

(12) NACH DEM VERTRAG ÜBER DIE INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES PATENTWESENS (PCT) VERÖFFENTLICHTE INTERNATIONALE ANMELDUNG

(19) Weltorganisation für geistiges Eigentum
Internationales Büro



(43) Internationales Veröffentlichungsdatum
31. Mai 2001 (31.05.2001)

PCT

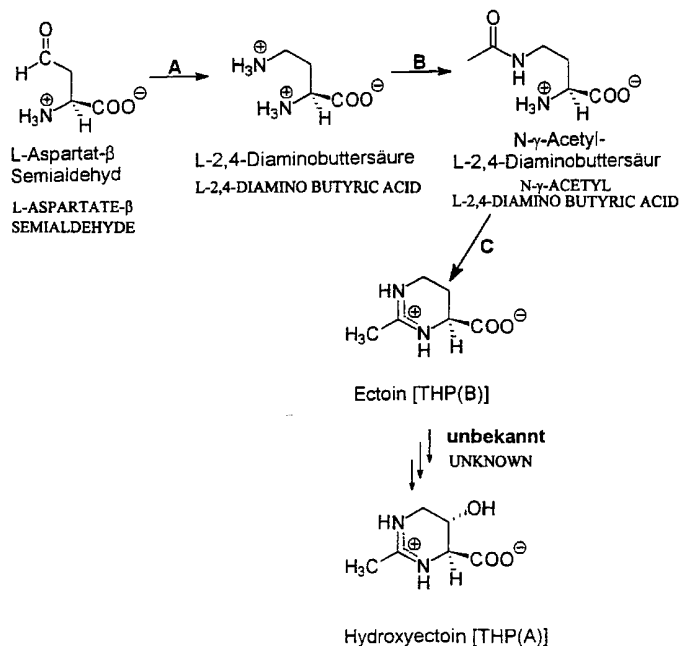
(10) Internationale Veröffentlichungsnummer
WO 01/38500 A2

- (51) Internationale Patentklassifikation⁷: C12N 9/00 [DE/DE]; Neuendorferstrasse 24 a, 16761 Henningsdorf (DE).
- (21) Internationales Aktenzeichen: PCT/DE00/04036
- (22) Internationales Anmeldedatum: 17. November 2000 (17.11.2000)
- (25) Einreichungssprache: Deutsch
- (26) Veröffentlichungssprache: Deutsch
- (30) Angaben zur Priorität: 199 57 470.7 24. November 1999 (24.11.1999) DE
- (71) Anmelder (für alle Bestimmungsstaaten mit Ausnahme von US): ACTINODRUG PHARMACEUTICALS GMBH
- (72) Erfinder; und
- (75) Erfinder/Anmelder (nur für US): KELLER, Ulrich [DE/DE]; Selbitzer Strasse 16c, 14890 Berlin (DE).
- (74) Anwalt: NOBBE, Matthias; Viering, Jentschura & Partner, Centroallee 263, 46047 Oberhausen (DE).
- (81) Bestimmungsstaaten (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

[Fortsetzung auf der nächsten Seite]

(54) Title: TETRAHYDROPYRIMIDINE OXYGENASE GENE, POLYPEPTIDES ENCODED BY SAID GENE AND METHOD FOR PRODUCING THE SAME

(54) Bezeichnung: TETRAHYDROPYRIMIDIN-DIOXYGENASE-GEN, DADURCH KODIERTE POLYPEPTIDE UND VERFAHREN ZUR DEREN HERSTELLUNG



(57) Abstract: The invention relates to enzymes having a tetrahydropyrimidine dioxygenase activity, to the genes encoding said enzyme, to the homologous and heterologous expression of said genes, to methods for producing enzymes having a tetrahydropyrimidine dioxygenase activity and to the use of said enzymes for the in vivo and in vitro production of hydroxylated tetrahydropyrimidine. The invention further relates to the biotechnological use of the hydroxylated tetrahydropyrimidine produced according to the inventive method.

[Fortsetzung auf der nächsten Seite]

WO 01/38500 A2



(84) **Bestimmungsstaaten** (*regional*): ARIPO-Patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), eurasisches Patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), europäisches Patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI-Patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Veröffentlicht:

— *Ohne internationalen Recherchenbericht und erneut zu veröffentlichen nach Erhalt des Berichts.*

Zur Erklärung der Zweibuchstaben-Codes, und der anderen Abkürzungen wird auf die Erklärungen ("Guidance Notes on Codes and Abbreviations") am Anfang jeder regulären Ausgabe der PCT-Gazette verwiesen.

(57) **Zusammenfassung:** Die vorliegende Erfindung betrifft Enzyme mit Tetrahydropyrimidin-Dioxygenase-Aktivität, die dafür kodierenden Gene, die homologe und heterologe Expression dieser Gene, Verfahren zur Herstellung von Enzymen mit Tetrahydropyrimidin-Dioxygenase-Aktivität und deren Verwendung dieser Enzyme zur *in vivo* und *in vitro* Produktion von hydroxyliertem Tetrahydropyrimidin. Zudem betrifft die Erfindung die biotechnologische Verwendung des nach einem der erfindungsgemässen Verfahren hergestellten hydroxylierten Tetrahydropyrimidins.

Tetrahydropyrimidin-Dioxygenase-Gen, dadurch kodierte
Polypeptide und Verfahren zur deren Herstellung

Die vorliegende Erfindung betrifft Enzyme mit
5 Tetrahydropyrimidin-Dioxygenase-Aktivität, die dafür
kodierenden Gene, die homologe und heterologe Expression
dieser Gene, Verfahren zur Herstellung von Enzymen mit
Tetrahydropyrimidin-Dioxygenase-Aktivität und deren Verwendung
dieser Enzyme zur in vivo und in vitro Produktion von
10 hydroxyliertem Tetrahydropyrimidin. Zudem betrifft die
Erfindung die biotechnologische Verwendung des nach einem der
erfindungsgemässen Verfahren hergestellten hydroxylierten
Tetrahydropyrimidins.

