酵素活性。大腸菌 B L 2 1 (D E 3) p l y s S x p E T H Y B rおよび大腸菌 B L 2 1 (D E 3) p l y s S x p E T H Y B R d 5からの抽出物(2 μ l)のDNAポリメラーゼ活性を、0 . 0 5ユニットの精度で測定した。この量を使用して、ハイブリッドポリメラーゼの逆転写酵素活性、および1 mMマンガニオンまたは4 mMマグネシウムイオンの作用を測定した。対照を、マンガニ依存性逆転写酵素としてT t h (0 . 2 5ユニット)、およびマグネシウム依存性逆転写酵素としてC . t h e r m . ポリメラーゼ (R o c h e M o l e c u l a r B i o c h e m i c a l s) とした。

20

【 図 1 】

配列番号 1

DNA 配列

1 ATGAGGGGCT CGCATCACCA TCACCATCAC GTCGCTGAGC ATGACGATAA
51 AATGAGGGGC ATGCTACCCG TATTTGAGCC CAAGGGCCGG GTCCCTCTGG
101 TCGACGGCCA CCACCTGGCC TACCAGCCTC TCCACGCCCT GAAGGCGCTC
151 ACCACGAGCC GGGGGGAGCC GGTGCGAGGC GTCTAGGGCT TCAGCAAGAG
201 CCTCTCAAG GCCTCAAGG AGGACGGGGA GGCCTGGAGG CCTACAGGCG
251 ACAGCAAGGC CCCTCTCTTC GGCCTGGAGG GGCCTGGAGG CCTACAGGCG
301 GGGCGGGGCC CCAGCGCGGA GGCCTGGAGG GGCCTGGAGG CCTACAGGCG
351 GGAGCTGGTG GACTCTCTTC GGCCTGGAGG GGCCTGGAGG CCTACAGGCG
401 AGCGGGACGA CTTCTCTGGC GGCCTGGAGG GGCCTGGAGG CCTACAGGCG
451 TCGAGGTCC CACTCTCTTC GGCCTGGAGG GGCCTGGAGG CCTACAGGCG
501 CGACCGGATC CAGCTCTCTC AGCTCTCTTC GGCCTGGAGG CCTACAGGCG
601 GCCTTGAGCA AAGTACGGCC CTGAGGCGCC ACCAGTGGCC GCATACCGCG
651 GGAGAGAGCC GCGAGGAAGC TTCTGGAGGA GTGGGGGAGC CTGGAAGGCC
701 TCCCTCAGAA CTTGGACCGG CTGAGGCGCC CCATCCGGGA GAAGATCCCTG
751 GCCACATGG ACGATCTGAA GCTCTCTTGG GACTCTCTTC AGCTCTCTTC
801 CGACTCGCCC CTGAGGCGCC ACTCTCGCAA AAGCGGGGAG CCGGACCGGG
851 AGAGGCTTAG GGCCTTTCTG GAGAGGCTTG AGTTTGGCAG CCTCTCTCAC
901 GAGTTGCGCC TTCTGGAAAG CCCCATGAC AACTACGTCA CCATCTCTGA
951 TGAAGAACA CTGAAGACCT GGATTCGAAA GCTGGAAAA GCGCCGGTAT
1001 TTGCATTGTA TACCAGAACG GACAGCCCTG ATACATCTCT TGCTAACCTG
1051 GTCGGGCTTT CTATTGCTAT CGAGCCAGGC GTAGCGGCAT ATATTCCGGT
1101 TGCTCATGAT TATCTTGATG CGCCGATCA AATCTCTCCG GAAGCTGCGC
1151 TCGAGTTGCT AAAACCCGTC CTGGAAGATG AAAAGGCGCT AAGGCGGCTG
1201 CAAAACCTGA AATACGATCG CGGTATTCTG GCGAATCTAG GCAATTGAAT
1251 CGGTGGGATT GCGTTTGATA CCAATCTGGA GTCCTACATT GTTGAACGCG
1301 TTGCGGGGCG TCACGATATG GACAGCCCTG CCGAAGCTTG GTTGAACGCG
1351 AAAACCATCA CTTTGAAGA GATTGCTGGT AAAGGCAAAA ATCAACTGAC
1401 CTTTAAACAG ATTGCCCTCG AAGAAGCCGG ACCTTACGCC CCGAAGGCG
1451 CAGATGTGAC CTTGCAAGTG CACTTGAATA TGTGGGAAA TGTGGCCGGA TCTGCAAAA
1501 CACGAGAGGC TCCTTTGGCT TTACCGGGAG GTGGAGAGGC CCCTTCCCG
1551 TGCTCTGGCC CACATGGAGG CCACGGGGG AGCTCGCCCG CCTCGAGGCC
1601 TCAGGGCCTT GTCCCTGGAG GTGGCCGAGG AGCTCGCCCG CCTCGAGGCC
1651 GAGGCTTCC GCTTGGCCCG CCACCCCTTC AACCTCACT CCGGGACCA
1701 GCTGGAAAGG GTCTCTTTG ACAGCTTAGG GCTTCCCGCC ATCGGAAGA
1751 CCGAAGAGAC CGGCATCTCT TCCACAGGCG CCGCCCTCTC GGAGGCCCTC
1801 CCGAGGCCCC ACGCATCTCT GGCATAGATC GGCAGATACC GTGAGCTCAC
1851 CAAGCTTAGG AGCACTTACA TTGACCCCTT GCGCGACTTC ATCCACCCA
1901 GAGGGGCGCG CCTCCACACC CGCTTCAACC AGACGGCCAC GGCACGGGCC
1951 AGGCTAAGTA GCTCCGATCC CAACCTCCAG AACATCCCGC TCCGCACCCC
2001 GCTTGGGCG AGGATCCGCC GGGCCTTCAT CCGCGAGGAG GGGTGGCTAT
2051 TGGTGGCCCT GGAATATAGC CAGATAGAGC TCAGGGTCTC GGCCCACTTC
2101 TCCGGCGAGC AGAACCCTGAT CCGGGTCTTC CAGGAGGGCC GGGACATCCA
2151 CACGGAGACC GCCAGCTGGA TGTTCGGCCT CCCCAGGAG CCGCTGGACC
2201 CCCTGATGCG CCGGGCGGCC AAGACATCA ACTTCCGGGT CCTCTACGCC
2251 ATGTCGGCCC ACCGCTCTC CAGAGAGCTA GCCATCCCTT ACGAGAGGCC
2301 CCGAGCCCTC ATTGAGCGCT ACTTTCAGAG CTTCCCAAG GTGCGGGCCT
2351 GGATTGAGAA GACCCTGGAG GAGGGCAGGA GCGGGGGGTA CCGTGGAGAC
2401 CTCTTCGGCC GCCCGCGCTA CCGTCCAGAC CTAGAGGCC CCGTGAAGAC
2451 CCGTCCGGGAG GCGGGCGGAG CCGTGGCCTT CACATCGCC GTCCAGGGCA
2501 CCGCCCGCGA CCTCATGAAG GTGCTATGTT TGAAGCTCTT CCCCAGGCTG

2551 GAGGAAATGG GGGCCAGGAT GCTCCTTCAG GTCCACGAGC AGCTGGTCTC
2601 CGAGGCCCCA AAGAGAGGCG CGAGGCCGCT GGCCCGGCTG GCCAAGGAGG
2651 TCATGGAGGG GGTGTATCCC CTGCCCCTGC CCCTGGAGGT GGAGGTGGGG
2701 ATAGGGGAGG ACTGGCTCTC CGCCAAGGAG TGA

