

(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
10. Januar 2002 (10.01.2002)

PCT

(10) Internationale Veröffentlichungsnummer
WO 02/02753 A2

- (51) Internationale Patentklassifikation⁷: C12N 9/00 Pfarrer-Engels-Str. 9a, 52428 Jülich (DE). **DEMIR, Ayhan, S.** [TR/TR]; Beril Sitesi 421. Sok. No. 30, Umitkoy, 06530 Ankara (TR).
- (21) Internationales Aktenzeichen: PCT/EP01/07426
- (22) Internationales Anmeldedatum: 29. Juni 2001 (29.06.2001) (74) **Anwalt: FITZNER, Uwe;** Lintorfer Str. 10, 40878 Ratingen (DE).
- (25) Einreichungssprache: Deutsch (81) **Bestimmungsstaat (national):** US.
- (26) Veröffentlichungssprache: Deutsch (84) **Bestimmungsstaaten (regional):** europäisches Patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR).
- (30) Angaben zur Priorität: 100 32 254.9 3. Juli 2000 (03.07.2000) DE **Veröffentlicht:**
— ohne internationalen Recherchenbericht und erneut zu veröffentlichen nach Erhalt des Berichts
— mit dem Sequenzprotokollteil der Beschreibung in elektronischer Form getrennt veröffentlicht; auf Antrag vom Internationalen Büro erhältlich
- (71) Anmelder (für alle Bestimmungsstaaten mit Ausnahme von US): **FORSCHUNGSZENTRUM JÜLICH GMBH** [DE/DE]; Leo-Brandt-Str., 52428 Jülich (DE).
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.
- (71) Anmelder und
(72) **Erfinder: POHL, Martina** [DE/DE]; Malmedyer Str. 102, 52066 Aachen (DE). **MÜLLER, Michael** [DE/DE];



WO 02/02753 A2

(54) **Title:** NUCLEOTIDE SEQUENCE ENCODING BENZALDEHYDE LYASE AND METHOD FOR THE STEREOSELECTIVE SYNTHESIS OF 2-HYDROXY KETONES

(54) **Bezeichnung:** NUKLEOTIDSEQUENZ KODIEREND FÜR EINE BENZALDEHYD-LYASE UND VERFAHREN ZUR STEREOSELEKTIVEN SYNTHESE VON 2-HYDROXYKETONEN

(57) **Abstract:** The invention relates to a nucleotide sequence that encodes a benzaldehyde lyase and to the use thereof in a method for the stereoselective production of compounds that contain 2-hydroxyketone groups.

(57) **Zusammenfassung:** Die vorliegende Erfindung betrifft eine Nukleotidsequenz kodierend für eine Benzaldehyd-Lyase sowie deren Einsatz in einem Verfahren zur stereoselektiven Herstellung von 2-Hydroxyketongruppen enthaltenden Verbindungen.

**Nukleotidsequenz kodierend für eine Benzaldehyd-Lyase und
Verfahren zur stereoselektiven Synthese von 2-Hydroxyketonen**

5 Die vorliegende Erfindung betrifft ein verbessertes Verfahren zur stereoselektiven Synthese von 2-Hydroxyketonen unter Einsatz einer Benzaldehyd-Lyase.

Eine Thiaminpyrophosphat (TPP)-abhängige Benzaldehyd-Lyase sowie
10 genetische Analysen des kodierenden Gens sind bei Vicuna et al. (J. Bacteriol., 1989, 171: 2401-2405) und Hinrichsen, P. et al. (Gene, 1994, 144: 137-138) beschrieben. Eine nähere Charakterisierung des Enzyms zeigte, daß diese Benzaldehyd-Lyase ausschließlich eine irreversible Spaltungsaktivität aufweist. So entstehen beispielsweise ausgehend von
15 Benzoin zwei Moleküle Benzaldehyd bzw. Anisoin wird zu zwei Molekülen Anisaldehyd gespalten. Eine Knüpfung von C-C-Verbindungen katalysiert durch eine Benzaldehyd-Lyase wird explizit ausgeschlossen.

Für eine enzymkatalysierte Synthese von 2-Hydroxyketonen kommen
20 insbesondere Decarboxylasen oder Transketolasen in Betracht, die in der Regel ebenfalls eine Abhängigkeit zu Thiaminpyrophosphat (TPP) aufweisen.

Whitesides et al. (J. Org. Chem., 1992, 57:5889-5907) und Turner et al.
25 (Tetrahedron Asymmetry, 1996, 7: 2185-2188) beschreiben Transketolasen, welche die Spaltung eines 2-Hydroxyketons unter gleichzeitiger Bildung eines anderen 2-Hydroxyketons katalysieren. Wirtschaftliche Produktionsverfahren zur Herstellung von 2-Hydroxyketonen mit Hilfe von Transketolasen sind hier jedoch nicht
30 beschrieben.

Ferner ist die enzymatische Synthese von 2-Hydroxyketonen in Gegenwart einer TPP-abhängigen Pyruvatdecarboxylase beschrieben (DE 195 23 269 und DE 197 36 104). Nachteilig bei den beschriebenen Verfahren ist jedoch, daß die Pyruvatdecarboxylase nur ein stark
5 beschränktes Substratspektrum akzeptiert. Außerdem kann eine spontane Racemisierung der entstehenden Produkte bedingt durch eine auftretende Keto-Enol-Tautomerie mit der Folge einer erniedrigten Enantioselektivität erfolgen.

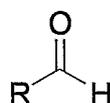
10 Wilcocks et al. (Biotech. Bioeng., 1992, 39: 1058-1063 und Appl. Env. Microbiol., 1992, 58: 1699-1704) beschreiben die Synthese von (S)-2-Hydroxy-1-phenyl-propanon ((S)-2-HPP) mit Hilfe einer Thiaminpyrophosphat-abhängigen Benzoylformiatdecarboxylase ausgehend von Benzoylformiat und Acetaldehyd. Allerdings entsteht auch
15 hier das 2-Hydroxyketon nicht enantiomerenrein, sondern lediglich mit einem Enantiomerenüberschuß von 92 %. Nachteilig ist außerdem die Bildung von Benzylalkohol als Nebenprodukt beim Einsatz ganzer Zellen.

Aufgabe der vorliegenden Erfindung ist daher, ein verbessertes
20 enzymatisches Verfahren zur Herstellung von 2-Hydroxyketonen zur Verfügung zu stellen, das die zuvor genannten Nachteile nicht mehr aufweist.

Diese Aufgabe wird durch das erfindungsgemäße Verfahren zur
25 stereoselektiven Herstellung von 2-Hydroxyketonen gelöst, wobei eine Benzaldehyd-Lyase eingesetzt wird, die in Anwesenheit wenigstens einer lösungsvermittelnden Verbindung die Umsetzung von zwei Aldehydgruppen enthaltenden Verbindungen, von denen wenigstens eine kein Acetaldehyd ist, oder von einer Aldehydgruppen enthaltenden
30 Verbindung mit einer 2-Hydroxyketongruppen enthaltenden Verbindung unter Beibehaltung der stereochemischen Anordnung der letzteren zu

einer weiteren 2-Hydroxyketongruppen enthaltenden Verbindung katalysiert.

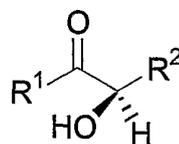
In einer Ausführungsvariante der vorliegenden Erfindung zeichnet sich das
5 zuvor genannte Verfahren dadurch aus, daß eine erste Aldehydruppen
enthaltende Verbindung der allgemeinen Formel I



wobei R = eine Komponente aus der Gruppe aliphatischer, aromatischer
10 oder heteroaromatischer Kohlenwasserstoffe ist, die in ortho-, meta- oder
para-Stellung einfach oder mehrfach substituiert sein kann, wobei die
Substituenten Alkyl-, Aryl- oder Aralkyl-Gruppen oder Heteroatome, wie
z.B. Cl, F, Br oder S, P, N oder Kombinationen davon sein können und R \neq
-CH₃ ist,

15

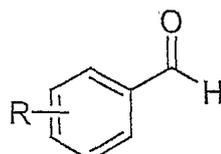
mit einer zweiten Verbindung der Formel I zu einer 2-
Hydroxyketongruppen enthaltenden Verbindung der allgemeinen Formel II
umgesetzt wird,



20

wobei R¹ und R² gleich oder verschieden sein können und R¹ und R²
dieselbe Bedeutung haben wie R in Formel I.

Eine Auswahl an Beispielen für Verbindungen der Formel I und damit für Substrate der erfindungsgemäß eingesetzten Benzaldehyd-Lyase sind nachfolgend aufgeführt:

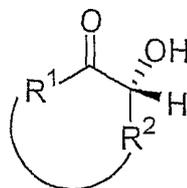


5

R = 2-F, 4-F, 2,4-F, 2-Br, 4-Br, 2-Cl, 4-Cl, 2-OCH₃, 3-OCH₃, 4-OCH₃, 2-OH, 2-CH₃

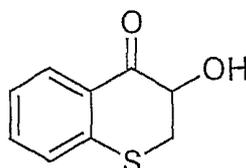
In einer bevorzugten Ausführungsvariante des erfindungsgemäßen Verfahrens wird als Verbindung der allgemeinen Formel I Benzaldehyd eingesetzt. Eine bevorzugte Verbindung der allgemeinen Formel II stellt das (R)-Benzoin dar.