15 In Folge von osmotischem Stress werden von Mikroorganismen
niedermolekulare intrazelluläre Verbindungen gebildet, die als
Osmolyte bezeichnet werden und zu denen auch die
Tetrahydropyrimidine zählen. Die Tetrahydropyrimidine, THP(B)
(2-Methyl-4-carboxy-3,4,5,6-tetrahydropyrimidin) und THP(A)
20 (2-Methyl-4-carboxy-5-hydroxy-3,4,5,6-tetrahydropyrimidin)
werden von verschiedenen Mikroorganismen als intrazelluläre
Komponenten des Cytoplasmas gebildet [Ventosa, A. et al.
(1998). *Biology of moderately halophilic aerobic bacteria.*
Microbiol. Mol. Biol. Rev. **62**, 504-544.] (Abb.1). Zu diesen
25 Mikroorganismen gehören die halophilen Eubakterien,
Actinomyceten, Bacillusarten und Brevibakterien. Die THPs
können aus diesen Organismen in reiner Form isoliert werden
und in vielfältiger Form in der Biotechnologie eingesetzt
werden, wie z.B. bei der Stabilisierung von Proteinen sowohl
30 in Lösung als auch bei deren Gefriertrocknung [Lippert, K. und
Galinski, E.A. (1992). *Enzyme stabilisation by ectoine-type
compatible solutes: protection against heating, freezing and
drying.* *Appl. Microbiol. Biotechnol.* **37**, 61-65.], bei der
Beeinflussung der Wechselwirkung von Proteinen und
35 Nukleinsäuren [Lapidot, A. et al. (1995). *Tetrahydropyrimidine
derivatives inhibit binding of a Tat-like, arginine-containing
peptide, to HIV TAR RNA in vitro.* *FEBS Lett.* **367**, 33-38.;

Malin, G. und Lapidot, A. (1996). Induction of synthesis of tetrahydropyrimidine derivatives in *Streptomyces* strains and their effect on *Escherichia coli* in response to osmotic and heat stress. *J. Bacteriol.* **178**, 385-395.; Malin, G. et al. (1999). Effect of tetrahydropyrimidine derivatives on protein-nucleic acids interaction. Type 11 restriction endonucleases as a model system. *J. Biol. Chem.* **274**, 6920-6929.] und bei der Rückfaltung von denaturierten Proteinen. Darüberhinaus kann durch Zugabe von THP(A) bzw. THP(B) zu Kulturmedien von verschiedenen Prokaryonten, wie z.B. *E. coli* eine erhöhte Salz- und Hitzetoleranz festgestellt werden [Malin, G. und Lapidot, A. (1996). Induction of synthesis of tetrahydropyrimidine derivatives in *Streptomyces* strains and their effect on *Escherichia coli* in response to osmotic and heat stress. *J. Bacteriol.* **178**, 385-395.] .

Der charakterisierte Biosyntheseweg von THP(A), gezeigt in Abb. 1, startet von dem L-Aspartat- β -semialdehyd. In der nachfolgenden Transaminierungsreaktion katalysiert die L-2,4-Diaminobuttersäure-Transaminase die Bildung von L-2,4-Diaminobuttersäure (DABA). Die nachfolgende N-Acetylierung von DABA erfolgt durch die L-2,4-Diaminobuttersäure-Acetyltransferase unter Verwendung von Acetyl-Coenzym A. Durch die N-Acetyl-(-L,2,4-Diaminobuttersäure-Cyclase katalysierte intramolekulare Kondensationsreaktion entsteht THP(B) [Peters, P. et al. (1990). The biosynthesis of ectoine. *FEMS Microbiol. Lett.* **71**, 157-162. ;, Ono, H. et al. (1999). Characterization of biosynthetic enzymes for ectoine as a compatible solute in a moderately halophilic eubacterium, *Halomonas elongata*. *J. Bacteriol.* **181**, 91-99.]. Demgegenüber verblieben die Schritte zur Biosynthese des 5-Hydroxyderivates von THP(B), dem (THP(A)) weiterhin ungeklärt.

Die Herstellung von THPs nach den im Stand der Technik bekannten Verfahren ist auf die aufwendige Extraktion dieser

Verbindungen aus den Zellen oder Kulturfiltraten von THP(A) und/oder THP(B) produzierenden Bakterien beschränkt. Bedingt durch die hohen Salzkonzentrationen im Kulturmedium sind diese Verfahren besonders kostenintensiv.

5

Eine chemische Synthese von THPs ist zwar möglich, jedoch umfasst diese Synthese eine Vielzahl von Syntheseschritten und den Einsatz einer aufwendigen Schutzgruppentechnik, so dass dieses komplexe Verfahren, insbesondere im großen Maßstab, nicht wirtschaftlich durchzuführen ist. Dies betrifft im besonderen die Synthese von THP(A), da die jeweiligen Vorstufen auf chemischen Wege nicht erhältlich sind und folglich die chemische Synthese von THP(A) zusätzlich erschwert ist.

10

15

Die Aufgabe der Erfindung besteht daher darin, ein verbessertes Verfahren zur Herstellung von hydroxilierteTHPs, insbesondere THP(A) bereitzustellen.

20

Die Aufgabe der Erfindung wird gelöst durch Bereitstellung eines Verfahrens unter Verwendung eines Polypeptides mit einer Tetrahydropyrimidin-Dioxygenase-Aktivität.

25

Ein Aspekt der Erfindung betrifft daher Nukleinsäuresequenzen, die für Polypeptide kodieren, die eine Tetrahydropyrimidin-Dioxygenase-Aktivität aufweisen. Diese Nukleinsäuresequenzen umfassen sowohl DNA- als auch RNA-Nukleinsäuresequenzen und sind aus der folgenden Gruppe von Nukleinsäuresequenzen ausgewählt:

30

(a) DNA-Nukleinsäuresequenz mit der in SEQ. ID. NO.: 1 dargestellten Nukleotidsequenz;

(b) DNA-Nukleinsäuresequenz, die sich als Ergebnis des degenerierten genetischen Codes von der in SEQ. ID. NO.: 1 dargestellten Nukleotidsequenz ableitet;

35

(c) DNA-Nukleinsäuresequenz, die mit mindestens einer DNA-Nukleinsäuresequenz, die eine DNA-Nukleinsäuresequenz gemäß (a) oder (b) aufweist, hybridisiert;

(d) DNA-Nukleinsäuresequenz, die Fragmente, allelische oder andere Variationen der DNA-Nukleinsäuresequenz mit der in SEQ. ID. NO.: 1 dargestellten Nukleotidsequenz, ist;

(e) RNA-Nukleinsäuresequenz, die von den DNA-
5 Nukleinsäuresequenz gemäß (a) bis (d) abgeleitet ist; und

(f) Nukleinsäuresequenz mit einem Grad der Homologie zu einer der Sequenzen gemäß (a) bis (e) von mehr als 60%.

Mit den für Polypeptide mit Tetrahydropyrimidin-Dioxygenase-
10 Aktivität kodierenden DNA-Nukleinsäuresequenzen sind sowohl genomische DNA-Nukleinsäuresequenzen als auch cDNA-Nukleinsäuresequenzen gemeint. Der Grad der Homologie der Nukleinsäuresequenz gemäß f) zu einer der Sequenzen gemäß (a) bis (e) beträgt mindestens 60%, bevorzugt 80% und besonders
15 bevorzugt mehr als 90%.