配列番号 7

アミノ酸配列

1 MRGSHHHHH AADDDDKMRG MLPLFEPKGR VLLVDGHLLA YRTFHALKGL
51 TTSRGEVQQA VYFAKSLIK ALKEDGDVAI VVFDKAPSF RHEAYGGYKA
101 GRAPTPEDFP RQALIKELV DLLGLARLEV PGYEADVDVA SLAKKAEKEG
151 YEVRLITADK DLYQLLSDRI HVLHPEGYLI TPWLWKEYG LRPDQWADYR
201 ALTGDESIDL PVKVGIGKPT ARKLLIEWGS LEALLKNLDR LKPAIRKIL
251 AHMDDLKLSW DLAKVRTDLP LEVDFAKRRE PDRERAFLE ERLEFGSLHL
301 EFGLESFYD NYVTILDEET LKAWIAKLEK BRLELLEKPL LEDEKALVKG
351 VGLSFAIEPG VAAIIPVADH YLDAPDQISR ERALELLKPL LEDEKALVKG
401 QNLKYDRGIL ANYGIELRGI AFDTMLESYI LNSVAGRDM DLSLAERWLKH
451 KTIITFEIAG KGNQLTFNQ IALEEAGRYA ABDADVTLQL HIKMFPDLQK
501 HERLLWLYRE VERPLSAVLA HMEATGVRLD VAYLRALSLE VAEVYARLEA
551 EVERLAGHPF NLSNRDQLER VLFDELGLPA IGTETKTKR STSAAVLEAL
601 REHRPIVEKI LQYRELKELV STYIDLPLDL IHPRTGRILH RENQATATATG
651 RLSSSDPNLQ NIPVPTPLQG RIRRAFVRE GWLLVALDYS QIELRVLAHL
701 SGENLIRVF QEGRDIHET ASMFMGVPRE AVDFLMRRAA KTFINFGVLYG
751 MSARHLSQEL AIPYEEAQA IERYFQSFPK VRAMIEKTL EGRRRYVET
801 LFRRRYVDP LEARVKSVE AEREMAFNMP VQGTADLMK LAMVKLFPRL
851 EEMGARMLLQ VHDLEVLLEAP KERAEAVALR AKEVMEGVYP LAVPLEVEVG
901 IGDWLSAKE

【 図 2 】

配列番号 2

DNA 配列

1 ATGAGGGGCT CGCATCACCA TCACCATCAC GTCGCTGAGC ATGACGATAA
51 AATGAGGGGC ATGCTACCCG TATTTGAGCC CAAGGGCCGG GTCCCTCTGG
101 TCGACGGCCA CCACCTGGCC TACCAGCCTC TCCACGCCCT GAAGGCGCTC
151 ACCACGAGCC GGGGGGAGCC GGTGCGAGGC GTCTAGGGCT TCAGCAAGAG
201 CCTCTCAAG GCCTCAAGG AGGACGGGGA GGCCTGGAGG CCTACAGGCG
251 ACAGCAAGGC CCCTCTCTTC GGCCTGGAGG GGCCTGGAGG CCTACAGGCG
301 GGGCGGGGCC CCAGCGCGGA GGCCTGGAGG GGCCTGGAGG CCTACAGGCG
351 GGAGCTGGTG GACTCTCTTC GGCCTGGAGG GGCCTGGAGG CCTACAGGCG
401 AGCGGGACGA CTTCTCTGGC GGCCTGGAGG GGCCTGGAGG CCTACAGGCG
451 TCGAGGTCC CACTCTCTTC GGCCTGGAGG GGCCTGGAGG CCTACAGGCG
501 CGACCGGATC CAGCTCTCTC AGCTCTCTTC GGCCTGGAGG CCTACAGGCG
601 GCCTTGAGCA AAGTACGGCC CTGAGGCGCC ACCAGTGGCC GCATACCGCG
651 GGAGAGAGCC GCGAGGAAGC TTCTGGAGGA GTGGGGGAGC CTGGAAGGCC
701 TCCCTCAGAA CTTGGACCGG CTGAGGCGCC CCATCCGGGA GAAGATCCCTG
751 GCCACATGG ACGATCTGAA GCTCTCTTGG GACTCTCTTC AGCTCTCTTC
801 CGACTCGCCC CTGAGGCGCC ACTCTCGCAA AAGCGGGGAG CCGGACCGGG
851 AGAGGCTTAG GGCCTTTCTG GAGAGGCTTG AGTTTGGCAG CCTCTCTCAC
901 GAGTTGCGCC TTCTGGAAAG CCCCATGAC AACTACGTCA CCATCTCTGA
951 TGAAGAACA CTGAAGACCT GGATTCGAAA GCTGGAAAA GCGCCGGTAT
1001 TTGCATTGTA TACCAGAACG GACAGCCCTG ATACATCTCT TGCTAACCTG
1051 GTCGGGCTTT CTATTGCTAT CGAGCCAGGC GTAGCGGCAT ATATTCCGGT
1101 TGCTCATGAT TATCTTGATG CGCCGATCA AATCTCTCCG GAAGCTGCGC
1151 TCGAGTTGCT AAAACCCGTC CTGGAAGATG AAAAGGCGCT AAGGCGGCTG
1201 CAAAACCTGA AATACGATCG CGGTATTCTG GCGAATCTAG GCAATTGAAT
1251 CGGTGGGATT GCGTTTGATA CCAATCTGGA GTCCTACATT GTTGAACGCG
1301 TTGCGGGGCG TCACGATATG GACAGCCCTG CCGAAGCTTG GTTGAACGCG
1351 AAAACCATCA CTTTGAAGA GATTGCTGGT AAAGGCAAAA ATCAACTGAC
1401 CTTTAAACAG ATTGCCCTCG AAGAAGCCGG ACCTTACGCC CCGAAGGCG
1451 CAGATGTGAC CTTGCAAGTG CACTTGAATA TGTGGGAAA TGTGGCCGGA TCTGCAAAA
1501 CACGAGAGGC TCCTTTGGCT TTACCGGGAG GTGGAGAGGC CCCTTCCCG
1551 TGCTCTGGCC CACATGGAGG CCACGGGGG AGCTCGCCCG CCTCGAGGCC
1601 TCAGGGCCTT GTCCCTGGAG GTGGCCGAGG AGCTCGCCCG CCTCGAGGCC
1651 GAGGCTTCC GCTTGGCCCG CCACCCCTTC AACCTCACT CCGGGACCA
1701 GCTGGAAAGG GTCTCTTTG ACAGCTTAGG GCTTCCCGCC ATCGGAAGA
1751 CCGAAGAGAC CGGCATCTCT TCCACAGGCG CCGCCCTCTC GGAGGCCCTC
1801 CCGAGGCCCC ACGCATCTCT GGCATAGATC GGCAGATACC GTGAGCTCAC
1851 CAAGCTTAGG AGCACTTACA TTGACCCCTT GCGCGACTTC ATCCACCCA
1901 GAGGGGCGCG CCTCCACACC CGCTTCAACC AGACGGCCAC GGCACGGGCC
1951 AGGCTAAGTA GCTCCGATCC CAACCTCCAG AACATCCCGC TCCGCACCCC
2001 GCTTGGGCG AGGATCCGCC GGGCCTTCAT CCGCGAGGAG GGGTGGCTAT
2051 TGGTGGCCCT GGAATATAGC CAGATAGAGC TCAGGGTCTC GGCCCACTTC
2101 TCCGGCGAGC AGAACCCTGAT CCGGGTCTTC CAGGAGGGCC GGGACATCCA
2151 CACGGAGACC GCCAGCTGGA TGTTCGGCCT CCCCAGGAG CCGCTGGACC
2201 CCCTGATGCG CCGGGCGGCC AAGACATCA ACTTCCGGGT CCTCTACGCC
2251 ATGTCGGCCC ACCGCTCTC CAGAGAGCTA GCCATCCCTT ACGAGAGGCC
2301 CCGAGCCCTC ATTGAGCGCT ACTTTCAGAG CTTCCCAAG GTGCGGGCCT
2351 GGATTGAGAA GACCCTGGAG GAGGGCAGGA GCGGGGGGTA CCGTGGAGAC
2401 CTCTTCGGCC GCCCGCGCTA CCGTCCAGAC CTAGAGGCC CCGTGAAGAC
2451 CCGTCCGGGAG GCGGGCGGAG CCGTGGCCTT CACATCGCC GTCCAGGGCA
2501 CCGCCCGCGA CCTCATGAAG GTGCTATGTT TGAAGCTCTT CCCCAGGCTG

2551 GAGGAAATGG GGGCCAGGAT GCTCCTTCAG GTCCACGAGC AGCTGGTCTC
2601 CGAGGCCCCA AAGAGAGGCG CGAGGCCGCT GGCCCGGCTG GCCAAGGAGG
2651 TCATGGAGGG GGTGTATCCC CTGCCCCTGC CCCTGGAGGT GGAGGTGGGG
2701 ATAGGGGAGG ACTGGCTCTC CGCCAAGGAG TGA