In einer weiteren Ausführungsvariante der vorliegenden Erfindung ist die Synthese von zyklischen Verbindungen möglich. Beispielsweise können durch intramolekulare Umsetzungen zweier Aldehydgruppen erfindungsgemäß z. B. steroidale Hydroxyketone entstehen. Zur näheren Erläuterung dieser zyklischen Verbindungen dient nachfolgende allgemeine Formel, wobei R¹ und R² die gleiche Bedeutung haben, wie in Formel II bzw. Formel I.



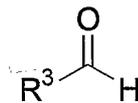
20

Als ein Beispiel für eine bevorzugte Verbindung dient die nachfolgende Formel, die jedoch lediglich zur Erläuterung dient und sich nicht limitierend auf die Erfindung auswirkt.



Ferner umfaßt die vorliegende Erfindung ein Verfahren bei dem Verbindungen der Formel II in Gegenwart von Aldehydgruppen enthaltenden Verbindungen der allgemeinen Formel III

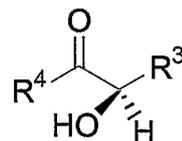
5



wobei $R^3 = -CH_3$ oder eine Komponente aus der Gruppe aliphatischer, aromatischer oder heteroaromatischer Kohlenwasserstoffe ist, die in ortho-, meta- oder para-Stellung einfach oder mehrfach substituiert sein kann, wobei die Substituenten Alkyl-, Aryl- oder Aralkyl-Gruppen oder Heteroatome, wie z.B. Cl, F, Br oder S, P, N oder Kombinationen davon sein können,

zu Verbindungen der Formel IV weiter umgesetzt werden

15



wobei $R^3 \neq R^4$ und $R^3 = R^3$ der Formel III und $R^4 = R$ der Formel I.

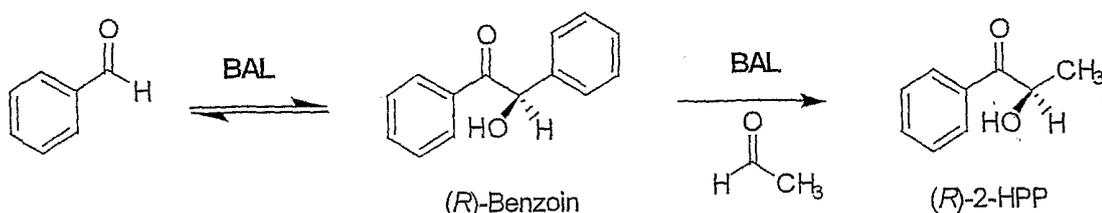
In einer bevorzugten Ausführungsvariante des vorliegenden Verfahrens wird als Verbindung der allgemeinen Formel III Acetaldehyd eingesetzt.

Eine bevorzugte Verbindung der allgemeinen Formel IV stellt (R)-2-Hydroxy-1-phenyl-propanon ((R)-2-HPP) dar.

In einer besonders bevorzugten Ausführungsvariante der vorliegenden Erfindung stellt das erfindungsgemäße Verfahren eine Kombination aus den beiden zuvor genannten Varianten dar, wobei in einem ersten

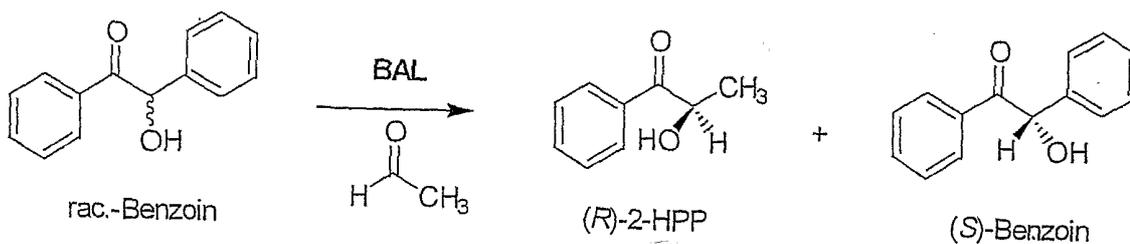
25

Abschnitt die Umsetzung von zwei Aldehydgruppen enthaltenden Verbindungen zu einer 2-Hydroxyketongruppen enthaltenden Verbindung erfolgt und sobald letztere vorliegt, auch die weitere Umsetzung dieser 2-Hydroxyketone zu einer weiteren 2-Hydroxyketongruppen enthaltenden
 5 Verbindung erfolgt. Eine Zusammenfassung des Reaktionsablaufs ist nachfolgend schematisch dargestellt:



Ferner umfaßt die vorliegende Erfindung auch ein Verfahren, wobei
 10 ausgehend von einem racemischen Gemisch enthaltend Verbindungen der Formel II selektiv das Enantiomer der Formel II unter Beibehaltung der stereochemischen Anordnung zu einer Verbindung der Formel IV weiter umgesetzt wird. Zur Verdeutlichung ist diese Variante nachfolgend schematisiert dargestellt.

15



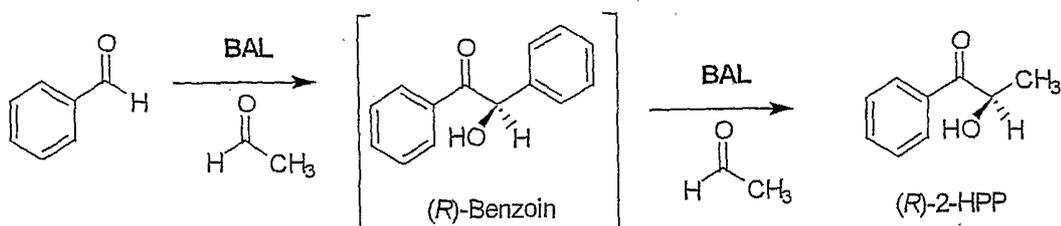
Hierbei soll die Anordnung der OH-Gruppe am α -C-Atom zur Ketogruppe über eine geschlängelte Linie das Vorliegen eines racemischen Gemisches einer 2-Hydroxyketongruppen enthaltenden Verbindung der
 20 Formel II darstellen. Hierbei kann es sich beispielsweise um ein racemisches Gemisch aus (S)- und (R)-Benzoin handeln.

Die vorliegende Erfindung zeichnet sich somit in vorteilhafter Weise dadurch aus, daß mit ein und demselben Katalysator sowohl eine hocheffiziente enantioselektive Herstellung von (R)-2-Hydroxyketonen möglich ist, als auch bei Vorliegen eines racemischen Gemisches von 2-Hydroxyketongruppen enthaltenden Verbindungen, z.B. der Formel II, eine hocheffiziente Trennung der beiden Enantiomere erfolgt. Beispielsweise kann durch das erfindungsgemäße Verfahren einerseits aus Benzaldehyd nahezu quantitativ (R)-Benzoin hergestellt werden und andererseits aus einem racemischen Gemisch von (S)- und (R)-Benzoin mit einer vergleichbar hohen Effizienz (S)-Benzoin und (R)-2-HPP gewonnen werden. Aufgrund des unterschiedlichen Löslichkeitsverhaltens der beiden zuletzt genannten 2-Hydroxyketongruppen enthaltenden Enantiomere ist eine quantitative Trennung der (S)- und (R)-Form des ursprünglich eingesetzten racemischen Gemisches möglich.

15

Eine weitere Variante des erfindungsgemäßen Verfahrens zeichnet sich dadurch aus, daß eine Aldehydgruppen enthaltende Verbindung der allgemeinen Formel I als ein erstes Substrat mit einem Acetaldehyd oder substituierten Acetaldehyd als ein zweites Substrat über eine Verbindung der Formel II als Zwischenstufe zu einer Verbindung der Formel IV umgesetzt wird. Die nachfolgende schematische Darstellung soll dies noch einmal verdeutlichen.

20



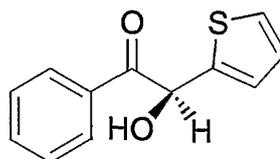
25 Der Begriff der Stereoselektivität bzw. Enantioselektivität im Sinne der vorliegenden Erfindung ist nachfolgend genauer erläutert. Generell ermöglicht das erfindungsgemäße Verfahren die stereoselektive

Herstellung von 2-Hydroxyketongruppen enthaltenden Verbindungen. Von den beiden möglichen Stereoisomeren (Enantiomeren) der gebildeten 2-Hydroxyketone, d.h. von dem (S)- oder (R)-Enantiomer, wird erfindungsgemäß aufgrund der eingesetzten Benzaldehyd-Lyase
5 ausschließlich das (R)-Enantiomer gebildet, d.h. mit einer Enantioselektivität von nahezu 100%. Dieses gebildete (R)-Enantiomer kann dann in einem weiteren Schritt des erfindungsgemäßen Verfahrens als ein weiteres Substrat der Benzaldehyd-Lyase fungieren und zu einer weiteren (R)-2-Hydroxyketongruppen enthaltenden Verbindung umgesetzt
10 werden.