Ein weiterer Aspekt der vorliegenden Erfindung betrifft Polypeptide, die eine Tetrahydropyrimidin-Dioxygenase-Aktivität aufweisen und die bevorzugt durch eine der
20 Nukleinsäuresequenzen gemäß (a), (b), (c), (d) (e) oder (f) kodiert sind. Besonders bevorzugt ist ein Polypeptid (SEQ. ID. NO.: 6), das durch die DNA-Nukleinsäuresequenz gemäß SEQ. ID. NO.: 1 kodiert ist.

25 In einem anderen Aspekt sind DNA- oder RNA-Nukleinsäuresequenzen gemäß (a), (b), (c), (d), (e) oder (f), die aus einem Archaeobakterium, einem Prokaryonten oder einem Eukaryonten isoliert wurden, Gegenstand der vorliegenden Erfindung. Besonders bevorzugt sind derartige
30 Nukleinsäuresequenzen, wenn sie aus einem halophilen Eubakterium, einem Actinomyceten, einem Bazillus oder einem Brevibakterium isoliert wurden.

Bevorzugt sind ebenfalls Nukleinsäurekonstrukte, die eine
35 operative Einheit umfassen, welche aus einer Promotersequenz und einer Nukleinsäuresequenz gemäß (a), (b), (c), (d), (e) oder (f) besteht.

Ebenfalls bevorzugt sind Nukleinsäurekonstrukte, die eine operative Einheit umfassen, welche aus einer Promotersequenz und einer Nukleinsäuresequenz gemäss (a), (b), (c), (d), (e) oder (f) besteht und zusätzlich eine Nukleinsäuresequenz umfassen, die für ein oder mehrere Sekretionssignale kodiert.

In einem weiteren Aspekt betrifft die vorliegende Erfindung replikative, rekombinante Vektoren, die eine Nukleinsäuresequenz gemäss (a), (b), (c), (d), (e) oder (f) umfassen.

Gegenstand der Erfindung sind des weiteren Zellen, die einen in dieser Zelle autonom replizierenden Vektor enthalten, der eine Nukleinsäuresequenz gemäss (a), (b), (c), (d), (e) oder (f) umfasst. Besonders bevorzugt sind dabei Bakterienzellen, Hefezellen oder Pflanzenzellen.

Bevorzugt sind Zellen, die eine durch nicht-natürliche Rekombination in das Genom integrierte Nukleinsäuresequenz gemäss (a), (b), (c), (d), (e) oder (f) aufweisen. Besonders bevorzugt sind dabei Zellen, bei denen die nicht-natürliche Rekombination in das Genom einer Bakterienzelle, einer Hefezelle oder einer Pflanzenzelle erfolgt ist.

In einem alternativen Aspekt betrifft die vorliegende Erfindung die Verwendung von einem Oligonukleotid als Nukleinsäure-Sonden zur Identifikation von Genen, wobei die Nukleotidsequenz des Oligonukleotids von jener der Nukleinsäuresequenzen gemäss (a), (b), (c), (d), (e) oder (f) abgeleitet worden sind und die Oligonukleotide zur Identifizierung von sich in in Archaeobakterien, Prokaryonten oder Eukaryonten befindlichen, chromosomalen oder extrachromosomalen Genen, die für ein Polypeptid kodieren, welches Tetrahydropyrimidin-Dioxygenase-Aktivität besitzt, verwendet werden.

Gegenstand der Erfindung ist auch eine Pflanze, die mit einer rekombinanten DNA, die eine DNA-Sequenz gemäß einem der

Ansprüche 1 bis 4 umfasst, transformiert wurde. Eine solche Pflanze zeichnet sich durch eine gesteigerte osmotische Toleranz aus.

5 Gegenstand der Erfindung ist des weiteren ein Verfahren zur Herstellung eines Polypeptids mit Tetrahydropyrimidin-Dioxygenase-Aktivität, wobei hierzu Zellen, die einen Nukleinsäureabschnitt mit einer Nukleinsäuresequenz gemäss (a), (b), (c), (d), (e) oder (f) oder ein
10 Nukleinsäurenkonstrukt der oben erläuterten Art enthalten, kultiviert werden und das Polypeptid mit Tetrahydropyrimidin-Dioxygenase-Aktivität isoliert wird.

Die für das Isolieren des Polypeptids verwendete Methode wird
15 in einer dem Fachmann bekannten Weise davon abhängen, ob die zur Herstellung verwendeten Zellen einen Nukleinsäureabschnitt mit einer Nukleinsäuresequenz gemäss (a), (b), (c), (d), (e) oder (f) oder ein Nukleinsäurekonstrukt der oben erläuterten Art enthalten. Bei der ersten Möglichkeit kann ein Isolieren
20 des hergestellten Polypeptids erst nach vorgeschaltetem Aufschluß der Zellen erfolgen, wobei bei der zweiten Möglichkeit, bei der die Zellen ein ein Sekretionssignal umfassendes Nukleinsäurenkonstrukt enthalten, das hergestellte Polypeptid direkt aus dem Kulturmedium isolierbar wäre.

25 Ein anderer Aspekt der Erfindung betrifft die Verwendung eines Polypeptids mit Tetrahydropyrimidin-Dioxygenase-Aktivität zur Herstellung von hydroxyliertem Tetrahydropyrimidin.

30 Bevorzugt erfolgt die Herstellung mittels eines Polypeptids mit Tetrahydropyrimidin-Dioxygenase-Aktivität in einer lebenden Zelle, wobei die lebende Zelle eine Bakterienzelle, eine Hefezelle oder eine Pflanzenzelle ist. Die das Polypeptid mit Tetrahydropyrimidin-Dioxygenase-Aktivität enthaltende,
35 lebende Zelle kann sich in einem Kulturmedium befinden, das Tetrahydropyrimidin enthält. In einem solchen Fall ist die lebende Zelle bevorzugt vor der Herstellung zu

permeabilisieren, um ein Inkontaktkommen des Polypeptids mit Tetrahydropyrimidin-Dioxygenase-Aktivität mit dem Tetrahydropyrimidin zu begünstigen.