配列番号 8

アミノ酸配列

1 MRGSHHHHH AADDDDKMRG MLPLFEPKGR VLLVDGHLLA YRTFHALKGL
51 TTSRGEVQQA VYFAKSLIK ALKEDGDVAI VVFDKAPSF RHEAYGGYKA
101 GRAPTPEDFP RQALIKELV DLLGLARLEV PGYEADVDVA SLAKKAEKEG
151 YEVRLITADK DLYQLLSDRI HVLHPEGYLI TPWLWKEYG LRPDQWADYR
201 ALTGDESIDL PVKVGIGKPT ARKLLIEWGS LEALLKNLDR LKPAIRKIL
251 AHMDDLKLSW DLAKVRTDLP LEVDFAKRRE PDRERAFLE ERLEFGSLHL
301 EFGLESFYD NYVTILDEET LKAWIAKLEK BRLELLEKPL LEDEKALVKG
351 VGLSFAIEPG VAAIIPVADH YLDAPDQISR ERALELLKPL LEDEKALVKG
401 QNLKYDRGIL ANYGIELRGI AFDTMLESYI LNSVAGRDM DLSLAERWLKH
451 KTIITFEIAG KGNQLTFNQ IALEEAGRYA ABDADVTLQL HIKMFPDLQK
501 HERLLWLYRE VERPLSAVLA HMEATGVRLD VAYLRALSLE VAEVYARLEA
551 EVERLAGHPF NLSNRDQLER VLFDELGLPA IGTETKTKR STSAAVLEAL
601 REHRPIVEKI LQYRELKELV STYIDLPLDL IHPRTGRILH RENQATATATG
651 RLSSSDPNLQ NIPVPTPLQG RIRRAFVRE GWLLVALDYS QIELRVLAHL
701 SGENLIRVF QEGRDIHET ASMFMGVPRE AVDFLMRRAA KTFINFGVLYG
751 MSARHLSQEL AIPYEEAQA IERYFQSFPK VRAMIEKTL EGRRRYVET
801 LFRRRYVDP LEARVKSVE AEREMAFNMP VQGTADLMK LAMVKLFPRL
851 EEMGARMLLQ VHDLEVLLEAP KERAEAVALR AKEVMEGVYP LAVPLEVEVG
901 IGDWLSAKE

【 図 3 】

配列番号 3

DNA 配列

1 ATGAGGGGCT CGCATCACCA TCACCATCAC GCTGCTGACG ATGACGATAA
51 AATGAGGGGC ATGCTACCCG TATTTGAGCC CAAGGGCCGG GTCTCTCTGG
101 TCGACGGCCA CCACCTGGCC TACCCGACCT TCACGCGCCT GAAGGGCCCTC
151 ACCACCAGCC GGGGGGAGCC GGTTCAGGCG GTCTACGGCT TCGCCAAAGAG
201 CCTCTCAAG GCCTCAAGG AGGACGGGGA CCGGCTGATC GTGGTCTTTG
251 ACCCCAAAGC CCCTCTCTTC GCGCACGGCC CTTACGAGTC GTGCTCTTTG
301 GCGCGGGCCC CCACCCGGGA GACCTTTCCC CGGCAACTCG CCTCATCAA
351 GGAGCTGGTG GACCTCTGGC GGCCTGGGCG CCTCGAGGTC CCGGGCTAGC
401 AGCGGGACGA CGTCTGGGCC AGCCTGGCCA AGAAGGGCGA AAGAGAGGGC
451 TAGAGGTTCC GCATCTCTAC CGCCGACAAA GACCTTTACC AGCTCCTTTC
501 CGACCGCATC CAGCTCTCCG ACCCGAGGCG GTACCTCATC ACCCCGGCCT
551 GGCTTTGGGA AAGTACGGCC CTAGGGCCCG ACCAGTGGCG CACTACCCGG
601 GCGCTGACCC GGGACGAGTC CGACAACTTT CCGGGGTCA AGGGCATCGG
651 GGAGAGAGCC GCGAGGAAGC TTCTGGAGGA GTGGGGGAGC CTGGAAGCCC
701 TCCTCAAGAA CTTGACCCGC CTAGAGCCCG CCATCCGGGA GAAGATCCTG
751 GCGCCATGCG ACATGCTGAA GCTCTCTCGG GACCTGGCCA AGGTGGCCAC
801 CGACCTGGCC CTGGAGGTGG ACTTCGCCAA AAGGGGGGAG CCGACCGGG
851 AGAGGCTTAG GGCTTTCTG GAGAGGCTTG AGTTTGGCAG CTCTCTCCAC
901 GAGTTCGGCC TTCTGGAAG CCCCCTGTT GGATACAGAA TAGTGAAGAA
951 CCTGGTGAAA TTTGAAAAC TCATAGAGAA ACTGAGAGAA TCCCTCTCGT
1001 TCGCCATAGA TCTTGGAGC TCTTCCCTCG ATCCTTTCCA CTGCGCAATT
1051 GTCGGTATCT CTGTGCTTTC CAACCAAAAG GAAGCGTACT ACATACCATT
1101 CCATCATAGA AACGCCCAAG ACCCTGGATG AAAAGAAGTT CTGAAAAGC
1151 TAAAAAATCT CCTGGAGGAC CCGCGAGCAA AGATCGTTGG TCAGAAATTTG
1201 AAAATTCGATT ACAAGGTGTT GATGGTAAAG GGTGTTGAAC CTGTCCCTCC
1251 TCACCTTCGAC ACGATGATAG CGGCTTACCT TCTTGGAGCC AACAAAAGA
1301 AGTTCATCTT GGACGATCTC GCATTGAAAT TTCTGGGATA CAAAATGACC
1351 TCTTACCAGG AACTCATGTC CTCTCTCTCT CCGCTGTTG GTTTCAGTTT
1401 TGCCGATGTT CCGTGTAGAA AAGCAGCGAA CTATTCTGTT GAAGATGCGG
1451 ACATCACCTA CAGACTCTAC AAGTCTCTGA GCTTAAAAC CTACGAGGCA
1501 AGGCTCCTTT GCGTTTACCG GGAGTCTGGG AGGCCCTTTT CCGCTGTCTT
1551 GGCCACACTG GAGGCGACGG GGTGCGCCTT GGACGTGGCC TATCTCAGGG
1601 CCTTGTCCCT GGAGGTGGCC GAGGAGATCG CCGGCTCGA GCGCCGAGTC
1651 TTCCGCTGGG CCGGCACCC CTTCACACTC AACTCCCGGG ACCAGCTGGA
1701 AAGGTCCTCT TTTGACGAGC TAGGGCTTCC CGCCATCGCC AAGACGGAGA
1751 AGACCGGCAA GCGCTTACC AGCGCCCGCC TCTGGAGGCG CCTCCGCGAG
1801 GCGCACCCCA TCGTGGAGAA GATCCTGCAG TACCGGAGC TCCCAAGCT
1851 GAAGAGCACC TACATTGACC CTTGCGCGGA CCTCATCCAC CCCAGGACGG
1901 GCGGCTCCCA CACCCGCTTC AACGAGAGCG CCAGCGCCAC GGGCAGGCTA
1951 AGTAGCTCCG ATCCCAACCT CCAGAACATC CCGCTCCGCA CCGCGTTGG
2001 GCAGAGGATC CCGCGGCGCT TCATCGCGGA GGAGGGGTGG CTATTGGTGG
2051 CCTTGGACTA TAGCCAGATA GAGCTCAGGG TGCTGGCCCA CCTCTCGGG
2101 GACGAGAAC TGATCCGGGT CTTCAGGAGG GCGCGGGACA TCCACACGGGA
2151 GACCGCCAGC TGGATGTTCC GCGTCCCGCC GGAGGCGGTG GACCCCTTGA
2201 TGCGCCGGCC GCGCAGAGCC ATCAACTTCC GGTCTCTCTA CCGCATGTCC
2251 GCGCACCGCC TCTCCAGGGA GCTAGCCATC CCTTACGAGG AGGCCAGGCG
2301 CTTCATTTAG CACTACTTTC AGAGCTTCCC CAAGGTGGGG GCTTGGATTG
2351 AGAAGACCCCT GGAGGAGGGC AGGAGGGGGG GGTACGTGGA GACCTCTTTC
2401 GCGCGCGGCC GCTACGTGCC AGACTAGAG GCCCGGTGGA AGAGCTGCGG
2451 GGAGCGGGCC GAGCGCATGG CCTTCAACAT CCGCGTCCAG GGCACCGCGC
2501 CGACCTCAT GAAGCTGGCT ATGTTGAAGC TCTTCCCGAG CTTGGAGGAA