Dabei ist unter einem (R)-Enantiomer erfindungsgemäß eine Verbindung zu verstehen, bei der die OH-Gruppe am α -C-Atom zur Ketogruppe aus der Papierebene hervortritt, während die Kohlenstoffatome des Carbonyl-
15 Grundgerüsts und die Reste R^1 und R^2 (in Formel II) bzw. R^3 und R^4 (in Formel IV) in der Papierebene liegen. Diese erfindungsgemäße Definition eines (R)-Enantiomers erfolgt dabei ohne Berücksichtigung der Priorität der Reste R^1/R^2 bzw. R^3/R^4 , sondern ausschließlich in Abhängigkeit von der Positionierung der Hydroxylgruppe am α -C-Atom zur Ketogruppe.
20 Folglich ist es möglich, daß die erfindungsgemäße (R)-Konfiguration nicht mit der gängigen Nomenklatur der stereoisomeren Verbindungen identisch ist.

Das zuvor gesagte ist im folgenden beispielhaft an einer Verbindung der Formel II erläutert:

25



Laut gängiger Nomenklatur handelt es sich um die Verbindung (S)-2-Hydroxy-1-phenyl-2-thiophen-2-yl-ethanon. Erfindungsgemäß liegt hier jedoch das (R)-Enantiomer vor, da die OH-Gruppe aus der Papierebene herausragt, während die beiden Kohlenstoffatome der Ketongruppe und die Reste R¹ und R² in der Papierebene liegen.

Unter dem Begriff „Beibehaltung der stereochemischen Anordnung“ ist erfindungsgemäß zu verstehen, daß die Konfiguration der OH-Gruppe am α -C-Atom zur Ketogruppe unverändert aus der enzymkatalysierten Reaktion hervorgeht, also erhalten bleibt.

Überraschender Weise verläuft das Enzym-katalysierte erfindungsgemäße Verfahren unter Einsatz einer lösungsvermittelnden Verbindung ausgehend von Aldehydgruppen enthaltenden Verbindungen der Formel I nahezu quantitativ in Richtung einer 2-Hydroxyketongruppen enthaltenden Verbindung der Formel II. Das zuvor Gesagte betrifft auch die Verfahrensvariante, bei der die in einem ersten Verfahrensabschnitt synthetisierten Verbindungen der Formel II weiter zu Verbindungen der Formel IV umgesetzt werden.

Erfindungsgemäß werden in einem der zuvor genannten erfindungsgemäßen Verfahren dem Reaktionsansatz oder dem Kulturmedium wenigstens 0,5-40%, bevorzugt 5-20%, besonders bevorzugt 7-12% und höchst bevorzugt 10% wenigstens einer lösungsvermittelnden Verbindung zugesetzt. Hierbei kann erfindungsgemäß als lösungsvermittelnde Verbindung ein mit Wasser mischbares und/oder wasserlösliches organisches Lösungsmittel und/oder eine lösungsmittelfreie Verbindung zugesetzt werden. Beispielhaft für mit Wasser mischbare und/oder wasserlösliche organische Lösungsmittel sind hier Dimethylsulfoxid (DMSO), Dimethylformamid oder Ethanol zu nennen. Als ein Beispiel für eine erfindungsgemäß einsetzbare lösungsmittelfreie

Verbindungen sind Cyclodextrine, beispielsweise β -Cyclodextrine (Cycloheptaamylose) zu nennen.

5 Ferner zeichnet sich das erfindungsgemäße Verfahren dadurch aus, daß ein Enantiomerenüberschuß (ee) an Verbindung der Formel II im Bereich von $\geq 96\%$, bevorzugt von $\geq 97\%$ besonders bevorzugt von $\geq 99\%$ und insbesondere von $99,9\%$ vorliegt.

10 Erfindungsgemäß ist das vorliegende Verfahren ferner dadurch gekennzeichnet, daß ein Enantiomerenüberschuß (ee) an Verbindung der Formel IV von wenigstens 50 bis 60%, bevorzugt 60 bis 90%, besonders bevorzugt von 92 bis 97% und höchst bevorzugt 97 bis $\geq 99,9\%$ erreicht wird.

15 Der Enantiomerenüberschuß (ee) ist das Maß für die Selektivität eines stereospezifischen Herstellungsverfahrens und kann durch folgende Formel dargestellt werden: $ee = \text{Betrag} (\%S - \%R) = (S-R) / (S+R)$. Hierunter ist der Betrag der Differenz der prozentualen Ausbeuten der beiden vorkommenden Isomere ((S)- und (R)-Enantiomer) im Produkt zu
20 verstehen. Liegt z. B. das eine Isomer zu 20% vor und das andere Isomer zu 80%, so ergibt sich ein Enantiomerenüberschuß von 60%.

In einer weiteren Variante der vorliegenden Erfindung zeichnet sich das erfindungsgemäße Verfahren dadurch aus, daß es in einer sogenannten
25 „fed-batch“ oder auch in kontinuierlicher Reaktionsführung durchgeführt wird. Geeignete Vorrichtungen hierzu sind zahlreich aus der Literatur bekannt. Beispielsweise eignet sich als Reaktionsbehälter ein klassischer Rührkessel (Fermenter), welcher mit einem geeigneten Medium zur Kultivierung von Zellen enthaltend eine erfindungsgemäße Benzaldehyd-
30 Lyase beschickt ist. Durch den kontinuierlichen Zufluß von Aldehydgruppen enthaltenden Verbindungen der allgemeinen Formel I

und gegebenenfalls kontinuierlicher oder diskontinuierlicher Zudosierung einer lösungsvermittelnden Verbindung kann das enzymatisch synthetisierte, stereospezifische Produkt einer (R)-2-Hydroxyketongruppen enthaltenden Verbindung kontinuierlich hergestellt und aus dem
5 Fermenter abgeführt werden.

In einer weiteren Verfahrensvariante der vorliegenden Erfindung erfolgt die Synthese von 2-Hydroxyketongruppen enthaltenden Verbindungen in einem Enzym-Membran-Reaktor, wie er beispielsweise bei Kula et al. (J.
10 Biotechnol., 1981, 23: 2789-2802) beschrieben ist.

Erfindungsgemäß können Produktausbeuten im Bereich von 96-100%, bevorzugt von 97%, besonders bevorzugt von 99% und insbesondere von mehr als 99% erreicht werden.

15 Gegenstand der vorliegenden Erfindung ist ferner ein Verfahren, das sich dadurch auszeichnet, daß eine Benzaldehyd-Lyase oder ein biologisch aktiver Teil davon kodiert durch eine Nukleotidsequenz gemäß SEQ ID No. 1 oder ein Allel, Homolog oder Derivat dieser Nukleotidsequenz oder eine mit dieser hybridisierende Nukleotidsequenz eingesetzt wird. Die
20 vorliegende Erfindung umfaßt dabei auch ein Verfahren, bei dem eine Benzaldehyd-Lyase mit einer Aminosäuresequenz gemäß SEQ ID No. 2 oder ein biologisch aktiver Teil davon, eine modifizierte Form, Isoenzyme oder Mischungen davon eingesetzt werden.

25 Unter Allelen sind erfindungsgemäß funktionell äquivalente, d. h. im wesentlichen gleichwirkende Nukleotidsequenzen zu verstehen. Funktionell äquivalente Sequenzen sind solche Sequenzen, welche trotz abweichender Nukleotidsequenz, beispielsweise durch die Degeneriertheit des genetischen Codes noch die gewünschten Funktionen besitzen.
30 Funktionelle Äquivalente umfassen somit natürlich vorkommende Varianten der hierin beschriebenen Sequenzen sowie künstliche, z. B.

durch chemische Synthese erhaltene und gegebenenfalls an den Kodon-Gebrauch des Wirtsorganismus angepaßte Nukleotid-Sequenzen. Darüber hinaus umfassen funktionell äquivalente Sequenzen solche, die eine veränderte Nukleotidsequenz aufweisen, welche dem Enzym
5 beispielsweise eine Desensitivität oder Resistenz gegenüber Inhibitoren verleiht. Funktionelle Äquivalente (Allele) sind auch solche Varianten, deren Funktion, verglichen mit dem Ausgangsgen bzw. Genfragment, abgeschwächt oder verstärkt ist.

10 Unter einem Allel versteht man insbesondere auch natürliche oder künstliche Mutationen einer ursprünglich isolierten Sequenz, welche weiterhin die gewünschte Funktion zeigen. Mutationen umfassen Substitutionen, Additionen, Deletionen, Vertauschungen oder Insertionen eines oder mehrerer Nukleotidreste. Inbegriffen sind hier auch sogenannte
15 Sinnmutationen (sense mutations), die auf Proteinebene beispielsweise zum Austausch konservierter Aminosäuren führen können, welche aber zu keiner grundsätzlichen Veränderung der Aktivität des Proteins führen und somit funktionsneutral sind. Dies beinhaltet auch Veränderungen der Nukleotidsequenz, die auf Proteinebene den N- oder C-Terminus eines
20 Proteins betreffen, ohne jedoch die Funktion des Proteins wesentlich zu beeinträchtigen. Diese Veränderungen können sogar stabilisierenden Einfluß auf die Proteinstruktur ausüben. Inbegriffen sind ferner Veränderungen, die beispielsweise eine spätere Proteinaufreinigung erleichtern, wie z.B. ein sogenannter aus mehreren aufeinander folgenden
25 Histidin-Resten bestehender „His-tag“.

Ferner werden beispielsweise auch solche Nukleotidsequenzen durch die vorliegende Erfindung mit umfaßt, welche man durch Modifikation der Nukleotidsequenz, resultierend in entsprechenden Derivaten, erhält. Ziel
30 einer solchen Modifikation kann z. B. die weitere Eingrenzung der darin enthaltenen kodierenden Sequenz oder auch die Einfügung weiterer

Erkennungsschnittstellen für Restriktionsenzyme oder der Einbau seltener Nukleotide sein.