- 5 Im folgenden wird eine bevorzugte Ausführungsform der Erfindung beschrieben, wobei diese jedoch nicht in der Weise verstanden werden soll, dass daraus eine Einschränkung des vom Erfinder beehrten Schutzzumfangs abgeleitet werden kann.
- 10 Zur Auffindung der THP-Biosynthesegene in *Streptomyces chrysomallus* wurde die L-2,4-Diaminobuttersäure-Acetyltransferase, ein Enzym der THP-Biosynthese aus *Streptomyces chrysomallus* gereinigt. Von tryptischen Peptidsequenzen der L-2,4-Diaminobuttersäure-Acetyltransferase
15 wurden Oligonukleotidsequenzen abgeleitet und für ein Screening einer Cosmidbank eingesetzt. Das Gen der L-2,4-Diaminobuttersäure-Acetyltransferase sowie flankierende Bereiche wurden sequenziert. Auf dem 8,7 kb großen BamHI/EcoRI Fragment wurden u.a. vier offene Leserahmen gefunden. Diese
20 Gene, die für THP-Biosyntheseenzyme kodieren, werden im folgenden als *thpA*, *thpB*, *thpC* und *thpD* bezeichnet (Abb. 2). Der als *thpD* bezeichnete offene Leserahmen kodiert für ein Enzym, das eine THP(B)-Dioxygenase-Aktivität aufweist und ein Molekulargewicht von 32,7 kDa besitzt. Es handelt sich bei
25 diesem Enzym um eine α -Ketoglutarat-abhängige Dioxygenase, die die irreversible Hydroxylierung von THP(B) zu THP(A) katalysiert. Das Gen *thpD* wurde in *E. coli* und *Streptomyces* exprimiert und das entsprechende Protein (THP(D)) nach im Stand der Technik bekannten Methoden gereinigt.
- 30 Die Verwendung des homolog oder heterolog exprimierten THP(D)-Proteins ermöglicht die *in vitro* Produktion von THP(A). Durch Inkubation des THP(D)-Proteins in Gegenwart von THP(B), α -Ketoglutarat, Ascorbinsäure, Eisen(II)sulfat und Katalase wird
35 eine vollständige Hydroxylierung von THP(B) zu THP(A) beobachtet.

Das Gen, das für die Tetrahydropyrimidin-Dioxygenase kodiert, liegt auf einem 8,7 kb großen *BamHI/EcoRI* Fragment. Durch die Expression der auf diesem DNA-Fragment enthaltenen Gene aus *Streptomyces chrysomallus* in Mikroorganismen, kann die
5 Synthese von hydroxyliertem Tetrahydropyrimidin erreicht werden.

Durch Expression des Gens *thpD* in Mikroorganismen, die THP(B) produzieren, wird die *in vivo* Produktion des 5-
10 Hydroxyderivates von Tetrahydropyrimidin (THP(A)) ermöglicht.

Dieses Verfahren ist bei einer Vielzahl von Mikroorganismen, wie z.B. Actinomyceten, Bazillen oder halophilen Bakterien anwendbar, sofern diese Nukleinsäuresequenzen besitzen, die
15 für Polypeptide mit THP-Dioxygenase-Aktivität kodieren. Unter dieser Voraussetzung kann die Expression zudem in jedem anderen prokaryontischen oder eukaryontischen Expressionssystem durchgeführt werden.

20 Im folgenden wird die Erfindung anhand von Beispielen näher beschrieben.

Beispiel 1: Detektion der L-2,4-Diaminobuttersäure-Acetyltransferase aus *Streptomyces chrysomallus*

25 Von einer internalen Peptidsequenz, der aus *Streptomyces chrysomallus* zuvor gereinigten L-2,4-Diaminobuttersäure-Acetyltransferase, wurde eine Nukleotidsequenz abgeleitet (Abb. 3). Diese Nukleotidsequenz wurde radioaktiv markiert und als Sonde zum Screenen einer Cosmidbank eingesetzt. Teile des
30 entsprechenden Cosmids wurden subkloniert und sequenziert. Hierbei wurden u.a. vier offene Leserahmen gefunden, wobei *thpA* für die L-Diaminobuttersäure-Acetyltransferase, *thpB* für die L-Aspartat- β -semialdehyd-Transaminase, *thpC* für die N-Acetyl- γ -L-2,4-Diaminobuttersäure-Cyclase und *thpD* für die
35 THP(B)-Dioxygenase kodiert (Abb. 2). Die Expression in andere

Mikroorganismen ist möglich und als Beispielorganismus dient *Streptomyces lividans* (siehe Beispiel 2).

Beispiel 2: Expression von *thpD* in *Streptomyces lividans*

5 Zur Expression von *thpD* in *Streptomyces lividans* wurde das Gen *thpD* als *SphI/HindIII* Fragment in den Expressionsvektor pSPIJ002 (Abb. 5) kloniert. Das resultierende Expressionsplasmid wurde als pSPIJthpD bezeichnet. Durch PCR Mutagenese wurden die Restriktionsschnittstellen *SphI* und
10 *HindIII* eingeführt (35 Zyklen; 1 min 95°C; 90 sec 55°C; 65 sec 72°C; Primer siehe Tabelle 1). Als Matrize für die PCR diente das Plasmid pQE30thpD. Bei diesem Plasmid handelt es sich um ein pQE30 Derivat, welches das *thpD* Gen als *BamHI/HindIII* PCR-Fragment enthält (35 Zyklen; 1 min 95°C; 90 sec 55°C; 65 sec
15 72°C; Primer siehe Tab.1). Die verwendeten Primer sind in Tabelle 1 als SEQ. ID. NO.: 2-5 gezeigt. Die Ligation und Transformation erfolgte nach Standardmethoden.

Beispiel 3: Reinigung von Hexa-His-THPD aus *Streptomyces lividans*

20 *lividans*
Von einer Vorkultur des *Streptomyces lividans* Stammes (transformiert mit pSPIJthpD, siehe Beispiel 2) wurde YEME-Medium (340 g/l Sucrose, 10 g/l Glucose, 3 g/l Hefeextrakt, 3 g/l Malzextrakt, 5 g/l Bacto-Pepton und 0,2 % MgCl₂) auf eine
25 OD₆₀₀ von 0.05 angeimpft und im Schüttler bei 28°C inkubiert. Nach 72 Stunden wurden die Zellen geerntet. Die Zellen wurden in Aufschlußpuffer (100 mM Phosphatpuffer pH 8.0, 10% Glycerol, 1 mM Benzamidin, 1 mM PMSF, 10 mM Imidazol und 300 mM NaCl) resuspendiert und durch zweimalige Passage durch eine
30 French-Press-Cell (16.000 psi; entspricht 1,105x10⁸ Pa) aufgeschlossen. Der zellfreie Rohextrakt wurde im SS34-Rotor 30 min bei 12000 upm zentrifugiert. Der zellfreie Überstand wurde an einer Ni-NTA-Matrix chromatographiert. Das Hexa-His-THPD Protein eluiert im Bereich von 30-100 mM Imidazol.
35

Beispiel 4: In vitro Synthese von 2-Methyl-4-carboxy-5-hydroxy-3,4,5,6-tetrahydropyrimidin (THP(A))

Die Bestimmung der THP(B)-Dioxygenase-Aktivität erfolgt unter Verwendung des folgenden Reaktionsansatzes:

- 5 Kaliumphosphatpuffer, pH 8 10 mM
 THP(B) 5 mM
 α -Ketoglutarat 5 mM
 Ascorbinsäure 5 mM
 Eisen(II)sulfat 1 mM
10 Hexa-His-THPD variabel

in einem Gesamtvolumen von 100 μ l.