配列番号 9

アミノ酸配列

1 MRGSHHHHH AADDDDKMRG MLPLFEPKGR VLLVDGHLLA YRTFHALKGL
51 TTSRGEVQQA VYFPAKSLK ALKEDGDAVI VVDFAKAPSF RHEAYGYGKA
101 GRAPTPEDFP RQLALIKELV DLLGLARLEV PGYVADDDVA SLAKKAEKGG
151 YEVRIITADK DLYQLSDRI HVLPPEGYLI TPWLWKEYG LRPDOWADYR
201 ALTGDESDNL PGVKIGEKT ARKLEEWGS LEALLKNLDR LKPAIRKILH
251 AHMDDLKLSW DLAKVRTDLP LEVDFAKRRE PDRERLRAFJ ERLEFGSLH
301 EFGLLSEPPV GYRIVKDLVE FEKLEIKLRE SPSFADIDET SSSDPFDCDI
351 VGIVSFKPK EAYIPLHRR NAQNLDEKVE LKLLKEILED PGAKIVGQNL
401 KFDYKVLVWK GVEPVPPHFD TMAIAYLLEP NEKKFNLLDD ALKFLGYKMT
451 SYQELMSFSS PLFGFSFADV PVEKAANYSC EDADITRYRL KILSLKLEHA
501 FRLAGHPFNL NSRDQLERLV FDELGLPAIG KTEKTKRST SAAVLEALRE
551 AHPIVEKILQ YRELTKLKST YIDPLEDLIH PRTRGLHTRF NQATATAGRL
601 SSSDPNLQNI PVRTPLGQRI RRAFIAREGW LLVALDYSQI ELRVLHLSG
701 DENLIRVQFE GRDHTETAS WMFGVPRVAV DFLMRRRAKT INRGVLYGMS
751 AHRLSQELAI PYEAQAFIE RYFQSFPRKV AWIEKLEEG RRRGVVETLF
801 GRRRYVPDLE ARVKSUREAA ERMAFNMPVQ GTAADLMKLA MVKLFPRLEE
851 MGARMLQVH DELVLEAPKE RAEAVARLAK EVMGVIYPLA VPLEVEVGI
901 EDWLSAKE

【 図 4 】

配列番号 4

DNA 配列

1 ATGAGGGGCT CGCATCACCA TCACCATCAC GCTGCTGACG ATGACGATAA
51 AATGAGGGGC ATGCTACCCG TATTTGAGCC CAAGGGCCGG GTCTCTCTGG
101 TCGACGGCCA CCACCTGGCC TACCCGACCT TCACGCGCCT GAAGGGCCCTC
151 ACCACCAGCC GGGGGGAGCC GGTTCAGGCG GTCTACGGCT TCGCCAAAGAG
201 CCTCTCAAG GCCTCAAGG AGGACGGGGA CCGGCTGATC GTGGTCTTTG
251 ACCCCAAAGC CCCTCTCTTC GCGCACGGCC CTTACGAGTC GTGCTCTTTG
301 GCGCGGGCCC CCACCCGGGA GACCTTTCCC CGGCAACTCG CCTCATCAA
351 GGAGCTGGTG GACCTCTGGC GGCCTGGGCG CCTCGAGGTC CCGGGCTAGC
401 AGCGGGACGA CGTCTGGGCC AGCCTGGCCA AGAAGGGCGA AAGAGAGGGC
451 TAGAGGTTCC GCATCTCTAC CGCCGACAAA GACCTTTACC AGCTCCTTTC
501 CGACCGCATC CAGCTCTCCG ACCCGAGGCG GTACCTCATC ACCCCGGCCT
551 GGCTTTGGGA AAGTACGGCC CTAGGGCCCG ACCAGTGGCG CACTACCCGG
601 GGGAGAGACC GCGAGGAAGC TTCTGGAGGA GTGGGGGAGC CTGGAAGCCC
651 TCCTCAAGAA CTTGACCCGC CTAGAGCCCG CCATCCGGGA GAAGATCCTG
701 GCGCCATGCG ACATGCTGAA GCTCTCTCGG GACCTGGCCA AGGTGGCCAC
801 CGACCTGGCC CTGGAGGTGG ACTTCGCCAA AAGGGGGGAG CCGACCGGG
851 AGAGGCTTAG GGCTTTCTG GAGAGGCTTG AGTTTGGCAG CTCTCTCCAC
901 GAGTTCGGCC TTCTGGAAG CCCCCTGTT GGATACAGAA TAGTGAAGAA
951 CCTGGTGAAA TTTGAAAAC TCATAGAGAA ACTGAGAGAA TCCCTCTCGT
1001 TCGCCATAGA TCTTGGAGC TCTTCCCTCG ATCCTTTCCA CTGCGCAATT
1051 GTCGGTATCT CTGTGCTTTC CAACCAAAAG GAAGCGTACT ACATACCATT
1101 CCATCATAGA AACGCCCAAG ACCCTGGATG AAAAGAAGTT CTGAAAAGC
1151 TAAAAAATCT CCTGGAGGAC CCGCGAGCAA AGATCGTTGG TCAGAAATTTG
1201 AAAATTCGATT ACAAGGTGTT GATGGTAAAG GGTGTTGAAC CTGTCCCTCC
1251 TCACCTTCGAC ACGATGATAG CGGCTTACCT TCTTGGAGCC AACAAAAGA
1301 AGTTCATCTT GGACGATCTC GCATTGAAAT TTCTGGGATA CAAAATGACC
1351 TCTTACCAGG AACTCATGTC CTCTCTCTCT CCGCTGTTG GTTTCAGTTT
1401 TGCCGATGTT CCGTGTAGAA AAGCAGCGAA CTATTCTGTT GAAGATGCGG
1451 ACATCACCTA CAGACTCTAC AAGTCTCTGA GCTTAAAAC CTACGAGGCA
1501 GATCTGGAGA ACGTGTCTTA CAAGATAGAA ATGCCTCTTG TGAGCGTGCT
1551 TGACCGGATG GAAGTGAAG GTGTGCGCCT GGACGTGGCC TATCTCAGGG
1601 CCTTGTCCCT GGAGGTGGCC GAGGAGATCG CCGGCTCGA GCGCCGAGTC
1651 TTCCGCTGGG CCGGCCACCC CTTCACACTC AACTCCCGGG ACCAGCTGGA
1701 AAGGTCCTCT TTTGACGAGC TAGGGCTTCC CGCCATCGCC AAGACGGAGA
1751 AGACCGGCAA GCGCTTACC AGCGCCCGCC TCTGGAGGCG CCTCCGCGAG
1801 GCGCACCCCA TCGTGGAGAA GATCCTGCAG TACCGGAGC TCCCAAGCT
1851 GAAGAGCACC TACATTGACC CTTGCGCGGA CCTCATCCAC CCCAGGACGG
1901 GCGGCTCCCA CACCCGCTTC AACGAGAGCG CCAGCGCCAC GGGCAGGCTA
1951 AGTAGCTCCG ATCCCAACCT CCAGAACATC CCGCTCCGCA CCGCGTTGG
2001 GCAGAGGATC CCGCGGCGCT TCATCGCGGA GGAGGGGTGG CTATTGGTGG
2051 CCTTGGACTA TAGCCAGATA GAGCTCAGGG TGCTGGCCCA CCTCTCGGG
2101 GACGAGAAC TGATCCGGGT CTTCAGGAGG GCGCGGGACA TCCACACGGGA
2151 GACCGCCAGC TGGATGTTCC GCGTCCCGCC GGAGGCGGTG GACCCCTTGA
2201 TGCGCCGGCC GCGCAGAGCC ATCAACTTCC GGTCTCTCTA CCGCATGTCC
2251 GCGCACCGCC TCTCCAGGGA GCTAGCCATC CCTTACGAGG AGGCCAGGCG
2301 CTTCATTTAG CACTACTTTC AGAGCTTCCC CAAGGTGGGG GCTTGGATTG
2351 AGAAGACCCCT GGAGGAGGGC AGGAGGGGGG GGTACGTGGA GACCTCTTTC
2401 GCGCGCGGCC GCTACGTGCC AGACTAGAG GCCCGGTGGA AGAGCTGCGG
2451 GGAGCGGGCC GAGCGCATGG CCTTCAACAT CCGCGTCCAG GGCACCGCGC
2501 CGACCTCAT GAAGCTGGCT ATGTTGAAGC TCTTCCCGAG CTTGGAGGAA