Außerdem sind artifizielle DNA-Sequenzen Gegenstand der vorliegenden
5 Erfindung, solange sie, wie oben beschrieben, die gewünschten
Eigenschaften vermitteln. Solche artifiziellen DNA-Sequenzen können
beispielsweise durch Rückübersetzung von mittels computergestützten
Programmen (molecular modelling) erstellten Proteinen oder durch in-
vitro-Selektion ermittelt werden. Besonders geeignet sind kodierende
10 DNA-Sequenzen, die durch Rückübersetzung einer Polypeptidsequenz
gemäß der für den Wirtsorganismus spezifischen Kodon-Nutzung erhalten
wurden. Die spezifische Kodon-Nutzung kann ein mit
molekulargenetischen Methoden vertrauter Fachmann durch
Computerauswertungen anderer, bereits bekannter Gene des zu
15 transformierenden Organismus leicht ermitteln.

Unter homologen Sequenzen sind erfindungsgemäß solche zu verstehen,
die zu den erfindungsgemäßen Nukleotidsequenzen komplementär sind
und/oder mit diesen hybridisieren. Der Begriff hybridisierende Sequenzen
20 schließt erfindungsgemäß substanziell ähnliche Nukleotidsequenzen aus
der Gruppe von DNA oder RNA ein, die unter an sich bekannten
stringenten Bedingungen eine spezifische Wechselwirkung (Bindung) mit
den zuvor genannten Nukleotidsequenzen eingehen. Hierzu zählen auch
kurze Nukleotidsequenzen mit einer Länge von beispielsweise 10 bis 30,
25 bevorzugt 12 bis 15 Nukleotiden. Dies umfaßt erfindungsgemäß u. a. auch
sogenannte „Primer“ oder Sonden.

Erfindungsgemäß sind auch die den kodierenden Bereichen
(Strukturgenen) vorausgehenden (5'- oder upstream) und/oder
30 nachfolgenden (3'- oder downstream) Sequenzbereiche eingeschlossen.
Insbesondere sind hierin Sequenzbereiche mit regulatorischer Funktion

inbegriffen. Sie können die Transkription, die RNA-Stabilität oder das RNA processing sowie die Translation beeinflussen. Beispiele für regulatorische Sequenzen sind u. a. Promotoren, Enhancer, Operatoren, Terminatoren oder Translationsverstärker.

5

Unter einem biologisch aktiven Teil des erfindungsgemäßen Enzyms ist erfindungsgemäß jede Polypeptidsequenz zu verstehen, die über die erfindungswesentliche und spezifische Enzymaktivität zur Knüpfung von C-C-Verbindungen verfügt. Die Länge eines biologisch aktiven Teils einer erfindungsgemäßen Benzaldehyd-Lyase kann zum Beispiel im Bereich von 560 ± 10 Aminosäurereste bis 560 ± 250 Aminosäurereste, bevorzugt von 560 ± 50 bis 560 ± 100 und besonders bevorzugt von 560 ± 25 bis 560 ± 50 Aminosäurereste variieren. Dabei entspricht die „Basiszahl“ von 560 Aminosäureresten erfindungsgemäß einer Polypeptidsequenz einer Benzaldehyd-Lyase gemäß SEQ ID NO. 2 kodiert durch eine Nukleotidsequenz gemäß SEQ ID NO. 1. Folglich kann die „Basiszahl“ des Polypeptids in Abhängigkeit von der sie kodierenden Nukleotidsequenz ebenfalls variieren.

10

Unter modifizierten Formen sind erfindungsgemäß Enzyme zu verstehen, bei denen Änderungen in der Sequenz, beispielsweise am N- und/oder C-Terminus des Polypeptids oder im Bereich konservierter Aminosäuren vorliegen, ohne jedoch die Funktion des Enzyms zu beeinträchtigen. Diese Veränderungen können durch den Austausch einer oder mehrerer Aminosäuren nach an sich bekannten Methoden vorgenommen werden.

20

Eine besondere Ausführungsvariante der vorliegenden Erfindung umfaßt Varianten der erfindungsgemäßen Polypeptide, deren Aktivität und/oder Spezifität beispielsweise durch Aminosäureaustausch, verglichen mit dem jeweiligen Ausgangsprotein, abgeschwächt oder verstärkt ist. Gleiches gilt für die Stabilität der erfindungsgemäßen Enzyme in den Zellen, die

30

beispielsweise gegenüber dem Abbau durch Proteasen verstärkt oder vermindert anfällig sind.

5 Ferner sind Polypeptide mit der Funktion einer Benzaldehyd-Lyase Gegenstand der vorliegenden Erfindung, die in ihrer Aminosäuresequenz derart verändert sind, daß sie gegenüber regulatorisch wirkenden Verbindungen, beispielsweise Endprodukte des Stoffwechselweges, desensitiv sind (feedback-desensitiv).

10 Unter Isoformen sind Enzyme mit gleicher oder vergleichbarer Substrat- und Wirkungsspezifität zu verstehen, die jedoch eine unterschiedliche Primärstruktur aufweisen.

15 Erfindungsgemäß werden in dem vorliegenden Verfahren eine isolierte Benzaldehyd-Lyase der zuvor beschriebenen Art oder Zellextrakte oder ganze Zellen enthaltend eine Benzaldehyd-Lyase eingesetzt.

20 Erfindungsgemäß kann eine isolierte Benzaldehyd-Lyase und/oder Zellextrakt enthaltend eine erfindungsgemäße Benzaldehyd-Lyase z. B. in einem Enzym-Membran-Reaktor zum Einsatz kommen. Darüber hinaus sind auch andere in-vitro Systeme zur erfindungsgemäß stereoselektiven Herstellung von 2-Hydroxyketongruppen enthaltenden Verbindungen umfaßt.

25 Als geeignete Systeme für den Einsatz von ganzen Zellen enthaltend eine erfindungsgemäße Benzaldehyd-Lyase sind an dieser Stelle beispielhaft gängige Kultivierungsverfahren im batch-Betrieb, fed-batch-Betrieb oder kontinuierliche Fermentationen zu nennen.

30 Das erfindungsgemäße Verfahren zeichnet sich ferner dadurch aus, daß eine Benzaldehyd-Lyase aus Mikroorganismen, bevorzugt der Gattung

Pseudomonas oder Acinetobacter eingesetzt wird. Bevorzugt wird eine Benzaldehyd-Lyase aus *Pseudomonas fluorescens* und insbesondere aus *Pseudomonas fluorescens* Biovar I eingesetzt. Hierbei umfaßt die vorliegende Erfindung auch eine Benzaldehyd-Lyase aus sogenannten
5 Produktionsstämmen. Diese können entweder auf natürliche oder künstliche Weise erhalten werden. Letztere schließt klassische Mutageneseverfahren oder beispielsweise gentechnische Methoden ein.

Ebenfalls sind die 2-Hydroxyketongruppen enthaltenden Verbindungen
10 (kurz: 2-Hydroxyketone), hergestellt nach einem Verfahren der zuvor beschriebenen Art, Gegenstand der vorliegenden Erfindung .

Die vorliegende Erfindung umfaßt ferner eine Benzaldehyd-Lyase zum Einsatz in ein zuvor genanntes Verfahren, kodiert durch eine
15 Nukleotidsequenz gemäß SEQ ID NO. 1 oder ein Allel, Homolog oder Derivat dieser Nukleotidsequenz oder eine mit dieser hybridisierende Nukleotidsequenz. Dies schließt ebenfalls eine Benzaldehyd-Lyase mit einer Aminosäuresequenz gemäß SEQ ID NO. 2 oder ein biologisch aktiver Teil oder eine modifizierte Form davon oder Isoenzyme oder
20 Mischungen daraus ein.

Die erfindungsgemäße Benzaldehyd-Lyase kann aus Mikroorganismen, bevorzugt der Gattung *Pseudomonas* oder *Acinetobacter* isoliert werden. Bevorzugt handelt es sich erfindungsgemäß um eine Benzaldehyd-Lyase
25 isoliert aus der Spezies *Pseudomonas fluorescens*, besonders bevorzugt der Art *Pseudomonas fluorescens* Biovar I. Diese Ausführungen sind exemplarisch, jedoch keinesfalls limitierend für die vorliegende Erfindung. Ebenso sind durch die vorliegende Erfindung sogenannte Produktionsstämmen für die Isolierung einer entsprechenden Benzaldehyd-
30 Lyase umfaßt. Gegenstand der vorliegenden Erfindung ist ferner eine

Nukleotidsequenz kodierend für eine Benzaldehyd-Lyase isoliert aus Organismen der zuvor beschriebenen Art.

5 Unter einer Nukleotidsequenz oder einem Nukleinsäurefragment ist erfindungsgemäß ein Polymer aus RNA oder DNA zu verstehen, das einzel- oder doppelsträngig sein kann und optional natürliche, chemisch synthetisierte, modifizierte oder artifizielle Nukleotide enthalten kann. Der Begriff DNA-Polymer schließt hierbei auch genomische DNA, cDNA oder Mischungen davon ein.

10

Gegenstand der vorliegenden Erfindung ist ferner eine Genstruktur enthaltend eine zuvor beschriebene Nukleotidsequenz kodierend für eine erfindungsgemäße Benzaldehyd-Lyase sowie mit dieser operativ verknüpfte regulatorische Sequenzen, welche die Expression der
15 kodierenden Sequenzen in der Wirtszelle steuern.