- Die Inkubation erfolgt bei 30°C für eine Stunde. Die Produktanalyse erfolgt durch Chromatographie an einer NH₂-Nucleosil-HPLC-Säule. Die Elution erfolgte mit 70% Acetonitril/H₂O. Das Elutionsprofil ist in Abb. 4 dargestellt. Unter Verwendung von Katalase (Rinderleber, Fa. Sigma) kann eine vollständige Hydroxylierung von THP(B) zu THP(A) erreicht werden.

- 20 Abb. 1: Biosynthese von THP(B) ausgehend von L-Aspartat- β -Semialdehyd. A: L-2,4-Diaminobuttersäure Transaminase, B: L-2,4-Diaminobuttersäure-Acetyltransferase, C: N-Acetyl-Diaminobuttersäure-Cyclase.

- 25 Abb. 2: Restriktionskarte des sequenzierten Bereichs sowie die Lage der aufgefundenen THP-Biosynthesegene. Unter den Restriktionsenzymen ist die entsprechende Basenposition angegeben.

- 30 Abb. 3: Interne tryptische Peptidsequenz der DABA-Acetyltransferase aus *Streptomyces chrysomallus* und davon abgeleitete Oligonukleotidsequenzen (170498A und 170498B).

- 35 Abb. 4: HPLC von THP(A) und THP(B) an einer NH₂-Nucleosil-Säule. Jeweils 1/10 des Reaktionsansatzes, der 5 mM THP(B), 5 mM α -Ketoglutarat, 5 mM Ascorbinsäure, 1 mM FeSO₄ und 2 mg/ml THPD-Fusionsprotein (A) bzw. Puffer D (B) enthielt, wurde aufgetragen. Die Elution erfolgte mit 70% Acetonitril bei einer Flußrate von 1 ml/min.

- Abb. 5: pSPIJ002: Shuttle-Vektor, der durch Ligation von

pSP72 (Bgl11) und plJ702 (Bgl11) entstanden ist. Dieser Expressionsvektor kann sowohl in *E. coli* als auch in *Streptomyces* verwendet werden.

5 **Tab. 1:** Für die Konstruktion von Plasmid pQE30thpD und pSPIJthpD verwendete PCR Primer.

PQE30thpD

FTHPD (30-mer):

10 5'-GCC TGA GGA TCC ATG ACC ACC GAA GTA CGC-3'
(SEQ. ID. NO.: 2)

RTHPD (27-mer):

5'-CCC GCC CGT GAA GCT TGC TAC TTC ACC-3'
(SEQ. ID. NO.: 3)

15

PSPIJthpD

FSPIJthp (30-mer):

5'-GAG GAG GAA TTC AGC ATG CGA GGA TCG CAT-3'
(SEQ. ID. NO.: 4)

20 RSPIJthp (25-mer):

5'-AGT CCA AGC TCA GCT AAT TAA GCT T-3'
(SEQ. ID. NO.: 5)

Patentansprüche

1. Nukleinsäuresequenz, wobei die Nukleinsäuresequenz für ein Polypeptid mit Tetrahydropyrimidin-Dioxygenase-Aktivität
5 kodiert und aus der folgenden Gruppe von Nukleinsäuresequenzen ausgewählt ist:

(a) DNA-Nukleinsäuresequenz mit der in SEQ. ID. NO.: 1 dargestellten Nukleotidsequenz;

10 (b) DNA-Nukleinsäuresequenz, die sich als Ergebnis des degenerierten genetischen Codes von der in SEQ. ID. NO.: 1 dargestellten Nukleotidsequenz ableitet;

(c) DNA-Nukleinsäuresequenz, die mit mindestens einer DNA-Nukleinsäuresequenz, die eine DNA-Nukleinsäuresequenz gemäß
15 (a) oder (b) aufweist, hybridisiert;

(d) DNA-Nukleinsäuresequenz, die Fragmente, allelische oder andere Variationen der DNA-Nukleinsäuresequenz mit der in SEQ.
ID. NO.: 1 dargestellten Nukleotidsequenz, ist;

(e) RNA-Nukleinsäuresequenz, die von einer DNA-Nukleinsäuresequenz gemäß (a) bis (d) abgeleitet ist; und

20 (f) Nukleinsäuresequenz mit einem Grad der Homologie zu einer der Sequenzen gemäß (a) bis (e) von mehr als 60%.

2. Nukleinsäuresequenz gemäß Anspruch 1, wobei die Nukleinsäuresequenz eine genomische DNA-Nukleinsäuresequenz
25 oder eine cDNA-Nukleinsäuresequenz ist.

3. Nukleinsäuresequenz gemäß einem der Ansprüche 1 oder 2, wobei die Nukleinsäuresequenz aus einem Archaeobakterium, einem Prokaryonten oder einem Eukaryonten isoliert wurde.
30

4. Nukleinsäuresequenz gemäß Anspruch 3, wobei die Nukleinsäuresequenz aus einem halophilen Eubakterium, einem Actinomyceten, einem Bazillus oder einem Brevibakterium isoliert wurde.
35

5. Nukleinsäurekonstrukt, wobei das Nukleinsäurekonstrukt

in einer operativen Einheit eine Promotersequenz und eine Nukleinsäuresequenz gemäß einem der Ansprüche 1 oder 2 umfasst.

5 6. Nukleinsäurekonstrukt gemäß Anspruch 5, wobei das Nukleinsäurekonstrukt zusätzlich eine Nukleinsäuresequenz umfasst, die für ein oder mehrere Sekretionssignale kodiert.

10 7. Polypeptid, wobei das Polypeptid durch eine Nukleinsäuresequenz gemäß einem der Ansprüche 1 oder 2 kodiert oder davon abgeleitet ist.

15 8. Polypeptid nach Anspruch 7, wobei das Polypeptid die Aminosäuresequenz SEQ. ID. NO.: 6 hat.

20 9. Replikativer, rekombinanter Vektor, wobei der Vektor eine Nukleinsäuresequenz gemäß einem der Ansprüche 1 oder 2 oder ein Nukleinsäurekonstrukt gemäß einem der Ansprüche 5 oder 6 umfasst.

25 10. Zelle, wobei die Zelle einen Vektor gemäß Anspruch 9 enthält.

30 11. Zelle, wobei die Zelle eine durch nicht-natürliche Rekombination in das Genom integrierte Nukleinsäuresequenz gemäß einem der Ansprüche 1 oder 2 aufweist.

35 12. Zelle gemäß Anspruch 10 oder 11, wobei die Zelle eine Bakterienzelle, eine Hefezelle oder eine Pflanzenzelle ist.

 13. Pflanze, die mit einer rekombinanten DNA, die eine DNA-Sequenz gemäß einem der Ansprüche 1 bis 4 umfasst, transformiert wurde.