2551 ATGAGGGGCA GATGCTCCT TCAGGTCCAC GACGAGCTGG TCCTCGAGGC
2601 CCAAAAAGAG AGGGCGGAGG CCGTGGCCCG GCTGGCCAA GAGGTCATGG
2651 AGGGGGTGTG TCCCTGGCC GTGCCCTGG AGGTGGAGGT GGGGATAGGG
2701 GAGGACTGGC TCTCCGCCAA GGAGTGA

配列番号 10

アミノ酸配列

1 MRGSHHHHH AADDDDKMRG MLPLFEPKGR VLLVDGHLLA YRTFHALKGL
51 TTSRGEVQQA VYFPAKSLK ALKEDGDAVI VVDFAKAPSF RHEAYGYGKA
101 GRAPTPEDFP RQLALIKELV DLLGLARLEV PGYVADDDVA SLAKKAEKGG
151 YEVRIITADK DLYQLSDRI HVLPPEGYLI TPWLWKEYG LRPDOWADYR
201 ALTGDESDNL PGVKIGEKT ARKLEEWGS LEALLKNLDR LKPAIRKILH
251 AHMDDLKLSW DLAKVRTDLP LEVDFAKRRE PDRERLRAFJ ERLEFGSLH
301 EFGLLSEPPV GYRIVKDLVE FEKLEIKLRE SPSFADIDET SSSDPFDCDI
351 VGIVSFKPK EAYIPLHRR NAQNLDEKVE LKLLKEILED PGAKIVGQNL
401 KFDYKVLVWK GVEPVPPHFD TMAIAYLLEP NEKKFNLLDD ALKFLGYKMT
451 SYQELMSFSS PLFGFSFADV PVEKAANYSC EDADITRYRL KILSLKLEHA
501 FRLAGHPFNL NSRDQLERLV FDELGLPAIG KTEKTKRST SAAVLEALRE
551 AHPIVEKILQ YRELTKLKST YIDPLEDLIH PRTRGLHTRF NQATATAGRL
601 SSSDPNLQNI PVRTPLGQRI RRAFIAREGW LLVALDYSQI ELRVLHLSG
701 DENLIRVQFE GRDHTETAS WMFGVPRVAV DFLMRRRAKT INRGVLYGMS
751 AHRLSQELAI PYEAQAFIE RYFQSFPRKV AWIEKLEEG RRRGVVETLF
801 GRRRYVPDLE ARVKSUREAA ERMAFNMPVQ GTAADLMKLA MVKLFPRLEE
851 MGARMLQVH DELVLEAPKE RAEAVARLAK EVMGVIYPLA VPLEVEVGI
901 EDWLSAKE

【 図 5 】

配列番号 5

DNA 配列

1 ATGAGGGGCT CGCATACCA TCACCATCAC GCTGCTGACC ATGACGATAA
51 AATGAGGGGC ATGCTACCCG TATTTGAGCC CAAGGGCCGG GTCTCCTGCG
101 TCGACGGCCA CCACCTGGCC TACCGCACTT TCCACGCCCT GAAGGGCCCTC
151 ACCACCAAGC GGGGGAGGCC GGTGCAGGCG GTCTACGGCT TCGCCAAGAG
201 CCTCTCAAG CCCTCAAGG AGGACGGGGA CTTCTAGGCT CCGCTATCA
251 AGCCCAAGCG CCCTCTCTTC GGCACGAGG CTTACGGGCG GTACCAAGGCG
301 GGGCCGGGCG CCACCTGGCC GACTTTCCTC CGGCAACTCG CCCTCATCAA
351 GGAGCTGGTG GACTTTCCTC AGCTTGGCCG CTTGAGGATC CCGGGTACG
401 ASGCGGACGA GACTTTCCTC AGCTTGGCCG AGAAGGCGGA AAGAGGAGGC
451 TAGAGGTGTC GCATCTCAC CCCTGAGAAA GACCTTTACC AGCTCTTTTC
501 CGACCGCATC CACTCTCTCC ACCCCGAGGG GTACCTTATC ACCCCGGCCTC
551 GGCTTTGGGA AAGATACGGC CTGAGGCCGC ACCAGTGGGC CGACTACCGG
601 CCCTTGACCG GGGACGAGTC CGACAACCTT CCGGGGTCCA AGGGCATCGG
651 GGAGAGAGCG GCGAGGAAAG TTTCTGGAGG GACCTGGCCA AGGTGCGCAC
701 TCTCAAGAAA CCTGACCGGC CTGAGGCCGC CCATCCGGGA AAGATCTCTG
751 GCCCACTGCG ACGATCTGAA GCTCTCTCTG GACCTGGCCA AGGTGCGCAC
801 CGACTCTGCG CTGAGAGTGG ACTTCCGCAA AAGCGGGGAG CCGGACCGGG
851 AGAGGCTTAG GGCTTTCTTG GAGAGGCTTG AGTTTGGCAG CCTCTCCAC
901 GAGTTCGGCC TTTCTGAAAG CCCCATCCA GCAGTGTGGG ACATCTTCGA
951 ATACGATATT CCTATTGCAA AGAGATACCT ATTCGACAAA GGCTTAATAC
1001 CAATGGAGGG GGAAGAGAGC CTAAGATATC TTGCCCTCGA TATAGAACC
1051 CTCTATCACG AAGAGAGAGA GTTGGGAAA GGCCCAATTA TAATGATTAG
1101 TTAATGAGAT GAAATGAGAG CAAAGGTGAT TACTTGGAAA AACATAGATT
1151 TTCCATACGT TGAGCTGTGA TCAAGCGAGA GAGAGATGAT AAGAGATTT
1201 CTCAGGATTA TCAGGAGAAA GGAATCTGAC ATTATAGTGA CTTATATATG
1251 AGACTCATTC GACTTCCCAT ATTTAGCGAA AAGGCGAGA AAACCTTGGGA
1301 TTAATATTAAC CATTGGAGA GATGGAAGCG AGCCCAAGAT GCAGAGATA
1351 GCGCATATGA CCGCTGTAGA AGTCAAGGGA AGAATACATT TCGACTTGT
1401 TCAATGATAA ACAAGGACAA TAAATCTCCC AACATACACA CTAGAGCGTG
1451 TATATGAAAC AATTTTGGAA AAGCCAAGG AGAAGGTATA CCGGACGGAG
1501 ATAGCAAAAG CCTGGGAAAG TGGAGGAGAC CTTGAGAGAG TTGCCAATA
1551 CTCGATGGAA GATGCAAAAG CAACTTATGA ACTCGGAAA GAATCTCTC
1601 CAATGGAAAT TCAGCTTCA GAGAGGCTCC TTTGGCTTTA CCGGGAGGTG
1651 GAGAGGCCCC TTTCCGCTGT CTTGCGCCAC ATGAGAGCCA CCGGGGTGCG
1701 CCGGACGTTG GCCTATCTCA GGGCTTGTCT CTTGAGGTG GCGGAGGARA
1751 TCGCCCGCCT GAGGCGGAG GTCCTCGCCG TCGCCGGCCA CCGCTTCAAC
1801 CTCACCTCCC GGGACCGACT GGAAGAGGTC GAAAGAGGTC CCGCTTCAAC
1851 TCCCGCATC GGCACGAGTC AAGAGAGGTC GAAAGAGGTC CCGCTTCAAC
1901 CCGTCTGGA GCGTCAACAA GCTGAGAGC ACCTACATTG ACCCTTGCC
2001 GAGTCTATC GACTTCCAGC CCGCCAGGA CCGCCCGCCT CCACACCGA
2051 CCGCCAGCGC CAGGCGAGG CTAAGTAGCT CCGATCCCAA CCTCCAGAAC
2101 ATCCCGTCCC GCACCCCGCT TGGGACAGAG ATCCCGCGGG CCTCATCGC
2151 CGAGGAGGGG TGCTATTGG TGCCCTGGA CTATAGCCAG ATAGAGCTCA
2201 GGGTGTGGC CCACCTCTCC GCGCAGAGA ACCTGATCCG GGTCTCCAG
2251 GAGGGCGGGC ACATCCACAC GAGACCGGCC AGCTGGATGT TCGCGTCCC
2301 CCGGAGGCC GTGACCCCC TGAITGCGCC GCGGCCAAG ACCATCAACT
2351 TCGGGTCTC CTAGCGCATG TCGGCCACC GCCTCTCCCA GGAGCTAGC
2401 ATCCCTTACG AGGAGGCCCA GGCCTTCAIT GAGCGTACT TTCAGAGCTT
2451 CCCAAGGTG CCGGCTCGA TTGAGAAGC CTTGAGGAGG GCGCAGGCG
2501 GGGGTACTT GGAACCCCTC TTCGCGCCG CCGCTACTG GCCAGACTTA