Unter einer operativen Verknüpfung versteht man die sequenzielle Anordnung beispielsweise von Promotor, kodierender Sequenz, Terminator und ggf. weiterer regulatorischer Elemente derart, daß jedes
20 der regulatorischen Elemente seine Funktion bei der Expression der kodierenden Sequenz bestimmungsgemäß erfüllen kann. Diese regulatorischen Nukleotidsequenzen können natürlichen Ursprungs sein oder durch chemische Synthese erhalten werden. Als Promotor ist grundsätzlich jeder Promotor geeignet, der die Genexpression in dem
25 entsprechenden Wirtsorganismus steuern kann. Hierbei kann es sich erfindungsgemäß auch um einen chemisch induzierbaren Promotor handeln, durch den die Expression der ihm unterliegenden Gene in der Wirtszelle zu einem bestimmten Zeitpunkt gesteuert werden kann. Beispielfhaft sei hier ein durch IPTG (Isopropyl- β -thiogalactosid)
30 induzierbarer Promotor genannt.

Die Herstellung einer Genstruktur erfolgt durch Fusion eines geeigneten Promotors mit wenigstens einer erfindungsgemäßen Nukleotidsequenz nach gängigen Rekombinations- und Klonierungstechniken, wie sie beispielsweise in T. Maniatis, E.F. Fritsch und J. Sambrook, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1994) beschrieben sind.

Für die Verbindung der DNA-Fragmente miteinander können an die Fragmente Adaptoren oder Linker angesetzt werden.

Darüber hinaus betrifft die vorliegende Erfindung einen Vektor enthaltend eine Nukleotidsequenz der zuvor beschriebenen Art kodierend für eine Benzaldehyd-Lyase, mit diesen operativ verknüpfte regulative Nukleotidsequenzen sowie zusätzliche Nukleotidsequenzen zur Selektion transformierter Wirtszellen, für die Replikation innerhalb der Wirtszelle oder zur Integration in das entsprechende Wirtszell-Genom. Ferner kann der erfindungsgemäße Vektor eine Genstruktur der vorgenannten Art enthalten.

Die vorliegende Erfindung betrifft ferner einen transformierten ein- oder mehrzelligen Organismus zum Einsatz in ein Verfahren der zuvor genannten Art, welcher eine Benzaldehyd-Lyase und/oder eine Nukleotidsequenz der zuvor genannten Art enthält.

Die Übertragung von Nukleinsäuresequenzen in eine Wirtszelle erfolgt nach gängigen gentechnischen Methoden. Als bevorzugtes Verfahren sei hier die Transformation und besonders bevorzugt die Übertragung von DNA durch Elektroporation genannt.

Erfindungsgemäß zeichnet sich dieser transformierte ein- oder mehrzellige Organismus dadurch aus, daß er ein Mikroorganismus, eine Hefe, ein Pilz, eine tierische oder eine pflanzliche Zelle ist. Bevorzugt gehört der erfindungsgemäß transformierte Organismus zu der Gruppe der

Enterobakterien. Besonders bevorzugt handelt es sich um einen transformierten Organismus der Art *Escherichia coli*. In der vorliegenden Erfindung inbegriffen sind ferner auch sogenannte Produktionsstämme, die sich zur Herstellung der erfindungsgemäßen Verbindungen eignen.

5

Gegenstand der vorliegenden Erfindung ist ferner die Verwendung der erfindungsgemäßen 2-Hydroxyketongruppen enthaltenden Verbindungen zur Herstellung von Verbindungen mit antiviraler und/oder antifungaler Wirkung und/oder mit der Wirkung von Enzyminhibitoren, die in Bereichen
10 der Pharmazie und/oder Medizin, beispielsweise zur Behandlung von Immunkrankheiten (AIDS) oder neurodegenerativen Erkrankungen (Epilepsie), eingesetzt werden können.

Hierbei kann es sich beispielsweise um (chirale) Vorstufen oder Bestandteile von Antibiotika, wie z.B. Chloramphenicol handeln. Als
15 Beispiele für Verbindungen mit antifungaler Wirkung seien hier z. B. Genaconazole genannt (Gala D. et al., 1996, Tetrahedron Letters, 37 (5): 611-614 oder Leuenberger H. G. W. et al., 1999, Chimia, 53: 536-540). Diese Aufzählungen sind nur zur Erläuterung der Verwendung der erfindungsgemäß hergestellten 2-Hydroxyketone zu verstehen, die sich
20 jedoch keinesfalls limitierend auf die vorliegende Erfindung auswirkt.

Die vorliegende Erfindung wird durch die nachfolgenden Ausführungsbeispiele näher charakterisiert, die jedoch nicht limitierend
25 sind.

Allgemeine genetische Methoden:

Die Isolierung von genomischer DNA und Plasmid-DNA sowie alle Techniken zur Restriktion, Klenow- und alkalische
30 Phosphatasebehandlung, ferner Methoden zur Klonierung und Sequenzierung von DNA sowie die Transformation, Kultivierung und

Selektion von Wirtszellen wurden nach T. Maniatis, E.F. Fritsch und J. Sambrook (Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1994)) durchgeführt.

5 Proteinexpression und Reinigung

Die Isolierung des für die Benzaldehyd-Lyase kodierenden Gens erfolgte nach bekannten Methoden aus *Pseudomonas fluorescens*. Anschließend wurde das für die Benzaldehyd-Lyase kodierende Gen in den induzierbaren Vektor pKK322-2 (Pharmacia Biotech) kloniert und an
10 seinem 3'-Ende mit sechs für Histidin kodierenden Triplets versehen.

Der resultierende Vektor pkk322-2::BAL-His wurde dann in den *E. coli* Stamm SG13009prep4 (Quiagen, Hilden) transformiert. Die transformierten Zellen wurden in LB-Medium in Gegenwart von Ampicillin (100 mg/ml) bei 37°C vermehrt. Die Induktion der Proteinexpression
15 erfolgt bei einer erreichten optischen Dichte der Zellen von $OD_{600} = 0.7$ mit 1 mM Isopropyl- β -D-thiogalactosid (IPTG). Die Zellernte wurde nach weiteren 16 h mittels Zentrifugation durchgeführt. Der Zellaufschluss erfolgte je nach Maßstab des Kulturansatzes mechanisch mit Glasperlen oder in einer Kugelmühle. Der erhaltene Zell-Rohextrakt wurde durch
20 Zentrifugation und anschließende Filtration von Zelltrümmern befreit und auf eine Nickelchelatchromatographiesäule (bevorzugt: Ni-NTA-Agarose; Qiagen, Hilden) aufgegeben. Die Säule wurde zuvor mit 50 mM Kaliumphosphatpuffer, pH 7.0, equilibriert. Durch Waschen mit dem Equilibrierungspuffer werden nichtbindende Bestandteile ausgespült.
25 Anschließend wurde derselbe Puffer mit einem Zusatz von 50 mM Imidazol verwendet, um schwach gebundene Proteine zu eluieren. Die Benzaldehyd-Lyase mit ihrem Histidin-Ende (BAL-His) eluiert selektiv in Gegenwart von 200 mM Imidazol. Die gesammelten Proteinfractionen wurden anschließend durch eine Gelfiltration in 50 mM
30 Kaliumphosphatpuffer, pH 6.5, mit 1 mM $MgSO_4$ und 0.01 mM

Thiaminpyrophosphat (TPP) umgepuffert. Abschließend erfolgte eine Lyophilisation der Benzaldehyd-Lyase und Lagerung bei -20°C.

Synthese von (R)-Benzoin ausgehend von Benzaldehyd

5
318 mg (3 mM) Benzaldehyd werden in einer Mischung von 20 ml DMSO und 100 ml Phosphatpuffer (50 mM, pH 7,0) enthaltend 2,5 mM MgSO₄ und 0,15 mM TPP gelöst. Durch Zugabe von 1 mg (20 U) Benzaldehyd-Lyase wird die Umsetzung gestartet und der Reaktionsansatz für 48
10 Stunden bei Raumtemperatur inkubiert. Anschließend werden nochmals 1 mg (20 U) Benzaldehyd-Lyase zugegeben. Die Beobachtung des Reaktionsverlaufs erfolgt mittels kombinierter Gaschromatographie/Massenspektroskopie (GC-MS) oder Hochdruckflüssigkeitschromatographie (HPLC). Nach 62 Stunden wird eine 97%ige Umsetzung
15 zu (R)-Benzoin erreicht. Der Reaktionsansatz wird mit 250 ml Dichlormethan extrahiert, die organische Phase mit 25 ml dest. Wasser und 25 ml gesättigter NaCl-Lösung gewaschen und über Na₂SO₄ getrocknet. Anschließend wird das Lösungsmittel im Vakuum entfernt. Durch ein Umkristallisieren des Rohprodukts werden 305 mg (96%) (R)-
20 Benzoin erhalten (Schmelzpunkt 134°C). Der Enantiomerenüberschuß liegt bei einem Wert von ee > 99%.

$[\alpha]_D^{22} = -115$ (c = 1.5, CH₃COCH₃). HPLC (Chiralpac AD, *iso*-Hexan/*iso*-Propanol 90:10; 0.75 ml/min, 20°C, 21 bar) Rt: 26.95 min.