 14. Verwendung von einem Oligonukleotid, das von einer Nukleinsäuresequenz gemäß einem der Ansprüche 1 oder 2

abgeleitet ist, als Nukleinsäure-Sonde zur Identifizierung von
chromosomalen oder extrachromosomalen, sich in
Archaeobakterien, Prokaryonten oder Eukaryonten befindlichen
Genen, die für ein Polypeptid kodieren, welches
5 Tetrahydropyrimidin-Dioxygenase-Aktivität besitzt.

10 15. Verwendung von einem Polypeptid mit
Tetrahydropyrimidin-Dioxygenase-Aktivität zur Herstellung von
hydroxyliertem Tetrahydropyrimidin.

16. Verwendung nach Anspruch 15, wobei das Polypeptid die
Aminosäuresequenz SEQ. ID. NO.: 6 hat.

15 17. Verwendung von einer Zelle nach einem der Ansprüche 10
- 12 zur Herstellung von hydroxyliertem Tetrahydropyrimidin.

18. Verfahren zur Herstellung eines Polypeptids mit
Tetrahydropyrimidin-Dioxygenase-Aktivität, wobei

20 Zellen gemäß einem der Ansprüche 10 bis 12, wobei die
Zellen das Polypeptid mit Tetrahydropyrimidin-Dioxygenase-
Aktivität exprimieren, in einem Kulturmedium kultiviert
werden; und

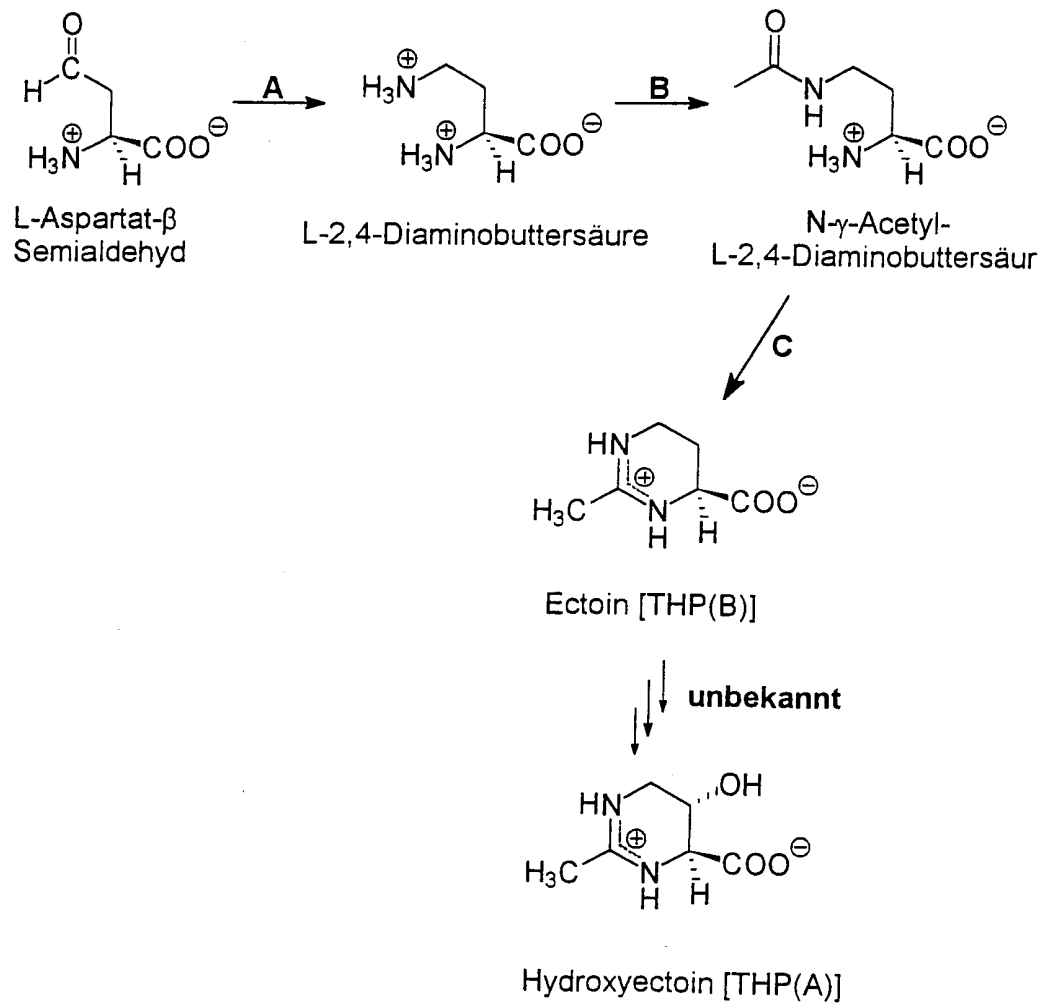
das Polypeptid mit Tetrahydropyrimidin-Dioxygenase-
Aktivität anschließend isoliert wird.

25

19. Verfahren gemäß Anspruch 18, wobei die lebende Zelle
permeabilisiert worden ist und Tetrahydropyrimidin im
Kulturmedium enthalten ist.