配列番号 1 1

アミノ酸配列

1 MRGSHHHHH AADDDKMRG MLPLFEPKGR VLLVDGHLLA YRTFHALKGL
51 TTSRGEVQA VVGFAKSLK ALKEDGDRAVI VVFDAPKPSF RHEAYGGYKA
101 GRAPTPEPFP RQALIKELV DDLGLARLEV PGYEDADVLA SLAKKAEKEG
151 YEVRILTDK DLYQLSDRI HVLHPEGLI TPWLEWERYG LRPDQWADYR
201 ALTGDESQNL PGVKGIGKKT ARKLEEWGWS LEALLNKDLR LKPAIREKIL
251 AHMDDLKLSV DLAKVRTDLP LEVDFAKRRE PDERLRAFEL ERLEFGSLIH
301 EFLGLESFPR AVVDIFEYDI PFARKYRLDK GLIPLMEPEG LKTLAFDIET
351 LYHGEHEEFG GPIIMISYAD ENEAKVITWK NIDLPFVSVV SSEREMIKRF
401 LRILIREKDP IIVTYNGDSF DFPYLAKRAE KLGKILITGR DGSEPKMQR
451 GDMTAVEVKG RIHFDLVHVI TRTINLPYIT LEAVYEIFLS KPEKQVYADE
501 IAKAWESQEN LERVAKYSME DAKATYELGK EFLMEIQLS ERLLWLYREV
551 ERPLSAVLAH MEATGVRLDV AYLRLALSLEV AELRLARLEA EHVPLAGHPFN
601 LNSRDQLERV LFDELGLPAI GKTEKTKRKS TSAAVLEALR EAHPIVEKIL
651 QVRELTKLKS TYIDPLPLDI HPRTRGLHTR FNQATATAGR LSSSDPNLQN
701 IPVRTPLGQR IRRAFIAEEG WLLVLDLXSO IELRVLHLHS GRENLRVRFQ
751 EGRDIHTETA SWMPGVPREA VDLPMRAAK TINFVGLVNG SAHRLYSQELA
801 IYPEEAQAFI ERYFQSPFKV RAWLEKTLER GRRRRLRVPL FGRHRLVPLD
851 EARVKSVREA ERMFAVMNPV QGTAAIDLML AMVKHVPREI EMGARMQLQV
901 HDLVLLEAPK ERABAVARLA KEVMEGVNPL AVLVEVEVGI GEWDLSAKE*

【 図 6 】

配列番号 6

DNA 配列

1 ATGAGGGGCT CGCATACCA TCACCATCAC GCTGCTGACC ATGACGATAA
51 AATGAGGGGC ATGCTACCCG TATTTGAGCC CAAGGGCCGG GTCTCCTGCG
101 TCGACGGCCA CCACCTGGCC TACCGCACTT TCCACGCCCT GAAGGGCCCTC
151 ACCACCAAGC GGGGGAGGCC GGTGCAGGCG GTCTACGGCT TCGCCAAGAG
201 CCTCTCAAG CCCTCAAGG AGGACGGGGA CCGCGTATC GTGGTCTTTG
251 AGCCCAAGCG CCCTCTCTTC GGCACGAGG CTTACGGGCT CCCTCATCAA
301 GGGCCGGGCG CCACCTGGCC GACTTTCCTC CGGCAACTCG CCCTCATCAA
351 GGAGCTGGTG GACTTTCCTC AGCTTGGCCG CTTGAGGATC CCGGGTACG
401 ASGCGGACGA GACTTTCCTC AGCTTGGCCG AGAAGGCGGA AAGAGGAGGC
451 TAGAGGTGTC GCATCTCAC CCCTGAGAAA GACCTTTACC AGCTCTTTTC
501 CGACTCTGCG CACTCTCTCC ACCCCGAGGG GTACCTTATC ACCCCGGCCTC
551 GGCTTTGGGA AAGATACGGC CTGAGGCCGC ACCAGTGGGC CGACTACCGG
601 CCCTTGACCG GGGACGAGTC CGACAACCTT CCGGGGTCCA AGGGCATCGG
651 GGAGAGAGCG GCGAGGAAAG TTTCTGGAGG GACCTGGCCA AGGTGCGCAC
701 TCTCAAGAAA CCTGACCGGC CTGAGGCCGC CCATCCGGGA AAGATCTCTG
751 GCCCACTGCG ACGATCTGAA GCTCTCTCTG GACCTGGCCA AGGTGCGCAC
801 CGACTCTGCG CTGAGAGTGG ACTTCCGCAA AAGCGGGGAG CCGGACCGGG
851 AGAGGCTTAG GGCTTTCTTG GAGAGGCTTG AGTTTGGCAG CCTCTCCAC
901 GAGTTCGGCC TTTCTGAAAG CCCCATCCA GCAGTGTGGG ACATCTTCGA
951 CATCTTCGAA TACGATATTC TATTTGCAA GAGATACTCT ATGCAACAAG
1001 GCCTAATACC AATGGAGGG GGAAGAGAGC TAAAGATTTT TGCCTTCGAT
1051 ATAGAACCCT TCTATCACGA AAGAGAGAGC TTTGAAAAG GCCCAATAT
1101 AATGATTAGT TCAATGAGAG AAGATGAGC AAGGTTGATT ACTTGGAAA
1151 ACATAGATTC TCCATACGTT GAGGTTGAT CAAGCGAGAG AGAGATGATA
1201 AAGAGATTTT TCAGGATTTT CAGGAGAGAG GATCTGACA TTATGATTAC
1251 TTAATGAGAT GACTCATTCG ACTTCCCATTA TTTAGCGAAA AGGCGAGAAA
1301 AACTTGGGAT TAAATTAACC ATTTGGAAGG ATGGAAGCGA CCGCAAGATG
1351 CAGAGATATG GCGATATGAC GGCTGTAGAA CTCGAAGGAA GAATACATTT
1401 CGACTTGTAT CATGTAATAA CAAGGACAAT AAATCTCCA ACATACACAC
1451 TAGAGGCTCT ATATGAGACA ATTTTGGAAA AGCCAAAGGA GAGGTATAC
1501 GCGCAGAGA TAGCAAAAAG CTGGGAAAGT GGAGAGAAC TTGAGAGAT
1551 TGCCAAATAC TCGATGAAAG ATGCAAAAGC AACTTATGAA CTCGGGAAAG
1601 AATTCCTTCC AATGGAAAAT GAGCTTTCAA GATTAGTTGG ACAACCTTTA
1651 TGGGATGTTT CAAAGTCAA GACAGGAAAC CTTGTAGAGT GGTCTTACT
1701 TAGGAAAGCC TACGAAAAGA ACAGAGTAGC TCCAACAGC CCAAGTGAAG
1751 AGGATATCA AAGAGAGCTC AGGAGAGCTC ACACAGGTGG ATTCGTGGCC
1801 CTGGACGTGG CCTATCTCAG GGCTTGTGCC CTTGAGGTGG CCGAGGAGAT
1851 CGCCCGCCTC GAGGCGGAGG TCTTCCGCTC GCGCCGCCAC CCCTTCAACC
1901 TCAACTCCCG GGACGAGCTG GAAAGGGTCC TCTTTGACGA CTTAGGGCTT
1951 CCGCCCATCG GCAAGAGCGA GAGACCGCGC AAGCGTCCA CCGCCCGCCG
2001 CGCTCTGGAG GCTCTCGCG AGGCTCCACC CACTGTSAG AAGTCTCTG
2051 AGTACCGGAC GCTCTCACG CTGAGAGCA CTTACTTGA CCGCTTSCCG
2101 GACTCATTC ACGGCGAGG TAAGTAGCTC CATCCCAAC CACCAAGACA
2151 GGCCTGGCC ACCGCGAGG GGGCGCCTC CACCCCGCTC TCACCGACAC
2201 TCCCGTCTCC CACCCCGCTT GGGCAGAGA TCCCGCGGCC CTTATCCGCC
2251 GAGGAGGCTG GCTATTGGT GGCCTGGAC TRTAGCCAGA TAGAGCTCAG
2301 GGTGCTGGC CACTCTCCG GCGCAGAGA CTTGATCCG GTCTCCAGG
2351 AAGGCGGGGA CATCCACAG GAGACCGCCA GCTGGATGT CCGCGTCCC
2401 CCGGAGCGCG TGGACCCCTG GATGCGCGG GCGGCCAAG CCATCAACTT
2451 CCGGCTCTC TACGGCATGT GCGCCACCG CTTCTCCAG GAGTAGGCA
2501 TCCCTTACGA GAGGCGCCG GCCTTCAITG AGCGTACTT TCAGACTTTC