25 Werden unter vergleichbaren Bedingungen dem Reaktionsansatz statt 20 ml nur 10 ml DMSO zugesetzt, fällt ein Teil des Produkts (R)-Benzoin kristallin aus und kann durch Filtration separiert werden.

30 Gewinnung von (S)-Benzoin und (R)-2-HPP ausgehend von einem racemischen Benzoin-Gemisch

414 mg (2 mM) eines racemischen Benzoin-Gemisches werden in einer Mischung von 20 ml DMSO und 100 ml Phosphatpuffer (50 mM, pH7,0) enthaltend 2,5 mM MgSO₄ und 0,15 mM TPP gelöst. Diesem Ansatz werden 88 mg (2 mM) Acetaldehyd zugesetzt. Durch Zugabe von 1 mg
5 (20 U) Benzaldehyd-Lyase wird die Reaktion gestartet und bei Raumtemperatur inkubiert. Nach 24 Stunden werden nochmals 1 mg (20 U) Benzaldehyd-Lyase und weitere 176 mg (4 mM) Acetaldehyd zugefügt. Die Beobachtung des Reaktionsverlaufs erfolgt mittels kombinierter Gaschromatographie/ Massenspektroskopie (GC-MS) oder
10 Hochdruckflüssigkeitschromatographie (HPLC). Diese Vorgehensweise wird alle 24 Stunden wiederholt bis das vorhandene (R)-Benzoin vollständig umgesetzt ist. HPLC-Analysen ergeben, daß nach 4 Tagen ausschließlich die Produkte (S)-Benzoin und (R)-2-HPP in der Lösung vorliegen. Der Reaktionsansatz wird mit 250 ml Dichlormethan extrahiert,
15 die organische Phase mit 25 ml dest. Wasser und 25 ml gesättigter NaCl-Lösung gewaschen und über Na₂SO₄ getrocknet. Anschließend wird das Lösungsmittel im Vakuum entfernt. Die Aufreinigung des Rohprodukts erfolgt mittels Säulenchromatographie an Kieselgel (CH₂Cl₂). Es resultieren 279 mg (93%) (R)-2-HPP, wobei der Enantiomerenüberschuß
20 ee > 99% beträgt ($[\alpha]_D^{25} = -123$ (c = 2, CHCl₃)) und 201 mg (95%) (S)-Benzoin, mit einem Wert für den Enantiomerenüberschuß von ebenfalls > 99% ($[\alpha]_D^{25} = 114$ (c = 1.5, CH₃COCH₃)); (Schmelzpunkt = 134 °C).

(R)-2-HPP: ¹H NMR (300 MHz, CDCl₃, 20°C): δ = 1.47 (d, 3H, ³J(H,H) =
25 7.0 Hz; CH₃), 3.81 (br, 1H; OH), 5.19 (q, 1H, ³J(H,H) = 7.0 Hz; CHOH),
7.52 (t, 2H, ³J(H,H) = 7.5 Hz; ar-H), 7.64 (tt, 1H, ³J(H,H) = 7.5 Hz, ⁴J(H,H)
= 1.3 Hz; ar-H), 7.95 (dd, 2H, ³J(H,H) = 7.5 Hz, ⁴J(H,H) = 1.3 Hz; ar-H);
¹³C NMR (75.5 MHz, CDCl₃, 20°C): δ = 22.74 (CH₃), 69.73 (CHOH),
129.09, 129.31 (CH), 133.71 (C_q), 134.44 (CH), 202.82 (CO); GCMS: R_t =
30 7.70 min; m/z (%) = 150 (0.2) [M⁺], 135 (1.3) [M⁺-CH₃], 105 (100)

[C₇H₅O⁺], 77 (57) [C₆H₅⁺], 51 (17) [C₅H₃⁺]; HPLC: (Chiralpac AD, *iso*-Hexan/*iso*-Propanol 90:10; 0.75 ml/min, 20°C, 21 bar) Rt: 15.78 min.

(*S*)-Benzoin: ¹H-NMR (300 MHz, CDCl₃, 20°C, TMS): δ = 4.56 (br, 1H; OH), 5.95 (s, 1H; CHOH), 7.5 (m, 6H; ar-H), 7.9 (m, 4H; ar-H); GCMS: R_t = 11.29 min; *m/z* (%) = 212 (0.5) [M⁺], 105 (100) [C₇H₅O⁺], 77 (40) [C₆H₅⁺], 51 (9.8) [C₄H₃⁺]; HPLC(Chiralpac AD, *iso*-Hexan/*iso*-Propanol 90:10; 0.75 ml/min, 20°C, 21 bar) Rt: 33.36 min.

10 Synthese von (R)-2-HPP ausgehend von Benzaldehyd

318 mg (3 mM) Benzaldehyd werden in einer Mischung von 20 ml DMSO und 100 ml Kaliumphosphatpuffer (50 mM, pH7,0) enthaltend 2,5 mM MgSO₄ und 0,15 mM TPP gelöst. Durch Zugabe von 1 mg (20 U) Benzaldehyd-Lyase wird die Umsetzung gestartet und bei Raumtemperatur inkubiert. Diesem Reaktionsansatz werden 88 mg (2 mM) Acetaldehyd zugegeben. Nach Zugabe von weiteren 1 mg (20 U) Benzaldehyd-Lyase wird der Reaktionsansatz weiter bei Raumtemperatur inkubiert. Nach 24 Stunden werden weitere 1 mg Benzaldehyd-Lyase und 88 mg Acetaldehyd zugegeben. Die Beobachtung des Reaktionsverlaufs erfolgt mittels kombinierter Gaschromatographie/ Massenspektroskopie (GC-MS) oder Hochdruckflüssigkeitschromatographie (HPLC). Diese Vorgehensweise wird alle 24 Stunden wiederholt, bis das vorhandene (R)-Benzoin vollständig umgesetzt ist. HPLC-Analysen ergeben, daß nach insgesamt 62 Stunden ausschließlich das Produkt (R)-2-HPP vorliegt, wobei eine 97%ige Umsetzung erfolgt ist. Der Reaktionsansatz wird mit 250 ml Dichlormethan extrahiert, die organische Phase mit 25 ml dest. Wasser und 25 ml gesättigter NaCl-Lösung gewaschen und über Na₂SO₄ getrocknet. Anschließend wird das Lösungsmittel im Vakuum entfernt. Die Aufreinigung des Rohprodukts erfolgt mittels Säulenchromatographie an

Kieselgel (CH₂Cl₂). Es resultieren 214 mg (95%) (R)-2-HPP, wobei der Enantiomerenüberschuß ee > 99% beträgt ($[\alpha]_D^{22} = -122$ (c = 2, CHCl₃)).

5

Legende zu den Figuren:

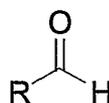
Fig. 1: Sequenzliste

10 enthaltend die Nukleotidsequenz kodierend für eine Benzaldehyd-
Lyase aus *Pseudomonas fluorescens* (SEQ ID NO. 1) und die
davon abgeleitete Aminosäuresequenz sowie eine separate
Darstellung der Aminosäuresequenz (SEQ ID NO. 2) der
Benzaldehyd-Lyase aus *Pseudomonas fluorescens*.

15

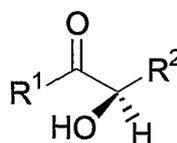
Patentansprüche:

1. Verfahren zur stereoselektiven Herstellung von 2-Hydroxyketonen, dadurch gekennzeichnet, daß eine Benzaldehyd-Lyase eingesetzt wird, die in Anwesenheit wenigstens einer lösungsvermittelnden Verbindung die Umsetzung von zwei Aldehydgruppen enthaltenden Verbindungen, von denen wenigstens eine kein Acetaldehyd ist, oder von einer Aldehydgruppen enthaltenden Verbindung mit einer 2-Hydroxyketongruppen enthaltenden Verbindung unter Beibehaltung der stereochemischen Anordnung der letzteren zu einer weiteren 2-Hydroxyketongruppen enthaltenden Verbindung katalysiert.
2. Verfahren gemäß Anspruch 1, dadurch gekennzeichnet, daß eine erste Aldehydgruppen enthaltende Verbindung der allgemeinen Formel I



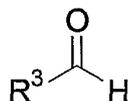
wobei R = eine Komponente aus der Gruppe aliphatischer, aromatischer oder heteroaromatischer Kohlenwasserstoffe ist, die in ortho-, meta- oder para-Stellung einfach oder mehrfach substituiert sein kann, wobei die Substituenten Alkyl-, Aryl- oder Aralkyl-Gruppen oder Heteroatome, wie z.B. Cl, F, Br oder S, N, P oder Kombinationen davon sein können und $\text{R} \neq -\text{CH}_3$ ist

mit einer zweiten Verbindung der Formel I zu einer 2-Hydroxyketongruppen enthaltenden Verbindung der allgemeinen Formel II umgesetzt wird,



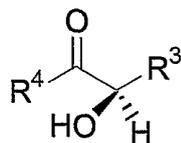
wobei R^1 und R^2 gleich oder verschieden sein können und R^1 und R^2 die dieselbe Bedeutung haben wie R in Formel I.