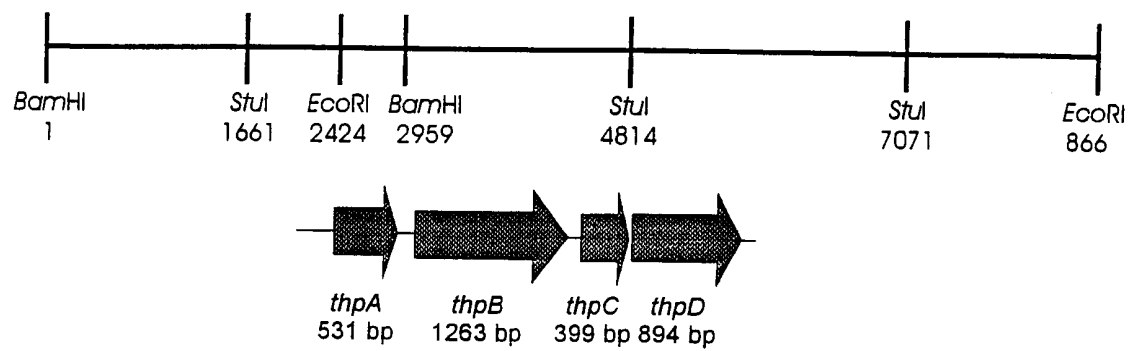
1/5

Abb. 1



2/5

Abb. 2



3/5

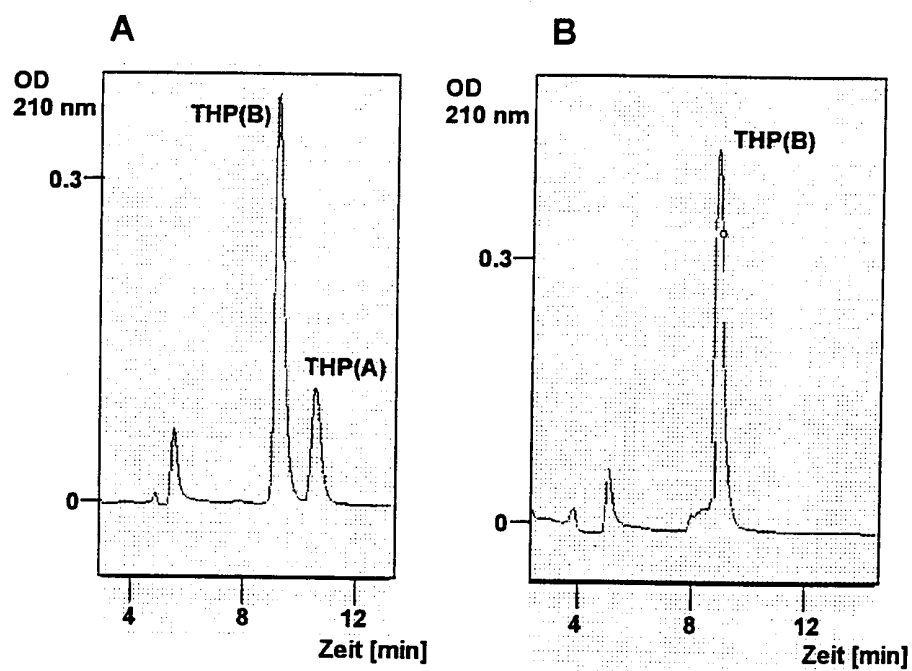
Abb. 3

Peptid Kel15P6: V L F D G E L F P E E T....
170498A: ttc gac ggc gag ctg ttc ccg gag gag acc

Peptid Kel15P6:E E T H L P E V L Y R
170498B: gag gag acc cac ctg ccg gag gtc ctg tac

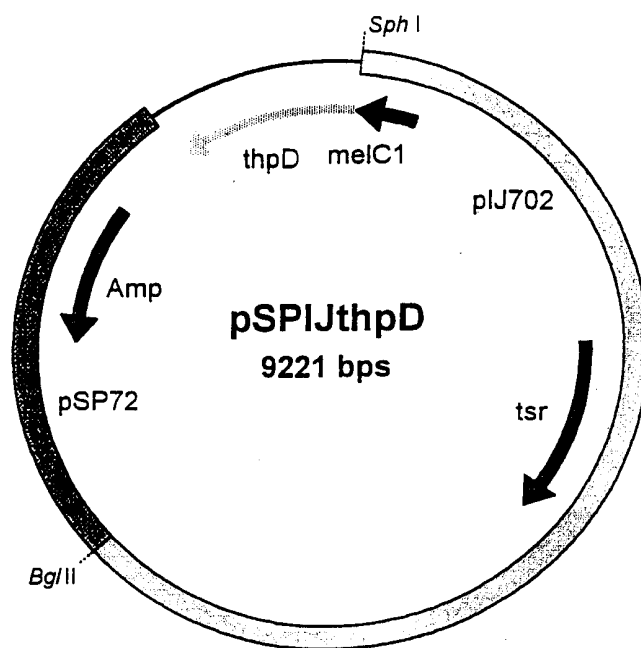
4/5

Abb. 4



5/5

Abb. 5



Sequenzprotokoll

- <110> Keller, Ulrich, Dr.
Grammel, Nicholas, Dr.
- <120> Tetrahydroxypyrimidin-Dioxygenase-Gen, Polypeptide
mit Tetrahydroxypyrimidin-Dioxygenase-Aktivität, Verfahren zur
Herstellung dieser Polypeptide sowie deren Verwendung zur Produktion
von hydroxyliertem Tetrahydroxypyrimidin
- <140> DE 199 57 470.7
- <141> 1999-11-24
- <160> 6
- <210> 1
- <211> 979
- <212> DNA
- <213> Streptomyces chrysomallus
- <400> 1

```

CCCCCGTC ACGGGACGGG AGGACCATGA CGAGAACGGT GTCTACCCAC TGCTGACCGA      60
GGAGGCTGA ACCACC ATG ACC ACC GAA GTA CGC GCC GAT CTG TAC CCC TCG      112
          Met Thr Thr Glu Val Arg Ala Asp Leu Tyr Pro Ser
          1              5              10
CGC GGC GCC GCC GAG ATG ACC ACT CCC CGC CAG GAC CCG GTC ATC TGG TCC      163
Arg Gly Ala Ala Glu Met Thr Thr Pro Arg Gln Asp Pro Val Ile Trp Ser
          15              20              25
GCG CCG GGC GCA CCG GGT CCG GTC GCC GCC AAG GAC CTC CAG GGA TAC GAG      214
Ala Pro Gly Ala Pro Gly Pro Val Ala Ala Lys Asp Leu Gln Gly Tyr Glu
          30              35              40              45
CAC GAC GGC TTC CTC ACC GTC GAC CAG CTC ATC GCC CCG GAC GAG GTC GCC      265
His Asp Gly Phe Leu Thr Val Asp Gln Leu Ile Ala Pro Asp Glu Val Ala
          50              55              60
GTC TAC CAG GCG GAG CTG AAC CGG CTG ATC TCC GAC CCG GCG GTC CGC GCC      316
Val Tyr Gln Ala Glu Leu Asn Arg Leu Ile Ser Asp Pro Ala Val Arg Ala
          65              70              75              80
GAC GAG CGC TCG ATC GTC GAG AAG CAG TCG CAG AAC GTA CGG TCC GTC TTC      367
Asp Glu Arg Ser Ile Val Glu Lys Gln Ser Gln Asn Val Arg Ser Val Phe
          85              90              95
GAG GTC CAC CGG ATC AGC GAG GTC TTC GCC GGT CTG GTC CGC GAC GAG CGG      418
Glu Val His Arg Ile Ser Glu Val Phe Ala Glu Leu Val Arg Asp Glu Arg
          100              105              110
GTG GTG GGC CGG GCC CGC CAG ATC CTC GGC TCG GAC GTG TAC GTC CAC CAG      469
Val Val Gly Arg Ala Arg Gln Ile Leu Gly Ser Asp Val Tyr Val His Gln
          115              120              125              130
TCC CGG ATC AAC GTG AAG CCG GGC TTC GGG GCC ACG GGC TTC TAC TGG CAC      520
Ser Arg Ile Asn Val Lys Pro Gly Phe Gly Ala Thr Gly Phe Tyr Trp His
          135              140              145
TCG GAC TTC GAG ACC TGG CAC GCG GAG GAC GGT CTG CCG AAC ATG CGG ACG      571
Ser Asp Phe Glu Thr Trp His Ala Glu Asp Gly Leu Pro Asn Met Arg Thr
          150              155              160              165
    
```