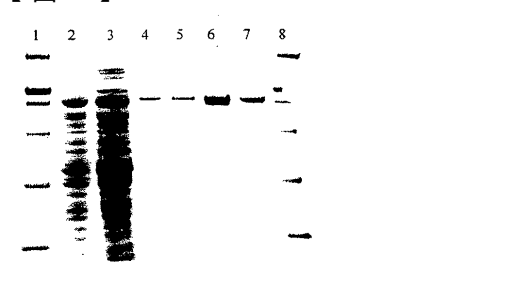
2551 CCCAAGGTG GGCCTGGAT TGAAGAGCC CTGGAGGAG GCGAGGAGCG
2601 GGGTACTGTC GAGACCCCTC TCGCCCGCCG CCCTCATGTC CCGACTTACG
2651 AGCCCGGGT GAGAGCGCTG CCGGAGCGCG CCGAGCCGAT GGCTTCAAG
2701 ATGCCCCGTC AGGACACCGC CCGCACCTC ATAGAGGGA CTAATGGTAA
2751 GCTCTTCCC AGGCTGGAG AATGGGGGGC CAGGATCCCT CTTAGGTC
2801 ACGACGAGCT GGCTCTGAG GCGCCAAAG AGAGGGCGCC CCGCTGGCC
2851 CCGCTGGCCA AGGAGTCAAT GGAGGGGGTG TATCCCTCG CCGTGGCCCT
2901 GGAGGTGGAG GTGGGGATAG GGGAGGACTG GCTCTCCCG AAGGAGTGA

配列番号 1 2

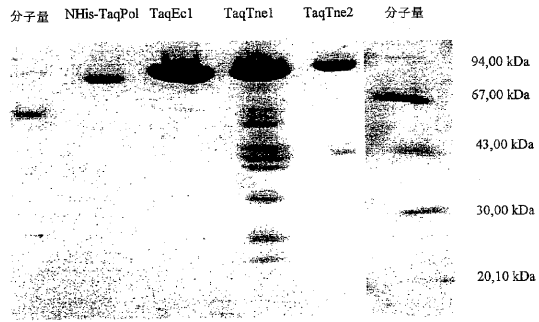
アミノ酸配列

1 MRGSHHHHH AADDDKMRG MLPLFEPKGR VLLVDGHLLA YRTFHALKGL
51 TTSRGEVQA VVGFAKSLK ALKEDGDRAVI VVFDAPKPSF RHEAYGGYKA
101 GRAPTPEPFP RQALIKELV DDLGLARLEV PGYEDADVLA SLAKKAEKEG
151 YEVRILTDK DLYQLSDRI HVLHPEGLI TPWLEWERYG LRPDQWADYR
201 ALTGDESQNL PGVKGIGKKT ARKLEEWGWS LEALLNKDLR LKPAIREKIL
251 AHMDDLKLSV DLAKVRTDLP LEVDFAKRRE PDERLRAFEL ERLEFGSLIH
301 EFLGLESFPR EHPAVDIFE YDIFPARKYL IDKGLIPMEG EEBELKILAF
351 IETLYHEGEE TVCSSTGSI YADENEAKVI TWINIDLPV EHVSSEREMI
401 KRFLRIIREK DFDIIVTYNG DSDFPFYLLAK RAEKLGKILK IGRDSEPKRM
451 QRIQDMTAVE VKGRIFHDLY HVITRTINLP TYTLEAVVEA IFGKPKQVFR
501 ADEIAKAWES GENLERVAKY SMEDAKATYE LGKEFLMEI QLSRLVQDPL
551 WDVSRSTGN LVWVFLRKA YERNEVAFNK PSESEYQRRL RESYTFQVFR
601 LDVYLRALS LEVAEILARL EAEVFLAGH PFLNLSRDLQ ERVLFDELGL
651 PAIKTEKTRG KRSTSAVLE ALREAHPIVE KILQYRELTK LKSTYIDPLF
701 DLHPRTRGL HTRFNQATA TGRSSSDPN LQNIQVPTPL QQRIRRFVFI
751 EEWLVLVALD YSGLERLRLA HLSGDENLIR VPOEGRDHTI ETASWMPGFE
801 REAVDPLMR AAKTINFGVL YGMSAHLRSQ ELAIPEYEAQ AFETERYQSF
851 PKVRAWIEKT LEEGRRRYV ETLFGRYRYV PDLERARVKS REAERMAFN
901 MPVQGTAADL MRLAMVLFKP RLEEMGARLS LQVHLDLVEI APRERAEAVA
951 RLAKVMEVGV YFLAVPLEVE VIGIGEDWLS KE*