- 5 3. Verfahren gemäß einem der Ansprüche 1 oder 2, bei dem Verbindungen der Formel II in Gegenwart von Aldehydgruppen enthaltenden Verbindungen der allgemeinen Formel III



- 10 wobei $R^3 = -\text{CH}_3$ oder eine Komponente aus der Gruppe aliphatischer, aromatischer oder heteroaromatischer Kohlenwasserstoffe ist, die in ortho-, meta- oder para-Stellung einfach oder mehrfach substituiert sein kann, wobei die Substituenten Alkyl-, Aryl- oder Aralkyl-Gruppen oder Heteroatome, wie z.B. Cl, F, Br oder S, N, P oder Kombinationen
15 davon sein können

zu Verbindungen der Formel IV weiter umgesetzt werden



- 20 wobei $R^3 \neq R^4$ und $R^3 = R^3$ der Formel III und $R^4 = R$ der Formel I.

4. Verfahren gemäß einem der Ansprüche 1 bis 3, dadurch gekennzeichnet, daß ausgehend von einem racemischen Gemisch enthaltend Verbindungen der Formel II selektiv das Enantiomer der

Formel II unter Beibehaltung der stereochemischen Anordnung zu einer Verbindung der Formel IV weiter umgesetzt wird.

- 5 5. Verfahren gemäß einem der Ansprüche 1 bis 4, dadurch gekennzeichnet, daß eine Aldehydgruppen enthaltende Verbindung der allgemeinen Formel I als ein erstes Substrat mit einem Acetaldehyd oder substituierten Acetaldehyd als ein zweites Substrat über eine Verbindung der Formel II als Zwischenstufe zu einer Verbindung der Formel IV umgesetzt wird.
- 10 6. Verfahren gemäß einem der Ansprüche 1 bis 5, dadurch gekennzeichnet, daß dem Reaktionsansatz oder dem Kulturmedium wenigstens 0,5-40%, bevorzugt 5-20%, besonders bevorzugt 7-12% und höchst bevorzugt 10% wenigstens einer lösungsvermittelnden Verbindung zugesetzt werden.
- 15 7. Verfahren gemäß einem der Ansprüche 1 bis 6, dadurch gekennzeichnet, daß als lösungsvermittelnde Verbindung ein mit Wasser mischbares und/oder wasserlösliches organisches Lösungsmittel und/oder eine lösungsmittelfreie Verbindung zugesetzt wird.
- 20 8. Verfahren gemäß einem der Ansprüche 1 bis 7, dadurch gekennzeichnet, daß als lösungsvermittelnde Verbindung Dimethylsulfoxid, Dimethylformamid, Ethanol und/oder ein Cyclodextrin zugesetzt wird.
- 25 9. Verfahren gemäß einem der Ansprüche 1 bis 8, dadurch gekennzeichnet, daß ein Enantiomerenüberschuß an Verbindung der Formel II von wenigstens 96%, bevorzugt von $\geq 97\%$, besonders bevorzugt von $\geq 99\%$ und insbesondere von 99,9% erreicht wird.
- 30

10. Verfahren gemäß einem der Ansprüche 1 bis 9, dadurch gekennzeichnet, daß ein Enantiomerenüberschuß an Verbindung der Formel IV von wenigstens 50 bis 60%, bevorzugt 60 bis 90%,
5 besonders bevorzugt von 92 bis 97% und höchst bevorzugt 97 bis \geq 99,9% erreicht wird.
11. Verfahren gemäß einem der Ansprüche 1 bis 10, dadurch gekennzeichnet, daß es in einer kontinuierlichen Reaktionsführung
10 durchgeführt wird.
12. Verfahren gemäß einem der Ansprüche 1 bis 11, dadurch gekennzeichnet, daß eine Benzaldehyd-Lyase oder ein biologisch aktiver Teil davon kodiert durch eine Nukleotidsequenz gemäß SEQ ID
15 No. 1 oder ein Allel, Homolog oder Derivat dieser Nukleotidsequenz oder eine mit dieser hybridisierende Nukleotidsequenz eingesetzt wird.
13. Verfahren gemäß einem der Ansprüche 1 bis 12, dadurch gekennzeichnet, daß eine Benzaldehyd-Lyase mit einer
20 Aminosäuresequenz gemäß SEQ ID No. 2 oder ein biologisch aktiver Teil davon, eine modifizierte Form, Isoenzyme oder Mischungen davon eingesetzt werden.
14. Verfahren gemäß einem der Ansprüche 1 bis 13, dadurch gekennzeichnet, daß eine isolierte Benzaldehyd-Lyase oder
25 Zellextrakte oder ganze Zellen enthaltend eine Benzaldehyd-Lyase eingesetzt werden.
15. Verfahren gemäß einem der Ansprüche 1 bis 14, dadurch gekennzeichnet, daß eine Benzaldehyd-Lyase aus Mikroorganismen,
30

bevorzugt der Gattung *Pseudomonas* oder *Acinetobacter* eingesetzt wird.

5 16. Verfahren gemäß einem der Ansprüche 1 bis 15, dadurch gekennzeichnet, daß eine Benzaldehyd-Lyase aus *Pseudomonas fluorescens* eingesetzt wird.

10 17. 2-Hydroxyketone hergestellt nach einem Verfahren gemäß einem der Ansprüche 1 bis 16.

18. Benzaldehyd-Lyase zum Einsatz in ein Verfahren gemäß einem der Ansprüche 1 bis 16 kodiert durch eine Nukleotidsequenz gemäß SEQ ID No. 1 oder ein Allel, Homolog oder Derivat dieser Nukleotidsequenz oder eine mit dieser hybridisierende Nukleotidsequenz.

15 19. Benzaldehyd-Lyase gemäß Anspruch 18 mit einer Aminosäuresequenz gemäß SEQ ID No. 2 oder ein biologisch aktiver Teil oder eine modifizierte Form davon oder Isoenzyme oder Mischungen daraus.

20 20. Benzaldehyd-Lyase gemäß einem der Ansprüche 18 oder 19 isoliert aus Mikroorganismen, bevorzugt der Gattung *Pseudomonas* oder *Acinetobacter*.

25 21. Benzaldehyd-Lyase gemäß einem der Ansprüche 18 bis 20 isoliert aus der Spezies *Pseudomonas fluorescens*.

22. Nukleotidsequenz kodierend für eine Benzaldehyd-Lyase gemäß einem der Ansprüche 19 bis 21.

30

23. Genstruktur enthaltend eine Nukleotidsequenz gemäß SEQ ID No. 1 oder ein Allel, Homolog oder Derivat dieser Nukleotidsequenz oder eine mit dieser hybridisierende Nukleotidsequenz kodierend für eine erfindungsgemäße Benzaldehyd-Lyase sowie mit dieser operativ verknüpfte regulatorische Sequenzen, welche die Expression der kodierenden Sequenz in der Wirtszelle steuern.
24. Vektor enthaltend eine Nukleotidsequenz gemäß SEQ ID No. 1 oder ein Allel, Homolog oder Derivat dieser Nukleotidsequenz oder eine mit dieser hybridisierende Nukleotidsequenz oder eine Genstruktur gemäß Anspruch 23, mit diesen Nukleotidsequenzen operativ verknüpfte regulative Nukleotidsequenzen sowie zusätzliche Nukleotidsequenzen zur Selektion von Wirtszellen, für die Replikation in Wirtszellen und/oder zur Integration in das Genom von Wirtszellen.
25. Transformierter ein- oder mehrzelliger Organismus zum Einsatz in ein Verfahren gemäß einem der Ansprüche 1 bis 16 enthaltend eine Benzaldehyd-Lyase gemäß einem der Ansprüche 18 bis 21 und/oder eine Nukleotidsequenz gemäß Anspruch 22 und/oder ein Genkonstrukt gemäß Anspruch 23 und/oder einen Vektor gemäß Anspruch 24.
26. Transformierter ein- oder mehrzelliger Organismus gemäß Anspruch 25, dadurch gekennzeichnet, daß er ein Mikroorganismus, eine Hefe, ein Pilz, eine tierische oder eine pflanzliche Zelle ist.
27. Transformierter ein- oder mehrzelliger Organismus gemäß einem der Ansprüche 25 oder 26, dadurch gekennzeichnet, daß er zu der Gruppe der Enterobacterien gehört.

28. Transformierter ein- oder mehrzelliger Organismus gemäß einem der Ansprüche 25 bis 27, dadurch gekennzeichnet, daß er zu der Gattung *Escherichia*, bevorzugt der Art *Escherichia coli* gehört.

- 5 29. Verwendung von 2-Hydroxyketonen gemäß Anspruch 17 zur Herstellung von Verbindungen mit antiviraler und/oder antifungaler Wirkung und/oder mit der Wirkung von Enzyminhibitoren, die in Bereichen der Pharmazie und/oder Medizin, beispielsweise zur Behandlung von Immunkrankheiten oder neurodegenerativen
- 10 Erkrankungen verwendet werden.