GTG TCC GTG TCG ATC GCG CTG ACC GAG AAC TTC GAC ACC AAC GGC GGG CTG	622
Val Ser Val Ser Ile Ala Leu Thr Glu Asn Phe Asp Thr Asn Gly Gly Leu	
170 175 180	
ATG ATC ATG CCC GGT TCG CAC AAG ACG TTC CTC GGC TGC GCG GGC GAG ACG	673
Met Ile Met Pro Gly Ser His Lys Thr Phe Leu Gly Cys Ala Gly Glu Thr	
185 190 195	
CCG AAG GAC AAC TAC AAG AAG TCG CTC CAG ATG CAG GAC GCC GGC ACC CCG	724
Pro Lys Asp Asn Tyr Lys Lys Ser Leu Gln Met Gln Asp Ala Gly Thr Pro	
200 205 210 215	
TCC GAC GAG GCG CTG ACG AAG ATG GCC GAC CGC CAC GGC ATC AGG CTC TTC	775
Ser Asp Glu Ala Leu Thr Lys Met Ala Asp Arg His Gly Ile Arg Leu Phe	
220 225 230	
ACG GGC AGG GCC GGT TCG GCG ACC TGG TTC GAC TGC AAC GCC ATG CAC GGC	826
Thr Gly Arg Ala Gly Ser Ala Thr Trp Phe Asp Cys Asn Ala Met His Gly	
235 240 245 250	
TCG GGC GAC AAC ATC ACC CCG TAC GCG CGC AGC AAC GTC TTC ATC GTC TTC	877
Ser Gly Asp Asn Ile Thr Pro Tyr Ala Arg Ser Asn Val Phe Ile Val Phe	
255 260 265	
AAC AGC GTG GAG AAC GCG GCC CAG GAG CCC TTC GCG GCT CCG ATC CGC CGC	928
Asn Ser Val Glu Asn Ala Ala Gln Glu Pro Phe Ala Ala Pro Ile Arg Arg	
270 275 280	
CCC GAG TTC ATC GGG GCG CGG GAC TTC ACC CCG GTG AAG TAG CGGGCCGCA	979
Pro Glu Phe Ile Gly Ala Arg Asp Phe Thr Pro Val Lys *	
285 290 295	

<210> 2
 <211> 30
 <212> DNA
 <213> künstliche Sequenz

<220>
 <223> Als PCR-Primer verwendet, stammend aus dem Plasmid pQEthpD. Das Plasmid pQEthpD ist das Plasmid pQE30, welches das thpD Gen als BamHI/HindIII PCR-Fragment enthält.

<400> 2

GCCTGAGGAT CCATGACCAC CGAAGTACGC 30

<210> 3
 <211> 27
 <212> DNA
 <213> künstliche Sequenz

<220>
 <223> Als PCR-Primer verwendet, stammend aus dem Plasmid pQEthpD. Das Plasmid pQEthpD ist das Plasmid pQE30, welches das thpD Gen als BamHI/HindIII PCR-Fragment enthält.

<400> 3

CCCGCCCGTG AAGCTTGCTA CTTCACC

27

<210> 4
 <211> 30
 <212> DNA
 <213> künstliche Sequenz

<220>
 <223> Als PCR-Primer verwendet, stammend aus dem Plasmid pQEthpD. Das Plasmid pQEthpD ist das Plasmid pQE30, welches das thpD Gen als BamHI/HindIII PCR-Fragment enthält.

<400> 4

GAGGAGGAAT TCAGCATGCG AGGATCGCAT

30

<210> 5
 <211> 25
 <212> DNA
 <213> künstliche Sequenz

<220>
 <223> Als PCR-Primer verwendet, stammend aus dem Plasmid pQEthpD. Das Plasmid pQEthpD ist das Plasmid pQE30, welches das thpD Gen als BamHI/HindIII PCR-Fragment enthält.

<400> 5

AGTCCAAGCT CAGCTAATTA AGCTT

25

<210> 6
 <211> 297
 <212> PRT
 <213> Streptomyces chrysomallus

<400> 6

Met Thr Thr Glu Val Arg Ala Asp Leu Tyr Pro Ser
 1 5 10

Arg Gly Ala Ala Glu Met Thr Thr Pro Arg Gln Asp Pro Val Ile Trp Ser
 15 20 25

Ala Pro Gly Ala Pro Gly Pro Val Ala Ala Lys Asp Leu Gln Gly Tyr Glu
 30 35 40 45

His Asp Gly Phe Leu Thr Val Asp Gln Leu Ile Ala Pro Asp Glu Val Ala
 50 55 60

Val Tyr Gln Ala Glu Leu Asn Arg Leu Ile Ser Asp Pro Ala Val Arg Ala
 65 70 75 80

Asp Glu Arg Ser Ile Val Glu Lys Gln Ser Gln Asn Val Arg Ser Val Phe
 85 90 95

Glu Val His Arg Ile Ser Glu Val Phe Ala Glu Leu Val Arg Asp Glu Arg

	100						105									110	
Val	Val	Gly	Arg	Ala	Arg	Gln	Ile	Leu	Gly	Ser	Asp	Val	Tyr	Val	His	Gln	
115					120					125					130		
Ser	Arg	Ile	Asn	Val	Lys	Pro	Gly	Phe	Gly	Ala	Thr	Gly	Phe	Tyr	Trp	His	
			135					140					145				
Ser	Asp	Phe	Glu	Thr	Trp	His	Ala	Glu	Asp	Gly	Leu	Pro	Asn	Met	Arg	Thr	
	150					155					160					165	
Val	Ser	Val	Ser	Ile	Ala	Leu	Thr	Glu	Asn	Phe	Asp	Thr	Asn	Gly	Gly	Leu	
				170					175					180			
Met	Ile	Met	Pro	Gly	Ser	His	Lys	Thr	Phe	Leu	Gly	Cys	Ala	Gly	Glu	Thr	
		185					190					195					
Pro	Lys	Asp	Asn	Tyr	Lys	Lys	Ser	Leu	Gln	Met	Gln	Asp	Ala	Gly	Thr	Pro	
200					205					210					215		
Ser	Asp	Glu	Ala	Leu	Thr	Lys	Met	Ala	Asp	Arg	His	Gly	Ile	Arg	Leu	Phe	
			220					225					230				
Thr	Gly	Arg	Ala	Gly	Ser	Ala	Thr	Trp	Phe	Asp	Cys	Asn	Ala	Met	His	Gly	
	235					240					245					250	
Ser	Gly	Asp	Asn	Ile	Thr	Pro	Tyr	Ala	Arg	Ser	Asn	Val	Phe	Ile	Val	Phe	
			255						260					265			
Asn	Ser	Val	Glu	Asn	Ala	Ala	Gln	Glu	Pro	Phe	Ala	Ala	Pro	Ile	Arg	Arg	
		270					275					280					
Pro	Glu	Phe	Ile	Gly	Ala	Arg	Asp	Phe	Thr	Pro	Val	Lys					
285					290					295							