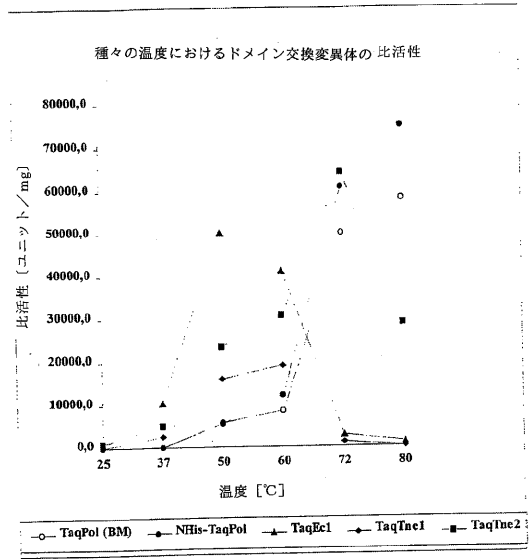
【 図 7 】



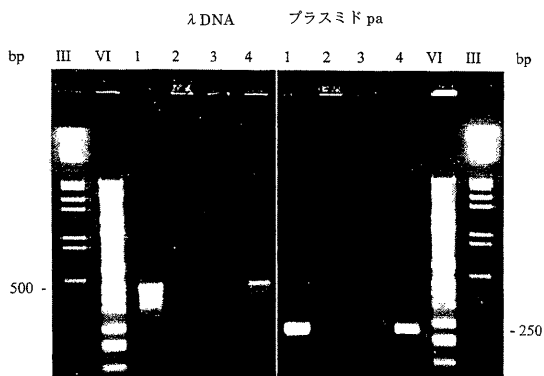
【 図 8 】



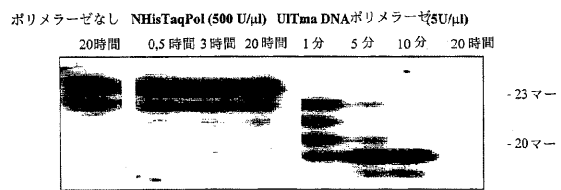
【 図 9 】



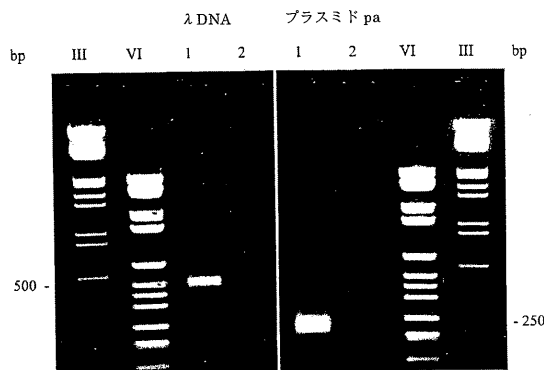
【 図 10 】



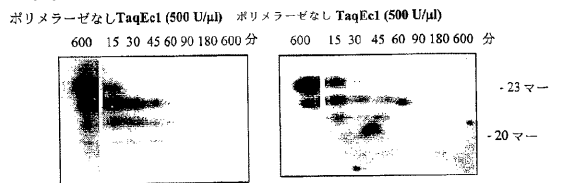
【 図 12 】



【 図 11 】



【 図 13 】



キメラ
 tne.rse
 ath.rse
 DP01_EC0LI

キメラ
 tne.rse
 ath.rse
 DP01_EC0LI

キメラ
 tne.rse
 ath.rse
 DP01_EC0LI

キメラ
 tne.rse
 ath.rse
 DP01_EC0LI

キメラ
 tne.rse
 ath.rse
 DP01_EC0LI

キメラ
 tne.rse
 ath.rse
 DP01_EC0LI

キメラ
 tne.rse
 ath.rse
 DP01_EC0LI

キメラ
 tne.rse
 ath.rse
 DP01_EC0LI

キメラ
 tne.rse
 ath.rse
 DP01_EC0LI

【 図 19 】

配列番号 3 4
 配列番号 3 5
 配列番号 3 6
 交差点 2
 キメラ
 tne.rse
 ath.rse
 DP01_EC0LI

TNE ポリメラーゼのヌクレオチド配列 1642~1689
 配列番号 3 7
 配列番号 3 8

Bam HI 部位
 1642 TCA CCG AAG CAG GTT TCA AGG ATC CTT TTT GAA AAA CTC GGC ATA AAA 1689
 548 S P K Q V S R L L F E K L G I K 563

配列番号 3 9
 配列番号 4 0
 配列番号 4 1
 配列番号 4 2

ATH ポリメラーゼのヌクレオチド配列 1513~1560
 1513 TCA CCG AAA CAG CTT TCT TAC ATT TTG TTT GAA AAG CTA AAA CTT CCT 1560
 505 S P K Q L S Y I L F E K L K L P 520
 5' CA CCG AAA CAG CTT TCT agg acc ctg TTT GAA AAG CTA AAA CTT CCT G 3'
3'GT GGC TTT GTC GAA AGA Ccc taq gAc AAA CTT TTC GAT TTT GAA GGA C 5'
 <-.....m2

【 図 18 】

配列番号 1 9
 配列番号 2 0
 配列番号 2 1
 配列番号 2 2
 TNE UP 5' CTG ACC ATG GCG AGA CTA TTT CTC TTT G -3'
 TNE LOW 5' TCT GTC GAC CTT CAC ACC GTT CAG TTC CAT CC -3'
 ATH UP 5' - AAG GTC GAC AGA GAT GGC CTC ATC CAA TAT ACC -3'
 ATH LOW 5' - TAG CAA GCT TCT ATT TTG TCT CAT ACC AGT -3'
 交差点 1
 A.

配列番号 2 3
 配列番号 2 4
 配列番号 2 5
 キメラ
 tne.rse
 ath.rse

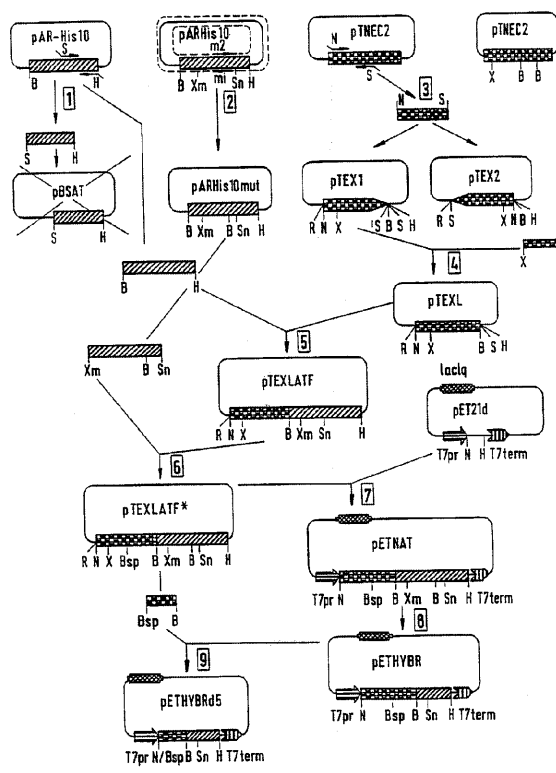
配列番号 1 9
 配列番号 2 6
 配列番号 2 7
 5' ctg acc atg gcg aga cta ttt ctc ttt g -3'
 TNEUP |-----|
 ATG GCG AGA CTA TTT CTC TTT GAT GGA 27
 M A R L F L F D G 9

1
 配列番号 2 8
 配列番号 2 9
 配列番号 2 0
 配列番号 2 1
 配列番号 3 0
 配列番号 3 1

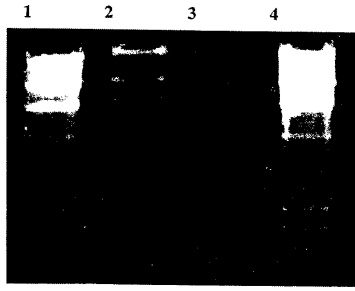
1512 CCG ATG GAA CTG AAC GGT GTG TAC GTG GAC ACA GAT TTC CTG AAG AAA CTC 1563
 505 R M E L N G V Y V D T E F L K R L 521
 3'CC CAT CTT GAC TTG CCA CAC ctc Cag ctg tct 5'
 <-----| TNELOW
 [Sal I 部位]
 ATHUP |-----|
 5' AAG Gtc Gac AGA GAT GGC CTC ATC CAA TAT ACC -3'
 1387 ATG GAA AAA ACA GGA TTT AAG GTG GAT AGA GAT GCC CTC ATC CAA TAT ACC 1435
 463 M E K T G F K V D R D A C L I Q Y T 479

配列番号 3 2
 配列番号 3 3
 配列番号 2 2
 2526 GGA CTG AAC TGG TAT GAG ACA AAA TAG 2553
 643 G L N W Y E T K *
 3'TG ACC ATA CTC TGT TTT ATC ttcgaacgat 5'
 <-----| ATHLOW

【 図 20 】

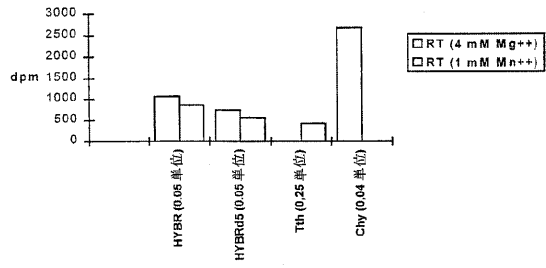


【 2 1 】



【 2 2 】

Tnc/Ath ハイブリッドポリメラーゼと、Tth および C.therm.ポリメラーゼとの
逆転写酵素活性の比較



フロントページの続き

- (72)発明者 ヴィルブランド,ブリッタ
ドイツ連邦共和国 デー - 3 8 1 0 2 ブラウンシュヴェグ,ファサネンシュトラッセ 19
- (72)発明者 ションベルク,ディマー
ドイツ連邦共和国 デー - 5 0 3 7 4 エルフスタット,リチャードシュトラッセ 35
- (72)発明者 ソベック,ハラルド
ドイツ連邦共和国 デー - 8 2 3 7 7 ペンツベルク,バーケンシュトラッセ 29
- (72)発明者 アンケンパウアー,ウォルトラウド
ドイツ連邦共和国 デー - 8 2 3 7 7 ペンツベルク,オベレンジャー 18

審査官 深草 亜子

- (56)参考文献 特開平07 - 147990 (JP, A)
国際公開第96 / 010640 (WO, A1)
米国特許第05489523 (US, A)
J Biol Chem, 1982年, 257, 1958-1964
Nucleic Acids Res, 1993年, 21, 259-265
Gene, 1992年, 112, 133-137
Proc Natl Acad Sci USA, 1995年 9月, 92, 9264-9268

(58)調査した分野(Int.Cl.⁷, DB名)

C12N 15/09 ZNA

C12N 9/12

C07K 19/00