SEQUENCE LISTING

<110> Forschungszentrum Juelich GmbH

<120> Nukleotidsequenz kodierend fuer eine Benzaldehyd-Lyase
und Verfahren zur stereoselektiven Synthese von
2-Hydroxyketonen

<130> 1

<140>

<141>

<160> 2

<170> PatentIn Ver. 2.1

<210> 1

<211> 1689

<212> DNA

<213> Pseudomonas fluorescens

<220>

<221> CDS

<222> (1)..(1689)

<400> 1

```

atg gcg atg att aca ggc ggc gaa ctg gtt gtt cgc acc cta ata aag 48
Met Ala Met Ile Thr Gly Gly Glu Leu Val Val Arg Thr Leu Ile Lys
  1                               5                               10                               15

gct ggg gtc gaa cat ctg ttc ggc ctg cac ggc gcg cat atc gat acg 96
Ala Gly Val Glu His Leu Phe Gly Leu His Gly Ala His Ile Asp Thr
                               20                               25                               30

att ttt caa gcc tgt ctc gat cat gat gtg ccg atc atc gac acc cgc 144
Ile Phe Gln Ala Cys Leu Asp His Asp Val Pro Ile Ile Asp Thr Arg
                               35                               40                               45

cat gag gcc gcc gca ggg cat gcg gcc gag ggc tat gcc cgc gct ggc 192
His Glu Ala Ala Ala Gly His Ala Ala Glu Gly Tyr Ala Arg Ala Gly
                               50                               55                               60

gcc aag ctg ggc gtg gcg ctg gtc acg gcg ggc ggg gga ttt acc aat 240
Ala Lys Leu Gly Val Ala Leu Val Thr Ala Gly Gly Gly Phe Thr Asn
  65                               70                               75                               80

gcg gtc acg ccc att gcc aac gct tgg ctg gat cgc acg ccg gtg ctc 288
Ala Val Thr Pro Ile Ala Asn Ala Trp Leu Asp Arg Thr Pro Val Leu
                               85                               90                               95

ttc ctc acc gga tcg ggc gcg ctg cgt gat gat gaa acc aac acg ttg 336
Phe Leu Thr Gly Ser Gly Ala Leu Arg Asp Asp Glu Thr Asn Thr Leu
                               100                              105                              110

cag gcg ggg att gat cag gtc gcc atg gcg gcg ccc att acc aaa tgg 384
Gln Ala Gly Ile Asp Gln Val Ala Met Ala Ala Pro Ile Thr Lys Trp
                               115                              120                              125

gcg cat cgg gtg atg gca acc gag cat atc cca cgg ctg gtg atg cag 432
Ala His Arg Val Met Ala Thr Glu His Ile Pro Arg Leu Val Met Gln
  130                              135                              140
    
```

gcg atc cgc gcc gcg ttg agc gcg cca cgc ggg ccg gtg ttg ctg gat	480
Ala Ile Arg Ala Ala Leu Ser Ala Pro Arg Gly Pro Val Leu Leu Asp	
145 150 155 160	
ctg ccg tgg gat att ctg atg aac cag att gat gag gat agc gtc att	528
Leu Pro Trp Asp Ile Leu Met Asn Gln Ile Asp Glu Asp Ser Val Ile	
165 170 175	
atc ccc gat ctg gtc ttg tcc gcg cat ggg gcc aga ccc gac cct gcc	576
Ile Pro Asp Leu Val Leu Ser Ala His Gly Ala Arg Pro Asp Pro Ala	
180 185 190	
gat ctg gat cag gct ctc gcg ctt ttg cgc aag gcg gag cgg ccg gtc	624
Asp Leu Asp Gln Ala Leu Ala Leu Leu Arg Lys Ala Glu Arg Pro Val	
195 200 205	
atc gtg ctc ggc tca gaa gcc tcg cgg aca gcg cgc aag acg gcg ctt	672
Ile Val Leu Gly Ser Glu Ala Ser Arg Thr Ala Arg Lys Thr Ala Leu	
210 215 220	
agc gcc ttc gtg gcg gcg act ggc gtg ccg gtg ttt gcc gat tat gaa	720
Ser Ala Phe Val Ala Ala Thr Gly Val Pro Val Phe Ala Asp Tyr Glu	
225 230 235 240	
ggg cta agc atg ctc tcg ggg ctg ccc gat gct atg cgg ggc ggg ctg	768
Gly Leu Ser Met Leu Ser Gly Leu Pro Asp Ala Met Arg Gly Gly Leu	
245 250 255	
gtg caa aac ctc tat tct ttt gcc aaa gcc gat gcc gcg cca gat ctc	816
Val Gln Asn Leu Tyr Ser Phe Ala Lys Ala Asp Ala Ala Pro Asp Leu	
260 265 270	
gtg ctg atg ctg ggg gcg cgc ttt ggc ctt aac acc ggg cat gga tct	864
Val Leu Met Leu Gly Ala Arg Phe Gly Leu Asn Thr Gly His Gly Ser	
275 280 285	
ggg cag ttg atc ccc cat agc gcg cag gtc att cag gtc gac cct gat	912
Gly Gln Leu Ile Pro His Ser Ala Gln Val Ile Gln Val Asp Pro Asp	
290 295 300	
gcc tgc gag ctg gga cgc ctg cag ggc atc gct ctg gcc att gtg gcc	960
Ala Cys Glu Leu Gly Arg Leu Gln Gly Ile Ala Leu Gly Ile Val Ala	
305 310 315 320	
gat gtg ggt ggg acc atc gag gct ttg gcg cag gcc acc gcg caa gat	1008
Asp Val Gly Gly Thr Ile Glu Ala Leu Ala Gln Ala Thr Ala Gln Asp	
325 330 335	
gcg gct tgg ccg gat cgc ggc gac tgg tgc gcc aaa gtg acg gat ctg	1056
Ala Ala Trp Pro Asp Arg Gly Asp Trp Cys Ala Lys Val Thr Asp Leu	
340 345 350	
gcg caa gag cgc tat gcc agc atc gct gcg aaa tcg agc agc gag cat	1104
Ala Gln Glu Arg Tyr Ala Ser Ile Ala Ala Lys Ser Ser Ser Glu His	
355 360 365	
gcg ctc cac ccc ttt cac gcc tcg cag gtc att gcc aaa cac gtc gat	1152
Ala Leu His Pro Phe His Ala Ser Gln Val Ile Ala Lys His Val Asp	
370 375 380	
gca ggg gtg acg gtg gta gcg gat ggt gcg ctg acc tat ctc tgg ctg	1200
Ala Gly Val Thr Val Val Ala Asp Gly Ala Leu Thr Tyr Leu Trp Leu	
385 390 395 400	

tcc gaa gtg atg agc cgc gtg aaa ccc ggc ggt ttt ctc tgc cac ggc 1248
 Ser Glu Val Met Ser Arg Val Lys Pro Gly Gly Phe Leu Cys His Gly
 405 410 415

tat cta ggc tcg atg ggc gtg ggc ttc ggc acg gcg ctg ggc gcg caa 1296
 Tyr Leu Gly Ser Met Gly Val Gly Phe Gly Thr Ala Leu Gly Ala Gln
 420 425 430

gtg gcc gat ctt gaa gca ggc cgc cgc acg atc ctt gtg acc ggc gat 1344
 Val Ala Asp Leu Glu Ala Gly Arg Arg Thr Ile Leu Val Thr Gly Asp
 435 440 445

ggc tcg gtg ggc tat agc atc ggt gaa ttt gat acg ctg gtg cgc aaa 1392
 Gly Ser Val Gly Tyr Ser Ile Gly Glu Phe Asp Thr Leu Val Arg Lys
 450 455 460

caa ttg ccg ctg atc gtc atc atc atg aac aac caa agc tgg ggg gcg 1440
 Gln Leu Pro Leu Ile Val Ile Ile Met Asn Asn Gln Ser Trp Gly Ala
 465 470 475 480

aca ttg cat ttc cag caa ttg gcc gtc ggc ccc aat cgc gtg acg ggc 1488
 Thr Leu His Phe Gln Gln Leu Ala Val Gly Pro Asn Arg Val Thr Gly
 485 490 495

acc cgt ttg gaa aat ggc tcc tat cac ggg gtg gcc gcc gcc ttt ggc 1536
 Thr Arg Leu Glu Asn Gly Ser Tyr His Gly Val Ala Ala Ala Phe Gly
 500 505 510

gcg gat ggc tat cat gtc gac agt gtg gag agc ttt tct gcg gct ctg 1584
 Ala Asp Gly Tyr His Val Asp Ser Val Glu Ser Phe Ser Ala Ala Leu
 515 520 525

gcc caa gcg ctc gcc cat aat cgc ccc gcc tgc atc aat gtc gcg gtc 1632
 Ala Gln Ala Leu Ala His Asn Arg Pro Ala Cys Ile Asn Val Ala Val
 530 535 540

gcg ctc gat ccg atc ccg ccc gaa gaa ctc att ctg atc ggc atg gac 1680
 Ala Leu Asp Pro Ile Pro Pro Glu Glu Leu Ile Leu Ile Gly Met Asp
 545 550 555 560

ccc ttc taa 1689
 Pro Phe

<210> 2
 <211> 562
 <212> PRT
 <213> Pseudomonas fluorescens

<400> 2
 Met Ala Met Ile Thr Gly Gly Glu Leu Val Val Arg Thr Leu Ile Lys
 1 5 10 15
 Ala Gly Val Glu His Leu Phe Gly Leu His Gly Ala His Ile Asp Thr
 20 25 30
 Ile Phe Gln Ala Cys Leu Asp His Asp Val Pro Ile Ile Asp Thr Arg
 35 40 45
 His Glu Ala Ala Ala Gly His Ala Ala Glu Gly Tyr Ala Arg Ala Gly
 50 55 60
 Ala Lys Leu Gly Val Ala Leu Val Thr Ala Gly Gly Gly Phe Thr Asn
 65 70 75 80
 Ala Val Thr Pro Ile Ala Asn Ala Trp Leu Asp Arg Thr Pro Val Leu

